

# Schreibe Dein Programm!

Einführung in die Programmierung

Michael Sperber

Herbert Klaeren

2. Februar 2022

Copyright © 2007-2021 by Michael Sperber and Herbert Klaeren

Dieses Buch ist lizenziert unter der Creative-Commons-Lizenz *Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International (CC BY-SA 4.0)*, nachzulesen unter <https://creativecommons.org/licenses/by-sa/4.0/deed.de>.

# INHALTSVERZEICHNIS

<b>o</b>	<b>Wie Du das Meiste aus diesem Buch herausholst</b>	<b>3</b>
o.1	Vorkenntnisse . . . . .	3
o.2	Anrede . . . . .	3
o.3	Software . . . . .	3
o.4	Programmiersprache . . . . .	4
o.5	Englisch . . . . .	4
o.6	Exkurse . . . . .	4
o.7	Aufgaben . . . . .	5
o.8	Bürokratische Didaktik . . . . .	5
o.9	Zusatzmaterial . . . . .	6
<b>1</b>	<b>Elemente des Programmierens</b>	<b>7</b>
1.1	DrRacket . . . . .	7
1.2	Ausdrücke und die REPL . . . . .	8
1.3	Das Definitionsfenster . . . . .	11
1.4	Rechnen ohne Zahlen . . . . .	13
1.5	Abstraktion und Applikation . . . . .	17
1.6	Information und Daten . . . . .	23
1.7	Programme systematisch konstruieren . . . . .	24
1.8	Kurzbeschreibung . . . . .	25
1.9	Signatur . . . . .	26
1.10	Tests . . . . .	27
1.11	Konstruktionsanleitungen . . . . .	30
1.12	Die Macht der Abstraktion . . . . .	34
1.13	Mantras für die Programmierung . . . . .	36
<b>2</b>	<b>Fallunterscheidungen und Verzweigungen</b>	<b>43</b>
2.1	Rechnen mit booleschen Werten . . . . .	43
2.2	Verzweigungen . . . . .	45
2.3	Abdeckung . . . . .	48
2.4	Aufzählungen . . . . .	49
2.5	Zahlenbereiche . . . . .	51
2.6	Datenanalyse und Schablonen . . . . .	65

---

2.7	Exkurs: Verzweigungen in der Mathematik . . . . .	69
2.8	Boolesche Fallunterscheidungen und binäre Verzweigungen . . . . .	70
2.9	Syntaktischer Zucker . . . . .	72
2.10	Unsinnige Daten abfangen . . . . .	73
<b>3</b>	<b>Zusammengesetzte Daten</b>	<b>79</b>
3.1	Computer konfigurieren . . . . .	79
3.2	Zusammengesetzte Daten selbst konstruieren . . . . .	86
3.3	Konstruktionsanleitungen für zusammengesetzte Daten . . . . .	92
3.4	Ein- und Ausgabe zusammengesetzter Daten . . . . .	94
3.5	Alternativen bei den Konstruktionsanleitungen . . . . .	99
<b>4</b>	<b>Gemischte Daten</b>	<b>109</b>
4.1	Gemischte Daten . . . . .	109
4.2	Gemischte Daten als Ausgabe . . . . .	115
4.3	Die Lebensmittelampel . . . . .	117
4.4	Fehler repräsentieren . . . . .	124
<b>5</b>	<b>Programmieren mit Selbstbezügen und Kombinatoren</b>	<b>131</b>
5.1	Flüsse abbilden . . . . .	131
5.2	Bilder modellieren . . . . .	139
5.3	Finanzverträge abbilden . . . . .	140
<b>6</b>	<b>Programmieren mit Listen</b>	<b>165</b>
6.1	Listen repräsentieren . . . . .	165
6.2	Listen mit anderen Elementen . . . . .	172
6.3	Konstruktionsanleitung . . . . .	175
6.4	Eingebaute Listen . . . . .	176
6.5	Parametrische Polymorphie . . . . .	177
6.6	Funktionen, die Listen produzieren . . . . .	178
6.7	Minimum einer Liste und lokale Variablen . . . . .	185
6.8	Funktionen auf nichtleeren Listen . . . . .	189
<b>7</b>	<b>Natürliche Zahlen</b>	<b>199</b>
7.1	Zahlen, die zählen . . . . .	199
7.2	Natürliche Zahlen und Listen . . . . .	202
7.3	Andersrum zählen . . . . .	205

7.4	Exkurs: Potenz optimieren . . . . .	208
<b>8</b>	<b>Exkurs: Induktive Beweise und Definitionen</b>	<b>213</b>
8.1	Aussagen über natürliche Zahlen . . . . .	213
8.2	Induktive Beweise führen . . . . .	217
8.3	Struktur der natürlichen Zahlen . . . . .	219
8.4	Endliche Folgen . . . . .	221
8.5	Notation für induktive Definitionen . . . . .	225
8.6	Strukturelle Rekursion . . . . .	226
8.7	Strukturelle Induktion . . . . .	229
<b>9</b>	<b>Higher-Order-Programmierung</b>	<b>237</b>
9.1	Fußball-Fakten ermitteln . . . . .	237
9.2	Higher-Order-Funktionen auf Listen . . . . .	244
9.3	Listen umwandeln . . . . .	251
9.4	Listen zusammenfalten . . . . .	257
9.5	Schönfinkeln . . . . .	263
9.6	Lexikalische Bindung . . . . .	267
<b>10</b>	<b>Programmieren mit Akkumulatoren</b>	<b>283</b>
10.1	Zwischenergebnisse mitführen . . . . .	283
10.2	Schablonen für Funktionen mit Akkumulator . . . . .	289
10.3	Über natürliche Zahlen akkumulieren . . . . .	294
10.4	Aktienkurse analysieren . . . . .	297
10.5	Kontext und Endrekursion . . . . .	301
10.6	Das Phänomen der umgedrehten Liste . . . . .	303
10.7	Listen iterativ zusammenfalten . . . . .	305
<b>II</b>	<b>Videospiele programmieren</b>	<b>311</b>
II.1	Bilder mit <code>image.rkt</code> . . . . .	311
II.2	Modelle und Ansichten . . . . .	320
II.3	Den Highway modellieren . . . . .	324
II.4	Straßenansichten . . . . .	326
II.5	Bewegte Bilder mit <code>universe.rkt</code> . . . . .	330
II.6	Autos und Gürteltiere auf dem Highway . . . . .	331
II.7	Die Welt ein Highway . . . . .	337
II.8	Überfahren konkret . . . . .	340

---

11.9	Reaktive Animation . . . . .	345
<b>12</b>	<b>Bäume</b>	<b>353</b>
12.1	Stammbäume . . . . .	353
12.2	Binärbäume . . . . .	357
12.3	Bäume für's Suchen . . . . .	363
12.4	Sortierte Bäume herstellen . . . . .	368
12.5	Suchbäume . . . . .	371
12.6	Sortierte Bäume sind effizienter als Listen . . . . .	375
12.7	Suchbäume balancieren . . . . .	377
12.8	Ausbalancieren durch Rotation . . . . .	388
<b>13</b>	<b>Eigenschaften von Funktionen</b>	<b>397</b>
13.1	Korrektheit und Tests . . . . .	397
13.2	Wie <code>check-property</code> funktioniert . . . . .	400
13.3	Mehr Eigenschaften und inexakte Zahlen . . . . .	401
13.4	Relationale Probleme . . . . .	405
13.5	Konstruktionsanleitungen für Testfälle? . . . . .	412
13.6	Suchbäume testen . . . . .	413
13.7	Algebraische Eigenschaften . . . . .	423
13.8	Eigenschaften von Funktionen auf Listen . . . . .	435
13.9	Eigenschaften von Funktionen höherer Ordnung . . . . .	443
13.10	Programme beweisen . . . . .	445
13.11	Rekursive Programme beweisen . . . . .	446
13.12	Invarianten . . . . .	451
13.13	Korrektheit von Suchbäumen beweisen . . . . .	454
<b>14</b>	<b>Fortgeschrittenes Programmieren mit Rekursion</b>	<b>461</b>
14.1	Lastwagen optimal beladen . . . . .	461
<b>15</b>	<b>Schrittweise Verfeinerung</b>	<b>469</b>
15.1	Löschen in Suchbäumen . . . . .	469
15.2	Datenverfeinerung . . . . .	478
<b>16</b>	<b>Der <math>\lambda</math>-Kalkül</b>	<b>489</b>
16.1	Sprache und Reduktionssemantik . . . . .	489
16.2	Normalformen . . . . .	495

---

16.3	Der $\lambda$ -Kalkül als Programmiersprache . . . . .	497
16.4	Auswertungsstrategien . . . . .	503
16.5	Die Auswertungsstrategie der Lehrsprachen . . . . .	506
<b>17</b>	<b>Die SECD-Maschine</b>	<b>509</b>
17.1	Der angewandte $\lambda$ -Kalkül . . . . .	509
17.2	Die einfache SECD-Maschine . . . . .	511
17.3	Quote und Symbole . . . . .	516
17.4	Implementierung der SECD-Maschine . . . . .	520
17.5	Die endrekursive SECD-Maschine . . . . .	537
17.6	Der $\lambda$ -Kalkül mit Zustand . . . . .	539
17.7	Die SECDH-Maschine . . . . .	542
17.8	Implementierung der SECDH-Maschine . . . . .	544
<b>A</b>	<b>Mathematische Grundlagen</b>	<b>557</b>
A.1	Aussagenlogik . . . . .	557
A.2	Mengen . . . . .	559
A.3	Prädikatenlogik . . . . .	562
A.4	Multimengen . . . . .	563
A.5	Relationen und Abbildungen . . . . .	564
A.6	Ordnungen . . . . .	566





## VORWORT

Schriftsteller fürchten die leere Seite. Maler fürchten die weiße Leinwand. Für Architekten ist es die grüne Wiese. Für Programmierer ist es der Cursor, der fordernd im leeren Fenster des Editors blinkt... Wie geht man eine Programmieraufgabe an, wenn zu Beginn keine einzige Zeile Code existiert, wenn keine Funktion oder Datenstruktur Orientierung bietet, und wenn zwar *alles* möglich aber gerade deshalb schwer greifbar erscheint?

Mike Sperber und Herbert Klaeren haben in *Schreibe Dein Programm!* hierzu gute Nachrichten. Programme lassen sich nämlich sehr systematisch konstruieren und—teilweise nahezu automatisch—aus der gegebenen Problemstellung ableiten. Über das gesamte Buch verstreut finden sich *Konstruktionsanleitungen*, deren Programmskelette konkrete Hinweise auf die Struktur der Lösung des Problems geben. Steht das Skelett erst einmal, lassen sich dessen Lücken oft unmittelbar vervollständigen, teils durch die Nutzung weiterer derartiger Konstruktionen, teils weil sich die Komplettierung direkt aus dem Problem ergibt. Programmierung mittels dieser Konstruktionsanleitungen kommt dem disziplinierten ingenieurmäßigen Bauen näher als alles andere.

Aus meinen Vorlesungen *Informatik I* an der Universität Tübingen kommen die Studierenden nicht heraus, ohne sich ordentlich «die Finger schmutzig gemacht zu haben». In der Vorlesung selbst, den wöchentlichen Übungsblättern und zusätzlichen betreuten Programmierstunden wird Code geschrieben, was das Zeug hält. Dabei konnte ich bereits Hunderte von Studierenden—typischerweise echte Anfänger, die noch keine Programmiersprache kennen—beobachten, wie mittels Konstruktionsanleitungen die Herausforderung des leeren Editors angenommen und gemeistert wurde. Die resultierenden Programme sind (zumeist ; -)) korrekt, kompakt und, vor allem, lesbar und nachvollziehbar. Wir können so in der *Informatik I* erstaunlich anspruchsvolle Aufgaben angehen und sind am Ende des zweiten Drittels der Vorlesung in der Lage, die Studierenden *PacMan*, *Tetris* oder *Asteroids* nachbauen zu lassen.

*Schreibe Dein Programm!* ist gleichzeitig eine Einführung in die funktionale Programmierung mit Scheme. Scheme selbst tritt mit seiner einfachen Syntax in den Hintergrund<sup>1</sup> und gibt den Blick frei auf die essentiellen Ideen des Denkens und Programmierens mit (mathematischen) Funktionen. Genau wie Mike Sperber und Herbert Klaeren bin ich überzeugt, dass dieses elegante Programmierparadigma «gekommen ist, um zu bleiben» und auf lange Sicht Programmierstil und -sprachen weitreichend beeinflussen wird. Zusammen mit den Konstruktionsanleitungen gibt die-

---

<sup>1</sup> Wo wir üblicherweise einen Zoo von Notationen  $x + 1$ ,  $\sin(x)$ ,  $x^2$ ,  $|x|$ ,  $\sqrt{x}$ ,  $x!$ , ... nutzen, begnügt sich Scheme uniform mit der Notation  $(f\ x)$ .

ses Buch Programmieranfängern daher gleich zwei Superkräfte mit auf den Weg. Das nenne ich einen *Deal*!

Torsten Grust  
Tübingen, Juli 2019

## DANKSAGUNGEN

Wir, die Autoren, haben bei der Erstellung dieses Buchs immens von der Hilfe anderer profitiert.

Martin Gasbichler half bei der Entwicklung der DMdA-Erweiterungen. Peter Thiemann, Martin Gasbichler, Andreas Schilling, Torsten Grust und Michael Hanus hielten Vorlesungen auf Basis dieses Buchs und lieferten wertvolle Rückmeldungen. Besonderer Dank gebührt den Tutorinnen und Tutoren und Studierenden unserer Vorlesung *Informatik I*, die eine Fülle hilfreicher Kritik und exzellenter Verbesserungsvorschläge lieferten.

Seitdem die Entwicklung des Buches offen auf Github stattfindet, haben außerdem viele Leserinnen und Leser Verbesserungsvorschläge und Pull-Requests gemacht, darunter Nicolas Neuß, Joachim Breitner, Raphael Borun Das Gupta, Tobias Huttner, Nicolai Mainiero, Johannes Maier, Sibylle Hasse, Manuela Reinisch, Noah Haasis, mi-skam, Aramís, Beat Hagenlocher, Erika Bor und Christina Zeller.

Davor brachten Robert Giegerich, Ulrich Güntzer, Martin Plümicke, Christoph Schmitz und Volker Klaeren, Eric Knaul, Marcus Crestani, Sabine Sperber, Jan-Georg Smaus und Mayte Fleischer viele Korrekturen und Verbesserungen ein, insbesondere zum Vorgängerbuch *Vom Problem zum Programm*.

Wir sind außerdem dankbar für die Arbeit unserer Kolleginnen und Kollegen, die Pionierarbeit in der Entwicklung von Konzepten für die Programmierausbildung geliefert haben. Eine besondere Stellung nehmen Matthias Felleisen, Robert Bruce Findler, Matthew Flatt und Shriram Krishnamurthi und ihr Buch *How to Design Programs* [FELLEISEN et al., 2014] ein, das entscheidende didaktische Impulse für dieses Buch gegeben hat. Felleisens Arbeit im Rahmen des PLT-Projekts hat uns stark beeinflusst; das PLT-DrRacket-System ist eine entscheidende Grundlage für die Arbeit mit diesem Buch.

Herbert Klaeren  
Michael Sperber

Tübingen, Dezember 2021



## ○ WIE DU DAS MEISTE AUS DIESEM BUCH HERAUSHOLST

Der Inhalt dieses Buchs ist über viele Jahre weiterentwickelt worden, was dazu geführt hat, dass es sich von vielen anderen Büchern zur Einführung in die Programmierung deutlich unterscheidet.

Bevor es mit dem eigentlichen Stoff losgeht, enthält dieses Kapitel einige Hinweise dazu, wie Du am effektivsten damit umgehst, damit Du möglichst schnell Deine eigenen Programme selbst entwickeln kannst

### 0.1 VORKENNTNISSE

Das Buch richtet sich an alle Leserinnen und Leser, die Programmieren lernen möchten. Es kann außerdem als Einführung in die sogenannte *funktionale Programmierung* für Leserinnen und Leser dienen, die bereits in einem anderen Paradigma (zum Beispiel objektorientierte Programmierung) programmieren können.

Die benötigten Vorkenntnisse haben wir versucht, auf ein Minimum zu reduzieren. Grundkenntnisse in Mathematik sind hilfreich, aber nur in geringem Maße notwendig. Anhang A erläutert die im Buch verwendeten mathematischen Notationen und Termini.

### 0.2 ANREDE

Dieses Buch benutzt in der Anrede «Du» und «wir». Mit «Du» bist Du, also die Leserin, der Leser oder das Lesery gemeint. Mit «wir» meinen wir uns, die Autoren. Wir benutzen oft Formulierungen wie «Zunächst schreiben wir eine Signatur», wenn wir über die Beispiele im Buch schreiben, die wir uns für Dich ausgedacht haben. Wenn das Buch Dich mit «Du» anspricht, haben wir ein konkretes Anliegen für Dich.

### 0.3 SOFTWARE

Die Programmierbeispiele dieses Buchs bauen auf der Programmierumgebung DrRacket auf, die speziell für die Programmierausbildung entwickelt wurde. DrRacket ist kostenlos im Internet auf

der Seite

<https://www.racket-lang.org/>

erhältlich und läuft auf Windows-, Mac- und Unix-/Linux-Rechnern.

## 0.4 PROGRAMMIERSPRACHE

Wer Programmieren lernt, muss sich mit ziemlich vielen Dingen herumschlagen: den Konzepten einer konkreten Programmiersprache, den dort benutzten Schreibweisen, einer Programmierungsumgebung und den sonstigen Werkzeugen, die für das Programmieren benötigt werden.

Die meisten Programmiersprachen und deren Werkzeuge sind für professionelle Entwickler entwickelt worden und setzen damit viele Kenntnisse voraus, die Du ja erst noch lernen willst.

Aus diesem Grund verwendet der vorliegende Text eine Serie von speziell für die Lehre entwickelten Programmiersprachen, die auf *Racket* und *Scheme* basieren und beim Racket-System mitgeliefert. Diese Lehrsprachen sind über die DrRacket-Entwicklungsumgebung Anfängerinnen und Anfängern besonders gut zugänglich.

## 0.5 ENGLISCH

Das Buch ist natürlich auf Deutsch, und auch die Programmierungsumgebung DrRacket kann ihre Menüs und sonstigen Texte auf Deutsch darstellen.

Die Programmiersprachen benutzen allerdings englische Wörter, und wir benutzen in Programmen englische Namen für alles mögliche. Das erfordert keine fortschgeschrittenen Englisch-Kenntnisse. Bei selteneren Wörtern liefern wir die Übersetzung. (Du kannst gern deutsche Namen in Deinen Programmen benutzen.)

Außerdem haben wir noch nicht alle Fehlermeldungen ins Deutsche übersetzt, arbeiten aber daran.

## 0.6 EXKURSE

Generell wurde das Buch so geschrieben, dass die Kapitel und Abschnitte innerhalb der Kapitel aufeinander aufbauen. Ab und zu jedoch gibt es Abschnitte (in der Regel solche mit Mathema-

tik), die wir interessant finden, die aber nicht notwendig sind, um das darauf folgende Material zu verstehen. Diese Abschnitte sind mit dem Wort «Exkurs» gekennzeichnet.

## 0.7 AUFGABEN

Das Buch enthält viele Übungsaufgaben. Es ist unerlässlich, dass Du zumindest einige davon bearbeitest, wenn Du lernen willst, eigenständig zu programmieren.

In jedem Kapitel gibt es Aufgaben am Ende: Davon kannst Du Dir die aussuchen, die Dir gefallen. Es gibt aber auch Aufgaben mittendrin, die sich auf den unmittelbar vorangehenden Stoff beziehen. (Das Zeichen □ markiert jeweils, wenn die Aufgabe zu Ende ist und es mit dem Text weitergeht.) Diese solltest Du bearbeiten, um sicherzustellen, dass Du diesen Stoff verstanden hast.

## 0.8 BÜROKRATISCHE DIDAKTIK

Wir wollen ehrlich sein: Programmieren ist eine komplexe Tätigkeit und damit schwierig. Wir – die Autoren – verbringen bereits weit über die Hälfte unseres Lebens damit und lernen immer noch dazu. Aber Programmieren macht auch immens Freude, wenn man erfolgreich eigenständig aus puren Gedanken etwas machen kann, das läuft und kommuniziert. Der Lernaufwand lohnt sich also.

Aber niemand hat Freude daran, bei der Lösung einer Programmieraufgabe steckenzubleiben und in einer Sackgasse zu landen. Wir wollen Dir darum mit diesem Buch nicht einfach nur schöne Beispielprogramme präsentieren und sagen «jetzt mach halt selber», sondern Dir Schritt für Schritt vermitteln, *wie* das geht.

Die Methode, die wir dafür entwickelt haben und die erfolgreich hunderten von Programmieranfängerinnen und -anfängern angewendet haben, wird Dir auf den ersten Augenblick enorm bürokratisch erscheinen: Dieses Buch behandelt sozusagen die deutsche Beamtenmethode des Programmierens.

Gelegentlich wirst Du Dich gegängelt fühlen und vielleicht auch etwas genervt. Aber Du wirst (hoffentlich!) diese Techniken nach und nach einsetzen können, um nicht nur Übungsaufgaben anderer zu lösen, sondern auch Deine eigenen Ideen umsetzen und Deiner eigenen Kreativität freien Lauf zu lassen: Dafür brauchst Du die Methoden dieses Buchs, inklusive der Bürokratie.

Zwei Elemente tragen besonders zum Lernerfolg bei:

- Die *Konstruktionsanleitungen* schreiben Dir Schritte vor, mit denen Du von einer Problemstellung zur Lösung kommst – die deutschen Beamtenvorschriften sozusagen.
- Die *Mantras* sind wiederkehrende Prinzipien der Programmierung, die es lohnt, immer parat zu haben. Zum Glück gibt es nur eine Handvoll.

## 0.9 ZUSATZMATERIAL

Das Buch ist aus dem DEINPROGRAMM-Projekt entstanden, dass sich allgemein mit der effektiven Ausbildung im Programmieren beschäftigt. Das Projekt hat seine Homepage hier:

<https://www.deinprogramm.de/>

Dort befindet sich weiteres Material zum Buch, insbesondere Quelltext für alle Programmbeispiele zum Herunterladen und Publikationen mit Hintergrundinformationen zum Projekt.



# 1 ELEMENTE DES PROGRAMMIERENS

Dieses Kapitel gibt einen Überblick über das wichtigste Handwerkszeug des Programmierens: Welche Software Du dafür brauchst, wie Du diese Software bedienst, wie kleine Programme aussehen und aus welchen Elementen sie bestehen.

## 1.1 DRRACKET

Zunächst einmal benötigst Du die Software *Racket*, die es zum kostenlosen Download auf der Seite <https://racket-lang.org/> für alle gängigen Plattformen gibt. Auf der Seite ist ein Download-Knopf – folge den dortigen Instruktionen für die Installation.<sup>1</sup>

Zu Racket gehört ein Programm namens *DrRacket*: starte es. Es erscheint ein Fenster, das ungefähr so aussehen sollte wie in Abbildung 1.1. Die Benutzeroberfläche kannst Du auf Deutsch umstellen, indem Du im Help-Menü auf **Deutsche Benutzeroberfläche für DrRacket** drückst. Wenn Du die Auswahl bestätigst, wird DrRacket danach beendet und Du musst es noch einmal starten; dann sollten die Menüs auf Deutsch sein.

DrRacket ist eine *Entwicklungsumgebung*, mit der Du Programme schreiben und ausführen kannst. DrRacket unterstützt nicht nur eine Programmiersprache, darum musst Du die richtige Programmiersprache für dieses Buch noch auswählen. Dazu wählst Du den Menüpunkt **Sprache** → **Sprache auswählen** (beziehungsweise **Language** → **Choose language** in der englischen Fassung), worauf ein Dialog erscheint. In dem Dialog gibt es eine Abteilung **Lehrsprachen**, und darunter eine Überschrift namens **Dein Programm**, unterhalb dessen mehrere Einträge erscheinen, die speziell auf die Kapitel dieses Buchs zugeschnitten sind.

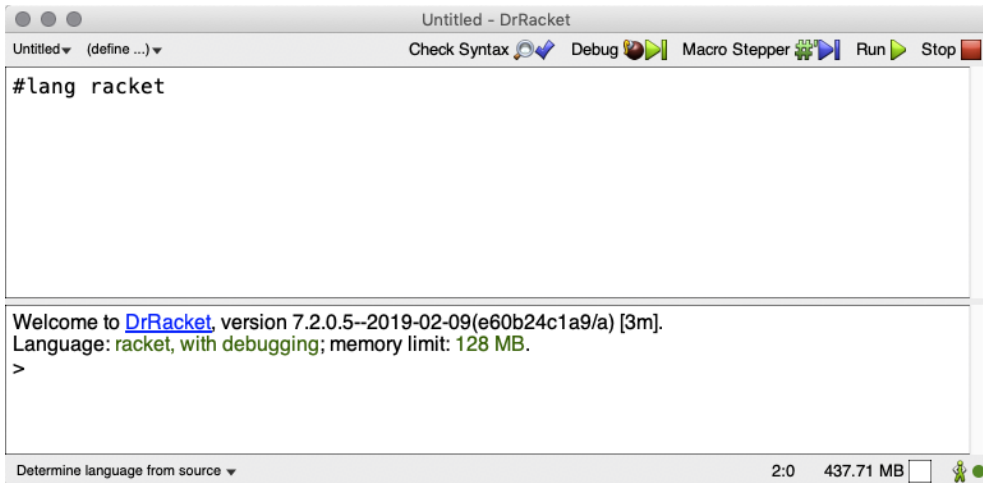
Für den ersten Teil des Buchs ist die Ebene **Schreibe Dein Programm! – Anfänger** zuständig: Wähle diese aus und drücke dann einmal oben rechts auf den Knopf **Start** (beziehungsweise **Run**), damit die Auswahl aktiv wird. Das Ergebnis sollte dann so aussehen wie in Abbildung 1.2.

Das DrRacket-Fenster besteht aus zwei Teilen:

1. In der oberen Hälfte des Fensters (dem *Editor* oder *Definitions-fenster*) steht der Programmtext. Der Editor funktioniert ähnlich wie ein reguläres Textverarbeitungsprogramm. Was dort steht,

---

<sup>1</sup> Achte darauf, dass Du mindestens Version 8.0 bekommst – das ist besonders relevant, wenn Du einen anderen Installationsweg wählst, wie zum Beispiel einen Package-Manager für eine Linux-Distribution.



**Abbildung 1.1:** DrRacket nach dem ersten Start

kannst Du abspeichern.

2. In der unteren Hälfte des Fensters – dem *Interaktionsfenster* oder der sogenannten *REPL*<sup>2</sup> werden die Ausgaben des Programms angezeigt. Außerdem kannst Du hier «Fragen» an das Programm stellen, um einzelne Programmteile gezielt auszuprobieren.

## 1.2 AUSDRÜCKE UND DIE REPL

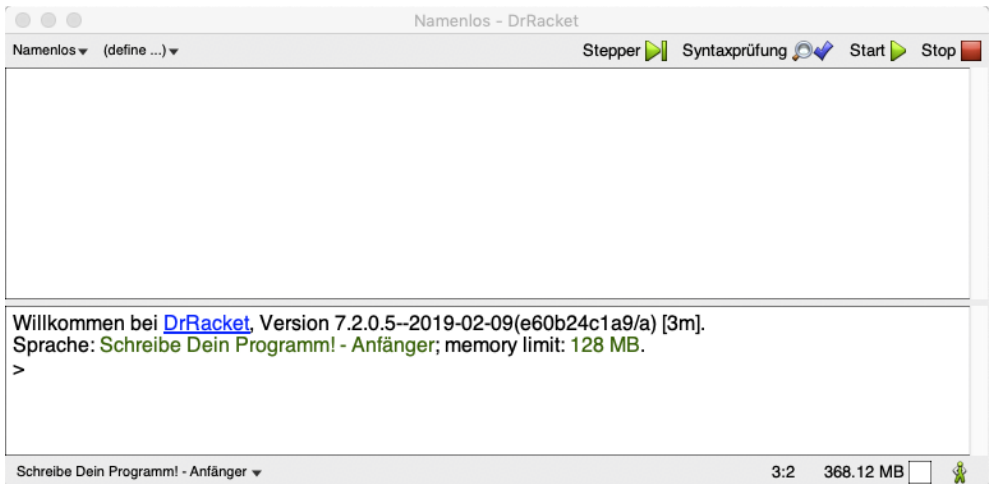
Fangen wir mit der REPL an: Wenn Du gerade **Start** gedrückt hast, dann erscheint der Cursor rechts von dem **>**-Zeichen: Du kannst da etwas eingeben und Return drücken – DrRacket zeigt dann das Ergebnis darunter an.

Wenn Du zum Beispiel **123** eintippst, zeigt DrRacket gleich darunter **123** an. DrRacket kann auch rechnen. Dafür musst Du allerdings die Rechenaufgaben etwas anders aufschreiben als sonst. Zum Beispiel so:

**(+ 123 42)**

Wenn Du das in die REPL tippst, zeigt DrRacket **165** an, die *Summe* von 123 und 42. Diese «Rechenaufgabe» ist ein sogenannter *Ausdruck*. Ausdrücke, die etwas ausrechnen sollten, haben in den Lehrsprachen immer die gleiche Form, und die sieht so aus:

<sup>2</sup> «REPL» steht für «Read-Eval-Print-Loop» – wir werden später zeigen, warum.



**Abbildung 1.2:** DrRacket mit ausgewählter Lehrsprache

*(Operator Operand ...)*

Das heißt, da stehen *immer* Klammern drumherum (und die können auch nirgendwo sonst stehen). Dann steht da der *Operator*, der bestimmt, *was* gemacht wird, also die *Operation*. Danach kommen die *Operanden*, welche die Eingaben für die Operation bestimmen.

Zum Merken ist es hilfreich, Ausdrücke entsprechend vorzulesen: Also nicht mehr «Hundertdreißig plus zweiundvierzig», sondern «die *Summe* von Hundertdreißig und zweiundvierzig».

Wie schon gesagt, die Klammern sind wichtig. Wenn Du sie vergisst, können verwirrende Ergebnisse herauskommen. Wenn Du zum Beispiel `+ 123 42` in der REPL eintippst, sieht das so aus:

```
> + 123 42
#<function:+>
123
42
```

Das liegt daran, dass ohne Klammern `+ 123 32` aus *drei* Ausdrücken besteht und die REPL darum auch drei Ergebnisse ausdrückt: Die Funktion `+`, dann die `123`, dann `42`. (Die Operation `+` ist eine sogenannte Funktion – das erläutern wir später noch genauer.) Ähnlich verhält es sich mit `123 + 42` – probiere es aus!

Wenn Du versehentlich den Operator dazwischen schreibst, erscheint in der REPL eine Fehlermeldung:

```
> (123 + 42)
```

Operator darf keine Zahl sein, ist aber 123

Etwas ähnliches passiert, wenn das Leerzeichen zwischen dem Operator und den Operanden fehlt – oder zwischen den Operanden. Das sieht zum Beispiel so aus:

```
> (+123 42)
```

Operator darf keine Zahl sein, ist aber 123

Das liegt daran, dass `+123` zusammen die Zahl «plus hundertdreiundzwanzig» bildet und nicht etwa in das `+`-Zeichen und `123` aufgeteilt wird.

Wenn ein Ausdruck wie `(+ 123 42)` ein Ergebnis wie 165 hat, schreiben wir das im Buch zukünftig so, etwas anders als die DrRacket-REPL:

```
(+ 123 42)
```

```
↪ 165
```

Es gehen natürlich nicht nur Summen, sondern auch Differenzen, Produkte und Quotienten:

```
(- 123 42)
```

```
↪ 81
```

```
(* 123 42)
```

```
↪ 5166
```

```
(/ 123 42)
```

```
↪ 2.9285714
```

Beim letzten Ausdruck ist zu sehen, dass Dezimalzahlen mit Punkt und nicht mit Komma geschrieben werden. Der Überstrich bei `2.9285714` ist eine *Periode*. Die Zahl ist also eigentlich

```
2.9285714285714285714285714285714...
```

Die REPL funktioniert also folgendermaßen: Sie *liest* einen Ausdruck ein (auf Englisch «read»), berechnet dessen Wert oder *wertet* diesen *aus* («eval») und zeigt das Ergebnis an oder *druckt* dieses *aus* («print») – und dann geht es von vorn los, wie in einer Schleife («loop»). Die Abfolge *Read-Eval-Print-Loop* gibt der REPL ihren Namen.

Ausdrücke können auch kombiniert werden, zum Beispiel so:

```
(* 123 (+ 20 22))
```

```
↪ 5166
```

```
(* 123 (+ (* 2 10) 22))
```

```
↪ 5166
```

Bei der Kombination ist wichtig, dass um jeden Teilausdruck wieder ein Klammernpaar kommt. Ist das nicht der Fall, erscheinen gelegentlich auch mal Fehlermeldungen wie diese hier:

```
(* 123 (+ * 2 10 22))
```

**+: Zahl als erstes Argument erwartet, #<function:\*> bekommen**

Das liegt daran, dass das `*` hier an der Stelle eines Operanden steht. Das Wort *Argument* steht für den Wert des Operanden einer Funktion, darum steht da sinngemäß, dass `+` eine Zahl als erstes Argument erwartet, aber stattdessen die Funktion `*` bekommen hat. (Das Wort «Argument» definieren wir genauer in Abschnitt 1 auf Seite 22.)

Abspeichern geht mit der REPL nicht, und der REPL-Inhalt verschwindet auch jedesmal, wenn Du **Start** beziehungsweise **Run** drückst. Du kannst aber frühere Eingaben zurückholen, indem Du **Strg-↑** beziehungsweise **Control-↑** (je nach Computertyp) drückst. (Das geht natürlich auch nach unten mit **Strg-↓** beziehungsweise **Control-↓**.) Das kannst Du auch benutzen, um einen fehlerhaften Ausdruck zu korrigieren, bei dem Du schon **Return** gedrückt hast.

#### AUFGABE 1.1

Schreibe folgende «mathematischen» Ausdrücke in der Notation der Lehrsprache in die REPL und lasse die REPL sie auswerten:

$$\begin{aligned}
 &55 * 27 \\
 &23 * (44 + 27) \\
 &\frac{23}{44} + \frac{44}{23} \\
 &(23 + 42) * (12 + (14 * 2))
 \end{aligned}$$

□

### 1.3 DAS DEFINITIONSFENSTER

Kommen wir zum Definitionsfenster oben. Dort schreibst Du Dein Programm, die REPL kannst Du dann benutzen, um es auszuprobieren. Schreib in das Definitionsfenster folgende *Definition*:

```
(define alles (+ 20 22))
```

Diese Definition besagt, dass der Name `alles` für das Ergebnis von `(+ 20 22)` steht. Um das auszuprobieren, drück auf den Knopf **Start** beziehungsweise **Run** rechts oben. Der Cursor landet dann wieder in der REPL, wo Du das Programm ausprobieren kannst:

```
alles  
↪ 42
```

Ein Name, der in einem Programm so definiert ist, heißt *Variable*.

Hier sind ein paar weitere Beispiele für Definitionen:

```
(define one 1)  
(define temperature 23)  
(define birgit-prinz 9)  
(define michael-ballack 13)
```

Bei der Gestaltung eines Namens gibt es weitgehende Freiheiten.<sup>3</sup> Nur Leerzeichen sind nicht erlaubt.

In einem Programm kannst Du Zeilenumbrüche und Einrückung benutzen, um Dein Programm übersichtlicher zu gestalten. Zum Beispiel kannst Du nach `alles` die Return-Taste drücken, das Ergebnis sieht so aus:

```
(define alles  
  (+ 20 22))
```

DrRacket rückt die zweite Zeile ein bisschen ein, um auszudrücken, dass sie noch in die Klammern vom `define` gehört. Bei komplizierteren Ausdrücken ist das hilfreich:

```
(define alles  
  (+ 20  
    (* 11 2)))
```

Hier stellt DrRacket die Operanden der Summe genau untereinander. DrRacket zwingt Dich nicht, die Einrückung genau so zu machen, und durch Änderungen im Programm gerät sie auch manchmal aus dem Lot. In dem Fall kannst Du die Tab-Taste drücken (auf den meisten Tastaturen steht  $\rightarrow$  | auf der Tab-Taste) und in der Zeile, in welcher der Cursor steht, wird die Einrückung korrigiert. Wenn Du einen Abschnitt im Programm markierst und dann Tab drückst, wird der ganze Abschnitt neu eingerückt. Es gibt außerdem einen Menüpunkt `Racket`  $\rightarrow$  `Alles einrücken` (beziehungsweise `Racket`  $\rightarrow$  `Reindent All`), der das für das gesamte Programm macht.

Ein weiterer praktischer Trick ist, dass Du einen geklammerten Ausdruck markieren (und dann ausschneiden) kannst, indem Du auf die öffnende oder schließende Klammer doppelt klickst. Insbesondere kannst Du den markierten Ausdruck dann mit einem Druck auf die Tab-Taste korrekt einrücken.

---


<sup>3</sup> Anders als in anderen Programmiersprachen sind auch Bindestriche in Namen möglich.

## AUFGABE 1.2

Bring bei einem mehrzeiligen Programm die Einrückung richtig durcheinander, zum Beispiel so:

```
(define nr
  (+ 12
    (- (* 42
        13)
      500))))
```

Benutze dann die Tab-Taste, um die Einrückung wieder zu korrigieren. □

Den Inhalt des Definitionsfenster kannst Du abspeichern, indem Du auf den Knopf mit dem Diskettensymbol  drückst.

## 1.4 RECHNEN OHNE ZAHLEN

Computerprogramme können nicht nur mit Zahlen rechnen. In diesem Abschnitt geht es um das Rechnen mit Text und das Rechnen mit Bildern.

### 1.4.1 RECHNEN MIT TEXT

Aus Text kann durch doppelte Anführungszeichen ein Wert werden:

```
"Mike Sperber"
"Herbert Klaeren"
"Schreibe Dein Programm!"
```

Solche Text-Werte heißen *Zeichenketten*.

Die einfachste Art, eine Zeichenkette herzustellen, ist, die Buchstaben hinzuschreiben, aus denen sie besteht. Die Anführungszeichen müssen drumherum, damit die Zeichenketten von anderen Ausdrücken unterschieden werden können. Die Anführungszeichen gehören aber nicht zu den Buchstaben dazu, aus denen die Zeichenkette besteht – "abc" besteht aus den drei Buchstaben abc.

Abbildung 1.3 beschreibt die genaue Schreibweise für solche «festen» Zeichenketten. Feste Schreibweisen für Werte heißen allgemein *Literale*. Das kennen wir schon von den Zahlen: Die

*Zeichenketten* (auf Englisch *Strings*) repräsentieren Text. Literale für Zeichenketten haben folgende Form:

```
"z1z2 ... zn"
```

Dabei sind die  $z_i$  beliebige einzelne Zeichen, außer " selbst. Beispiel:

```
"Herbert was here!"
```

Das Anführungszeichen (") kann nicht «ungeschützt» vorkommen, da es das Ende der Zeichenkette markiert. Es wird als Zeichen innerhalb einer Zeichenkette durch \ " dargestellt:

```
"Herbert sagt, Mike wäre \"doof\"!"
```

**Abbildung 1.3:** Zeichenketten

Zeichenfolge 123 steht für die Zahl «hundertdreißig». Kästen wie Abbildung 1.3 werden in diesem Buch noch oft dazu dienen, neue Sprachelemente einzuführen.

Mit Text kann ein Programm auch rechnen, und zwar mit der eingebauten Funktion `string-append`, die zwei Zeichenketten aneinanderhängt:

```
(string-append "Herbert" "Mike")
↪ "HerbertMike"
(string-append "Mike" " " "ist doof")
↪ "Mike ist doof"
```

Die eingebaute Funktion `string-length` liefert die Anzahl der Zeichen in einer Zeichenkette:

```
(string-length "Herbert")
↪ 7
(string-length "Mike")
↪ 4
```

Die Namen `string-append` und `string-length` sehen auf den ersten Blick «anders» aus als + und \* zum Beispiel, dieser Eindruck täuscht jedoch: Sie sind allesamt Namen von vordefinierten Operationen, die Programme benutzen können, ohne sie selbst definieren zu müssen.

Die vordefinierten Funktionen `string->number` und `number->string` konvertieren zwischen Zahlen und den Zeichenketten, die diese darstellen:

```
(string->number "23")
↪ 23
(number->string 23)
↪ "23"
```



## AUFGABE 1.3

Mache Dir den Unterschied zwischen der Zahl 23 und der Zeichenkette "23" klar. Probiere zum Beispiel aus:




```
(+ "23" 42)
(string-append 23 "42")
(number->string "23")
```

□

## 1.4.2 RECHNEN MIT BILDERN

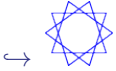
Programme können auch mit Bildern rechnen. Dazu wird eine Erweiterung zu DrRacket benötigt, ein sogenanntes *Teachpack*: Wähle den Menüpunkt *Sprache* → *Teachpack* hinzufügen (Language → Add Teachpack). In dem folgenden Dialog wähle `image.rkt` aus und drücke OK. Dann musst Du noch einmal auf *Start* drücken. Wenn alles geklappt hat, steht unten im REPL Fenster `Teachpack: image.rkt`.

Das Teachpack `image.rkt` enthält zusätzliche vordefinierte Funktionen, mit denen wir Bilder herstellen können. Zum Beispiel `square`, `circle` und `star-polygon`:

```
(square 40 "solid" "red")
↩ 
(circle 20 "solid" "green")
↩ 
(star-polygon 20 10 3 "solid" "blue")
↩ 
```

Schauen wir uns das eingebaute `star-polygon` etwas näher an: Es akzeptiert fünf Eingaben, die ersten drei davon sind Zahlen – die Seitenlänge, die Seitenanzahl und die Anzahl der Ecken, die für jede Seite übersprungen wird. Danach kommen zwei Zeichenketten – "`solid`" heißt, dass das Innere des Sterns ausgefüllt ist und "`blue`" ist die Farbe. Statt "`solid`" kannst Du auch "`outline`" schreiben, dann wird etwas klarer, was «überspringen» heißt:

```
(star-polygon 20 10 3 "outline" "blue")
```



#### AUFGABE 1.4

Wofür stehen die Zahleneingaben bei `square` und `circle`? Probiere unterschiedliche Zahlen aus! Es gibt auch eine eingebaute Funktion `rectangle`. Kannst Du ein funktionierendes Beispiel für den Einsatz von `rectangle` konstruieren? Außerdem gibt es eine eingebaute Funktion `ellipse`, die genauso benutzt wird wie `rectangle` – probiere sie aus! ☐

Bilder sind Werte wie Zahlen und Zeichenketten auch. Du kannst mit Definitionen auch Namen dafür vergeben:

```
(define s1 (square 40 "solid" "red"))
(define c1 (circle 40 "solid" "green"))
(define p1 (star-polygon 20 10 3 "solid" "blue"))
```

#### AUFGABE 1.5

Probiere diese Definitionen in der REPL aus, indem Du dort `s1`, `c1` und `p1` eingibst. Was passiert, wenn Du `(s1)` eingibst? Warum? ☐

#### AUFGABE 1.6


Du kannst auch Bilddateien oder Bilder von Webseiten in Dein Programm einfügen, wie in einem Textverarbeitungsprogramm. Das geht so:

- Um eine Bilddatei einzufügen, wähle den Menüpunkt **Einfügen** → **Bild einfügen** beziehungsweise **Insert** → **Insert Image**.
- Um ein Bild aus einer Webseite einzufügen, wähle aus dem Kontextmenü im Browser **Bild kopieren** (oder **Copy Image** oder so ähnlich) aus. Du kannst das Bild dann in DrRacket über den Menüpunkt **Bearbeiten** → **Einfügen** beziehungsweise **Edit** → **Paste einfügen**.

Probiere es aus und gib dem Bild einen Namen! ☐


Mit Bildern kann DrRacket auch rechnen:

```
(beside s1 p1)
```



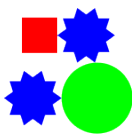
↪

```
(above s1 c1)
```



↪

```
(above (beside s1 p1) (beside p1 c1))
```



↪

### AUFGABE 1.7

Probiere in der der REPL folgende Ausdrücke aus:

```
(triangle 50 "outline" "red")
(square 100 "solid" "blue")
```

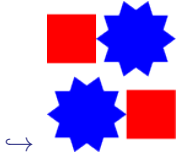
Schreibe mit Hilfe der Funktionen `triangle` und `square` einen Ausdruck, der ein ganz einfaches Haus berechnet. □

Das heißt, `beside` akzeptiert als Eingabe zwei (oder mehr) Bilder und macht daraus wieder ein Bild, in dem die Teilbilder nebeneinander stehen. Das gleiche gilt für `above`, nur dass die Bilder übereinander angeordnet werden. Selbstverständlich können `above` und `beside` auch kombiniert werden.

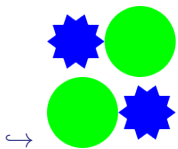
## 1.5 ABSTRAKTION UND APPLIKATION

Hier sind zwei Ausdrücke, welche die Bilder aus dem vorigen Abschnitt zu einem Muster kombinieren:

```
(above
  (beside s1 p1)
  (beside p1 s1))
```



```
(above
  (beside p1 c1)
  (beside c1 p1))
```



Beide Ausdrücke folgen dem gleichen Muster, sie «kacheln» jeweils zwei Bilder in einer quadratischen Anordnung. Das Muster könnte man so hinschreiben:

```
(above
  (beside a b)
  (beside b a))
```

Das erste Beispiel entsteht dann, indem für `a` `s1` eingesetzt wird und für `b` `p1`, im zweiten für `a` `p1` und für `b` dann `c1`.

Dieses Muster kannst Du in ein Programm gießen. Dann müssen wir nicht jedesmal `above ... beside ... beside` eintippen. Dazu schreibst Du es erst einmal genauso hin, also mit `a` und `b`. Wenn Du das Programm startest, erscheint folgende Meldung:

**a: Variable nicht definiert**

Na klar, es gibt ja auch keine Definition für `a`. Du könntest eine hinschreiben, zum Beispiel:

```
(define a s1)
```

Aber dann wäre `a` auf `s1` festgelegt. Wir wollen stattdessen das Muster verallgemeinern, so dass Du es mehrmals mit unterschiedlichen Werten für `a` und `b` verwenden kannst. Dieser Verallgemeinerungsprozess heißt beim Programmieren *Abstraktion*. Dafür müssen wir dem Muster noch etwas hinzufügen, um zu sagen, dass wir unterschiedliche Werte für `a` und `b` einsetzen wollen:

```
(lambda (a b)
  (above
    (beside a b)
    (beside b a)))
```

Das **lambda** ist eine Art Zauberwort, es sagt soviel wie: «Für die Variablen **a** und **b** möchte ich später (und vielleicht mehrmals) Werte einsetzen.» Wenn Du das Programm jetzt startest, geht die Fehlermeldung weg und in der REPL erscheint:

```
#<function>
```

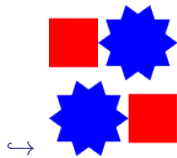
Das deutet darauf hin, dass bei dem **lambda** eine Funktion herauskommt. Um etwas damit anzufangen, gib ihr einen Namen mit **define**:

```
(define tile
  (lambda (a b)
    (above
      (beside a b)
      (beside b a))))
```

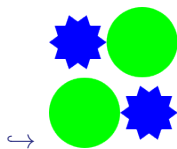
(Alles schön richtig einrücken! «Tile» heißt auf Deutsch «kacheln».)

Dadurch ist eine neue Operation entstanden. Die kannst Du direkt verwenden:

```
(tile s1 p1)
```



```
(tile p1 c1)
```



Was ist da passiert? Am einfachsten kann man das in einem speziellen Werkzeug sehen, dem *Stepper*. Um ihn zum Einsatz zu bringen, Sorge dafür, dass im Definitionsfenster folgendes Programm steht, also das, was weiter oben schon steht:

```
(define s1 (square 40 "solid" "red"))
(define c1 (circle 40 "solid" "green"))
(define p1 (star-polygon 20 10 3 "solid" "blue"))

(define tile
  (lambda (a b)
```

```
(above
  (beside a b)
  (beside b a)))
```

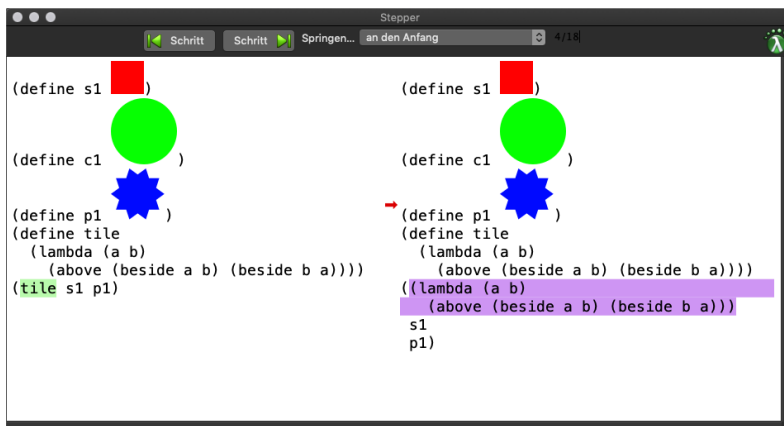
```
(tile s1 p1)
(tile p1 c1)
```

Dann drück auf den **Stepper**-Knopf oben rechts im DrRacket-Fenster. Es erscheint ein neues Fenster, das so aussieht:



**Abbildung 1.4:** Der Stepper in Aktion

Du kannst jetzt mit den Knöpfen mit dem Vorwärts- beziehungsweise dem Rückwärts-Pfeil zusehen, wie DrRacket Dein Programm ausführt. Wenn Du ein paar Schritte vorwärts gehst, sieht das so aus:



**Abbildung 1.5:** Stepper: Variable ersetzen

Du kannst sehen, wie DrRacket jeweils einen Ausdruck einen Schritt auf einmal auswertet – der ist auf der linken Seite grün – und dann rechts durch das Resultat ersetzt. Oben kannst Du sehen, wie

der Stepper die Variable `tile` durch den `lambda`-Ausdruck aus der Definition von `tile` ersetzt. Das macht der Stepper auch mit den Definitionen von `s1` und `p1`:

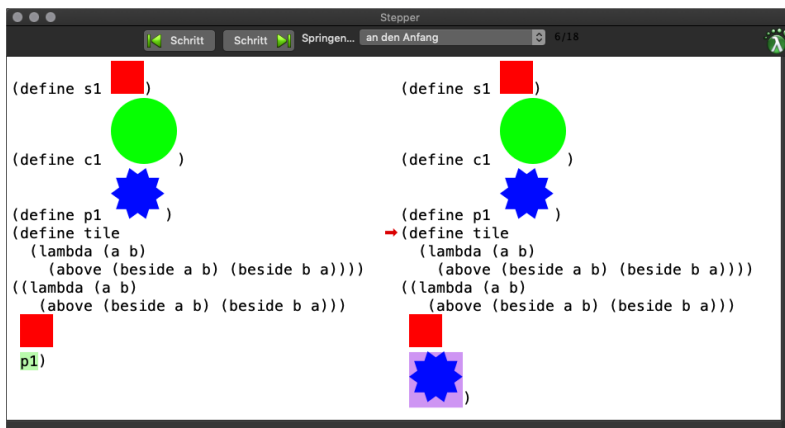


Abbildung 1.6: Stepper: Variable

Interessant wird es danach. Die Variablen aus den Definitionen sind alle ersetzt. Hinter dem `lambda`-Ausdruck stehen jetzt das Quadrat und der Stern, und die werden jetzt für die Variablen aus dem `lambda`-Ausdruck eingesetzt, also das Quadrat für `a` und der Stern für `b`:

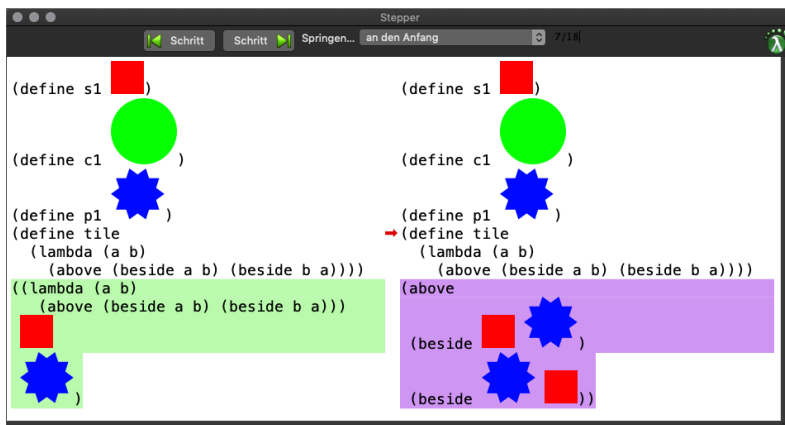


Abbildung 1.7: Stepper: Applikation

Dabei verschwindet auch das `lambda`.

Eine Abstraktion hat folgende Form:

```
(lambda (p1 ... pn) e)
```

Die  $p_i$  sind jeweils Namen, die *Parameter*, und  $e$  ist der *Rumpf*. In  $e$  dürfen die  $p_i$  vorkommen. Der Wert einer Abstraktion ist eine *Funktion*, welche für jeden Parameter eine Eingabe erwartet. Eine *Applikation* einer Funktion hat folgende Form:

```
(f a1 ... an)
```

$f$  ist ein Ausdruck, der eine Funktion ergeben muss, die  $a_i$  sind ebenfalls Ausdrücke, die *Argumente*. Bei der Auswertung einer Applikation werden zunächst  $f$  und die  $a_i$  ausgewertet; danach geht es mit der Auswertung des Rumpfes der Funktion weiter, wobei für die Parameter  $p_i$  jeweils die Werte der Argumente  $a_i$  eingesetzt werden.

**Abbildung 1.8:** Abstraktion und Applikation

Probiere das Beispiel im Stepper aus und klicke Dich bis zum Ende durch. Mache Dir dabei klar, wie jeweils die linke mit der rechten Seite zusammenhängt. □

`(tile s1 p1)` ist eine sogenannte *Applikation*: Das ist ein Ausdruck mit (natürlich) Klammern drum. Der Operator und die Operanden der Applikation sind ebenfalls Ausdrücke. Der Operator ergibt eine Funktion – entweder eingebaut oder aus einem `lambda`-Ausdruck. Eine Applikation wird auch oft *Funktionsaufruf* oder *Funktionsanwendung* genannt.

Wenn die Funktion ein `lambda`-Ausdruck ist wie in Abbildung 1.7, dann muss es genauso viele Operanden geben wie der `lambda`-Ausdruck Variablen hat. Dann wertet DrRacket den Ausdruck wie folgt aus:

1. Die Operanden werden ausgewertet und ergeben die Eingaben für die Funktion, die sogenannten *Argumente*.
2. Die Argumente werden für die Variablen des `lambda`-Ausdrucks – die sogenannten *Parameter* – im Innenteil des `lambda`-Ausdrucks, dem *Rumpf* eingesetzt.

`lambda`-Ausdrücke heißen auch *Abstraktionen*. Abbildung 1.8 fasst zusammen, wie Abstraktionen und Applikationen aufgebaut sind und ausgewertet werden.

Verwirrend ist vielleicht der Unterschied zwischen Abstraktion und Funktion: Die Abstraktion ist der `lambda`-Ausdruck im Programm, die Funktion ist der Wert, der bei der Programmauswertung dabei herauskommt. Wir verwenden im Buch aber häufig den Begriff «Funktion» für eine Abstraktion, weil uns interessiert, was bei der Abstraktion herauskommt.

#### AUFGABE 1.9

Kannst Du das Ergebnis von `tile` wieder als Eingabe für `tile` benutzen? Probier es aus! □



## 1.6 INFORMATION UND DATEN

Eine Definition wie

```
(define mehrwertsteuer 19)
```

suggeriert, dass die Zahl 19 an dieser Stelle eine Bedeutung «in der realen Welt» hat, zum Beispiel in einem Programm, das eine Registrierkasse steuert oder das bei der Steuererklärung hilft. Die Bedeutung könnte folgende Aussage sein: «Der Mehrwertsteuersatz beträgt 19%.» Dieser Satz repräsentiert *Information*, also einen Fakt über die Welt oder zumindest den Ausschnitt der Welt, in dem das Programm arbeiten soll. In Computerprogrammen wird Information in eine vereinfachte Form gebracht, mit der das Programm rechnen kann – in diesem Fall die Zahl 19. Diese vereinfachte Form heißt *Daten*: Daten sind *Repräsentationen* für Information.

Eine der wichtigsten Aufgaben beim Programmieren ist, die richtige Form für die Daten zu wählen, um die für das Programm relevanten Informationen darzustellen und die Informationen dann in Daten zu übersetzen.

Nicht immer ist offensichtlich, welche Information durch bestimmte Daten repräsentiert werden. Die Zahl 23 zum Beispiel könnte eine Reihe von Informationen darstellen:

- die Anzahl der Haare von Bruce Willis
- die aktuelle Außentemperatur in °C in Tübingen
- die Außentemperatur vom 1.7.2000 in °C in Tübingen
- die Größe in m<sup>2</sup> des Schlafzimmers
- die Rückennummer von Alexandra Popp

Welche gemeint ist, hängt vom Kontext ab. Damit andere unsere Programme lesen können, werden wir also immer wieder klarstellen müssen, was der Kontext ist und wie Information in Daten zu übersetzen ist und umgekehrt. An konkreten Beispielen können (und sollten) wir dazu einen Kommentar anbringen:

```
(define bruce-willis 23) ; Anzahl der Haare
(define temperatur-tübingen 23) ; 23°C in Tübingen
(define temperatur-tübingen-2000-07-01 23)
      ; 23°C in Tübingen am 7.1.2000
(define schlafzimmer 23) ; Schlafzimmer ist 23m2
(define alexandra-popp 23) ; Rückennummer
```

## 1.7 PROGRAMME SYSTEMATISCH KONSTRUIEREN

In Abschnitt 1.5 haben wir gezeigt, wie Abstraktion und Applikation bei der Programmauswertung funktionieren. Das Verständnis dafür ist wichtig, aber Du musst beim Schreiben Deines Programms nicht die ganze Zeit daran denken, wie DrRacket Dein Programm auswertet. Entsprechend kümmern wir uns in diesem Buch hauptsächlich darum, wie Du Programme systematisch konstruierst. Schauen wir uns das anhand einer kleinen aber realistischen Aufgabe an:

Betrachte folgende Stromtarife. Beide Tarife bestehen aus einer monatlichen Grundgebühr und einem Teil, der sich nach den verbrauchten Kilowattstunden (kWh) richtet.

	Grundgebühr pro Monat	Verbrauchspreis pro kWh
Tarif «Billig-Strom»	4,90 €	0,19 €
Tarif «Watt für wenig»	8,20 €	0,16 €

1. Schreibe ein Programm, das den Monatsverbrauch in Kilowattstunden akzeptiert und den im Tarif «Billig-Strom» zu zahlenden monatlichen Rechnungsbetrag berechnet.
2. Schreibe ein Programm, das den Monatsverbrauch in Kilowattstunden akzeptiert und den im Tarif «Watt für wenig» zu zahlenden monatlichen Rechnungsbetrag berechnet.

Fangen wir mit dem ersten Aufgabenteil an!

**Funktion** | Zunächst einmal: Da steht «Schreibe ein Programm ...». In den Lehrsprachen dieses Buchs heißt das immer «Schreibe eine *Funktion* ...». Später werden wir Programme schreiben, die neben der «Hauptfunktion» noch «Hilfsfunktionen» haben, aber das Prinzip bleibt immer das gleiche.

Wenn dort steht, «das den Monatsverbrauch in Kilowattstunden akzeptiert», dann bedeutet dies, dass die Funktion den Monatsverbrauch als Eingabe akzeptieren soll – also als Argument.

Ebenso bedeutet «und den ... zu zahlenden monatlichen Rechnungsbetrag berechnet», dass die Funktion diesen Rechnungsbetrag als Ergebnis liefern soll.

**Gerüst** | Funktionen, die ein Problem aus einer Aufgabenstellung lösen, sollten immer einen Namen haben – den sollten wir uns gleich am Anfang ausdenken. Der Name `billig-strom` bietet sich an. Wir wissen schon, dass die Funktion den Verbrauch in Kilowattstunden akzeptiert. Wenn wir uns für Kilowattstunden auch noch einen Namen ausdenken, können wir direkt ein unfertiges Programm hinschreiben, ohne groß nachzudenken:

```
(define billig-strom
  (lambda (kWh)
    ...))
```

*Wichtig:* Schreibe dieses *Gerüst* unbedingt hin, auch wenn Du noch keine Vorstellung hast, wie es danach weitergeht. Immer. Jedes Mal.

**Rumpf** | Für das Berechnen des Rechnungsbetrags können wir eine Formel aufschreiben: Er ist die Summe aus der Grundgebühr und dem Produkt aus Energie (in Kilowattstunden) und dem Preis einer Kilowattstunde. Entsprechend müssen wir den Rumpf der Funktion ergänzen:

```
(define billig-strom
  (lambda (kWh)
    (+ 4.90 (* 0.19 kWh))))
```

Die Funktion `billig-strom` kannst Du jetzt aufrufen:

```
(billig-strom 10)
↪ 6.8
(billig-strom 20)
↪ 8.7
(billig-strom 30)
↪ 10.6
```

Die Funktion `billig-strom` ist ein einigermaßen winziges Programm. Wenn Programme größer werden, ist häufig nicht mehr unmittelbar ersichtlich, was die Programmteile tun oder ob sie tatsächlich korrekt sind. Es ist darum sinnvoll, die Funktion mit einigen Programmelementen zu ergänzen, welche die Verständlichkeit erhöhen und auch die Wahrscheinlichkeit, dass die Funktion korrekt ist. In den nächsten Abschnitten nehmen wir uns diese Programmelemente einzeln vor.

## 1.8 KURZBESCHREIBUNG

Als erste Unterstützung der Lesbarkeit stellen wir einer Funktion eine kurze Beschreibung auf Deutsch voran, die sogenannte *Kurzbeschreibung*, die beschreibt, was die Funktion macht. Eine prägnante Zeile ist dafür genug. Das könnte so aussehen:

```
; monatlichen Rechnungsbetrag für Tarif Billig-Strom berechnen

(define billig-strom ...)
```

Ein Semikolon `;` kennzeichnet einen *Kommentar*. Der Kommentar erstreckt sich vom Semikolon bis zum Ende der Zeile und wird von DrRacket ignoriert.

**Abbildung 1.9:** Kommentare

Das Semikolon am Anfang kennzeichnet die Zeile als *Kommentar*. DrRacket ignoriert diese Zeile beim Auswerten, aber für menschliche Leserinnen ist sie nützlich. Abbildung 1.9 fasst zusammen, wie das Semikolon funktioniert.

## 1.9 SIGNATUR

Die Lesbarkeit des Programms wird außerdem durch die *Signatur* unterstützt, die beschreibt, welche Art von Daten die Funktion als Argumente erwartet und welche sie liefert.

Die Anzahl der Kilowattstunden auf dem Zählerableseblatt ist in der Regel eine ganze Zahl, die nicht negativ sein kann. Die Ausgabe kann hingegen zwei Nachkommestellen enthalten, ist also ein Bruch. Um die Signatur zu schreiben, benutzen wir für die Beschreibung dieser *Sorten* mathematische Namen. Eine ganze, nicht-negative Zahl heißt auch *natürliche Zahl*. Ein Bruch heißt auch *rationale Zahl*. Die Signatur-Deklaration für `billig-strom` sollte gleich nach der Kurzbeschreibung kommen und sieht so aus:

```
(: billig-strom (natural -> rational))
```

Die Signatur `natural` ist für die natürlichen Zahlen zuständig, `rational` für Brüche.

Diese Signatur-Deklaration besteht aus der Form `(: ...)`, dem Namen der Funktion und ihrer Signatur, `(natural -> rational)`. Diese Signatur hat einen Pfeil `->` in der Mitte und steht deshalb für eine Funktion. Links vom Pfeil steht, was die Funktion als Eingabe akzeptiert, rechts vom Pfeil die Ausgabe. Die Signatur-Deklaration ist also so zu lesen: «`Billig-strom` ist eine Funktion, die eine natürliche Zahl als Eingabe akzeptiert und eine rationale Zahl als Ausgabe produziert.»

Abbildungen 1.10 und 1.11 beschreiben das Format von Signaturen und Signatur-Deklarationen genauer.

Eine *Signatur* beschreibt eine Menge oder Sorte von Werten.

Die Signatur `string` beschreibt eine Zeichenkette.

Folgende Signaturen für Zahlen sind eingebaut:

<code>natural</code>	natürliche Zahlen (0, 1, 2 ...)
<code>integer</code>	ganze Zahlen (0, 1, -1, 2, -2, ...)
<code>rational</code>	Brüche beziehungsweise rationale Zahlen
<code>real</code>	reelle Zahlen
<code>number</code>	beliebige Zahlen

Diese Liste bildet einen «Turm»: Jede natürliche Zahl ist auch eine ganze Zahl (da kommen noch die negativen Zahlen hinzu), jede ganze Zahl ist auch eine rationale Zahl (Brüche mit Nenner ungleich 1 kommen hinzu), jede rationale Zahl ist eine reelle Zahl (nicht-rationale Zahlen wie  $\sqrt{2}$  kommen hinzu, und jede reelle Zahl ist auch eine Zahl (komplexe Zahlen kommen hinzu).

Signaturen für Funktionen haben die Form

`(s1 ... sn -> s)`

Die Signaturen  $s_1 \dots s_n$  sind die Signaturen für die Argumente der Funktion und  $s$  die Signatur für das Ergebnis.

**Abbildung 1.10:** Signaturen

Eine *Signatur-Deklaration* für eine Funktion hat folgende Form:

`(: f (s1 ... sn -> s))`

Sie sagt aus, dass das Verhalten der Funktion der Signatur entspricht: Wenn die Eingaben der Funktion  $f$  den Signaturen  $s_1 \dots s_n$  entspricht, dann entspricht die Ausgabe der Signatur  $s$ .

**Abbildung 1.11:** Signatur-Deklarationen

## 1.10 TESTS

Wir haben oben in der REPL die Funktion `billig-strom` ausprobiert, indem wir DrRacket mehrere Aufrufe mit unterschiedlichen Eingaben haben auswerten lassen. Anhand dieser Beispiele können wir auch (vielleicht mit einem Taschenrechner) überprüfen, ob die Funktion dabei korrekte Ergebnisse produziert hat.

Diese Beispielaufäufe sind noch hilfreicher für Leserinnen und Leser, wenn sie im Programm stehen. Gepaart mit den erwünschten Ergebnissen erleichtern sie das Verständnis. Im Programm schreiben wir das folgendermaßen, gleich nach der Signatur:

`(check-expect (billig-strom 10) 6.8)`

`(check-expect (billig-strom 20) 8.7)`

Ein *Test* hat die folgende Form:

```
(check-expect t e)
```

Wenn das Programm läuft, wertet DrRacket den Test-Ausdruck *t* aus und überprüft, ob er mit dem Wert des Ausdrucks *e* übereinstimmt. Wenn nicht, schlägt der Test fehl und wird von DrRacket protokolliert.

**Abbildung 1.12:** Tests

```
(check-expect (billig-strom 30) 10.6)
```

Wenn wir jetzt das Programm laufen lassen, steht in der REPL:

Alle 3 Tests bestanden!

Das neue Programmelement `check-expect` (siehe Abbildung 1.12) macht nämlich einen sogenannten *Testfall* oder *Test*. In diesem Falle meldet DrRacket, dass mit den Tests alles geklappt hat. Aber das Programm könnte einen Fehler enthalten. Wenn wir in der Definition von `billig-strom` statt `4.90` «versehentlich» `5.90` schreiben, sieht die Ausgabe anders aus:

Check-Fehler:

Der tatsächliche Wert 7.8 ist nicht der erwartete Wert 6.8.

in stromtarif.rkt, Zeile 5, Spalte 0

Der tatsächliche Wert 9.7 ist nicht der erwartete Wert 8.7.

in stromtarif.rkt, Zeile 6, Spalte 0

Der tatsächliche Wert 11.6 ist nicht der erwartete Wert 10.6.

in stromtarif.rkt, Zeile 7, Spalte 0

(Wir haben die den Code unter dem Namen `stromtarif.rkt` abgespeichert, darum taucht dieser Dateiname in den Fehlermeldungen auf.

Die Tests weisen uns also auf mögliche Fehler hin. Das ist so wertvoll, dass wir in Zukunft *immer* Tests schreiben werden, und zwar nach der Signatur. Insbesondere werden wir die Tests schreiben, *bevor* wir die Definition angehen. Das nervt manchmal, ist aber psychologisch sinnvoll, damit wir uns vorher überlegen wie das richtige Verhalten der Funktion aussehen soll. Wenn wir die Tests hinterher schreiben, ist die Versuchung groß, die Funktion einfach in der REPL aufzurufen und das tatsächliche Ergebnis in den Test zu übernehmen, ohne es auf Korrektheit zu überprüfen.

Bei den drei Tests oben müssten wir vielleicht, wenn wir sie vorher schrieben, einen Taschenrechner bemühen. Es gibt aber fast immer mindestens einen möglichen Test, der ganz einfach ist. In diesem Fall ist das:

```
(check-expect (billig-strom 0) 4.9)
```

Hier können wir das Ergebnis direkt der Tabelle entnehmen. Dieser überprüft zwar nicht, ob der richtige pro-kWh-Preis berechnet wird, ist aber so einfach zu schreiben, dass wir die Gelegenheit nicht auslassen sollten, ihn zu schreiben.

Wieviele Tests sollte man schreiben? Das hängt von der Komplexität der Funktion ab. Aber drei sollten es bei einfachen Funktionen mindestens sein, bei komplizierteren mehr. Ein einfaches weiteres Kriterium ist, dass jeder Teil des Programms, das wir geschrieben haben, im Laufe der Tests einmal ausgewertet werden sollte. DrRacket kann Dir dabei helfen.

Wenn Du das Programm nur auf die Funktionsdefinition reduzierst (also alle Tests und Beispielaufrufe entfernst) und dann laufen lässt, sollte Dir auffallen, dass der Rumpf von `billig-strom` blau hinterlegt ist. (Außerdem Teile der Signatur von `billig-strom`.) Diese Hinterlegung zeigt die sogenannte *Abdeckung* an: Die blau hinterlegten Programmteile sind noch nie gelaufen und wurden ergo auch noch nicht getestet. Es fehlen also noch Tests. In Zukunft achten wir also darauf, dass, nachdem das Programm gelaufen ist, nichts blaues übrigbleibt – andernfalls müssen wir mehr Tests schreiben. Wir kommen in Abschnitt 2.3 auf dieses Thema zurück.

Bei Programmen, die Bilder erzeugen, ist es in der Regel nicht möglich, die Testfälle vorher zu schreiben. Du kannst aber Beispiele ausprobieren, die Du Dir vorher überlegt hast, und diese visuell überprüfen. Wenn Dir ein Ergebnis richtig erscheint, kannst Du das Bild aus der REPL kopieren in das Programm kopieren. In der REPL könnte das so aussehen:

```
(square 40 "solid" "red")
```



Du kannst dann das Quadrat aus der REPL kopieren, also zum Beispiel mit der Maus markieren und dann mit `Ctrl-C` beziehungsweise `Strg-C` oder `Command-C` kopieren und dann ins Programm mit `Ctrl-V` beziehungsweise `Strg-V` oder `Command-V` einfügen und dann daraus folgenden Test machen:

```
(check-expect (square 40 "solid" "red") )
```

#### AUFGABE 1.10

Schreibe für `tile` zwei Tests!



Übrigens ist eine Signatur-Deklaration auch so eine Art Testfall. Das sehen wir, wenn wir absichtlich einen Fehler wie diesen hier machen:

```
(check-expect (billig-strom "30") 13.0)
```

Probier es aus! Es erscheint zusammen mit den Meldungen über die Tests eine Meldung über eine *Signaturverletzung* wie folgt:

**Signaturverletzungen:**

bekam "30" in stromtarif.rkt ..., Signatur in stromtarif.rkt ...

In der Meldung stehen klickbare Links, die Dir zeigen, wo die Signaturverletzung passiert ist und welche Signatur verletzt wurde.

Wenn Signaturverletzungen erscheinen, solltest Du diese genau lesen, weil sie auf ein Problem in Deinem Programm hinweisen und wertvolle Hinweise zu dessen Lösung geben.

## 1.11 KONSTRUKTIONSANLEITUNGEN

Wir haben im vorigen Abschnitt folgende Arbeitsschritte kennengelernt, die zu einer Funktion führen:

- Gerüst
- Kurzbeschreibung
- Signatur-Deklaration
- Tests
- Rumpf

Diese Reihenfolge ist der Reihenfolge ihrer Einführung in diesem Kapitel geschuldet. Zukünftig ist es sinnvoll, dass wir diese Schritte immer in der gleichen und zwar in folgender Reihenfolge durchführen:

- Kurzbeschreibung
- Signatur
- Tests
- Gerüst
- Rumpf

Dies entspricht auch der Reihenfolge, in der die entsprechenden Elemente im Programm stehen: Wenn Du diese Reihenfolge befolgst, kannst Du Dein Programm einfach «herunterschreiben».

Eine solche Anleitung zur Konstruktion von Programmen nennen wir in diesem Buch *Konstruktionsanleitung*. Da noch viele weitere Konstruktionsanleitungen hinzukommen werden, bekommen sie Nummern. Die erste Konstruktionsanleitung legt den Ablauf fest und bekommt darum die besondere Nummer 0. Wir fügen gleich noch zwei weitere Elemente hinzu – Datenanalyse und Schablonen – die wir in späteren Kapiteln erläutern werden.



### KONSTRUKTIONSANLEITUNG 0 (ABLAUF)

Gehe bei der Konstruktion einer Funktion in folgender Reihenfolge vor:

1. Kurzbeschreibung
2. Datenanalyse
3. Signatur-Deklaration
4. Tests
5. Gerüst
6. Schablonen
7. Rumpf

Für jeden dieser Schritte gibt es jeweils eine Konstruktionsanleitung, für manche mehrere je nach Situation. Es mag Dir übermäßig bürokratisch vorkommen, immer die gleiche Reihenfolge einzuhalten und immer nach der jeweiligen Konstruktionsanleitung vorzugehen. Bürokratisch ist das in jedem Fall, ist aber eine wertvolle Hilfestellung, die verhindert, dass Du in unnötige Sackgassen gerätst. Wenn Du irgendwann beim Lösen einer Aufgabe in eine Sackgasse gerätst, solltest Du als erstes überprüfen, ob Du auch nach den Konstruktionsanleitungen vorgegangen bist. Oft ist die Ursache des Problems, dass einer der Schritte fehlt.

Zunächst einmal führen wir Konstruktionsanleitungen für die Schritte des Ablaufs ein, die wir schon auf der Liste haben:

### KONSTRUKTIONSANLEITUNG 1 (KURZBESCHREIBUNG)

Schreibe für die Funktion zunächst einen Kommentar, der ihren Zweck kurz beschreibt. Ein Satz, der auf eine Zeile passen sollte, reicht. Beispiel:

```
; monatlichen Rechnungsbetrag für Tarif Billig-Strom berechnen
```

### KONSTRUKTIONSANLEITUNG 2 (SIGNATUR-DEKLARATION)

Schreibe für die Funktion direkt unter die Kurzbeschreibung eine Signatur-Deklaration. Dazu denke Dir zunächst einen möglichst aussagekräftigen Namen aus. Überlege dann, welche Sorten die Ein- und Ausgaben haben und schreibe dann eine Signatur, welche die Ein- und Ausgaben der Funktion möglichst präzise beschreiben. Beispiel:

```
(: billig-strom (natural -> rational))
```

Achte bei den Zahlen-Signaturen besonders auf eine möglichst präzise Signatur. Bei `billig-`

`strom` wäre auch die Signatur `(number -> number)` korrekt, aber nicht so genau.

### KONSTRUKTIONSANLEITUNG 3 (TESTS)

Schreibe unter die Signatur Tests für die Funktion. Denke Dir dafür möglichst einfache, aber auch möglichst interessante Beispiele für Aufrufe der Funktion aus und lege fest, was dabei herauskommen soll. Mache aus den Beispielen Tests mit `check-expect`. Beispiel:

```
(check-expect (billig-strom 0) 4.9)
(check-expect (billig-strom 10) 6.8)
(check-expect (billig-strom 20) 8.7)
(check-expect (billig-strom 30) 10.6)
```

Achte darauf, dass die Tests dafür sorgen, dass der Code Deiner Funktion durch die Tests vollständig abgedeckt wird.

Der letzte Satz der Konstruktionsanleitung ergibt, was das Beispiel `billig-strom` betrifft, keinen Sinn, da jeder einzelne der Tests bereits dafür sorgt, dass die Addition und die Multiplikation in `billig-strom` ausgewertet werden. In Abschnitt 2.3 werden wir aber Beispiele programmieren, bei denen es nicht selbstverständlich ist, dass die Tests alles auswerten.

### KONSTRUKTIONSANLEITUNG 4 (GERÜST)

Schreibe unter die Tests ein Gerüst für die Funktion. Dazu übernimmst Du den Namen aus der Signatur-Deklaration in eine Funktionsdefinition wie zum Beispiel:

```
(define billig-strom
  (lambda (...)
    ...))
```

Denke Dir Namen für die Eingaben der Funktion aus. Das müssen genauso viele sein, wie die Signatur Eingaben hat. Schreibe dann diese Namen als Eingaben in die `lambda`-Abstraktion. Beispiel:

```
(define billig-strom
  (lambda (kWh)
    ...))
```

Für den Rumpf, schreibe zunächst drei Punkte ... (auch genannt *Ellipse*) als Erinnerung daran, dass Du hier später noch etwas ausfüllen musst.

#### KONSTRUKTIONSANLEITUNG 5 (RUMPF)

Als letzten Schritt fülle mit Hilfe des Wissens über das Problem den Rumpf der Funktion aus.

```
(define billig-strom
  (lambda (kWh)
    (+ 4.90 (* 0.19 kWh))))
```

Diesen Ablauf demonstrieren wir anhand des zweiten Teils der Stromtarif-Aufgabe. Zur Erinnerung: Es ging um den Tarif «Watt für Wenig» mit einem Grundbetrag von 8,20 € und einem Verbrauchspreis pro kWh von 0,16 €.

**Kurzbeschreibung** | Fangen wir mit der Kurzbeschreibung an. Die können wir entsprechend zu «Billig-Strom» formulieren:

; monatlichen Rechnungsbetrag für Tarif Watt für wenig berechnen

**Signatur** | Auch die Signatur-Deklaration können wir analog zu «Billig-Strom» formulieren – wir müssen uns nur einen neuen Namen aussuchen:

```
(: watt-für-wenig (natural -> rational))
```

**Tests** | Die Tests müssen wir neu formulieren. Es ist immer sinnvoll, mit dem einfachsten Beispiel anzufangen:

```
(check-expect (watt-für-wenig 0) 8.2)
```

Ansonsten nehmen wir wieder progressiv größere Verbrauchswerte und rechnen den Rechnungsbetrag im Kopf aus:

```
(check-expect (watt-für-wenig 10) 9.8)
(check-expect (watt-für-wenig 20) 11.4)
(check-expect (watt-für-wenig 30) 13.0)
```

**Gerüst** | Das Gerüst ergibt sich direkt aus der Signatur:

```
(define watt-für-wenig
  (lambda (kWh)
    ...))
```

**Rumpf** | Schließlich müssen wir noch den Rumpf vervollständigen, indem wir die entsprechende Formel hineinschreiben:

```
(define watt-für-wenig
  (lambda (kWh)
    (+ 8.20 (* 0.16 kWh))))
```

Fertig!

## 1.12 DIE MACHT DER ABSTRAKTION

Wir haben bei den Stromtarifen für die beiden Aufgabenteile völlig voneinander unabhängige Lösungen geschrieben. Diese unterscheiden sich allerdings nur in Details – in beiden Fällen wird der Stromtarif aus einer Grundgebühr und einem Verbrauchspreis pro kWh mit der gleichen Formel berechnet. Die Definitionen ähneln sich dementsprechend:

```
(define billig-strom
  (lambda (kWh)
    (+ 4.90 (* 0.19 kWh))))

(define watt-für-wenig
  (lambda (kWh)
    (+ 8.20 (* 0.16 kWh))))
```

Diese «Verdoppelung» ist unbefriedigend und kann auch später Probleme machen: Wenn ein Fehler bekannt wird, müssen möglicherweise zwei Funktionen korrigiert werden.

Es wäre also gut, wenn wir die Gemeinsamkeiten der beiden Funktionen irgendwie zusammenfassen könnten, mithin über `billig-strom` und `watt-für-wenig` *abstrahieren* könnten. Dazu kopieren wir die Funktion ein letztes Mal und benennen sie um. Außerdem ersetzen wir alle Stellen, bei denen sich die beiden Funktionen unterscheiden, jeweils durch eine Variable, zum Beispiel `grundgebühr` und `pro-kWh`:

```
(define stromtarif-rechnungsbetrag
  (lambda (kWh)
    (+ grundgebühr (* pro-kWh kWh))))
```

Die beiden neuen Variablen sind noch nicht gebunden, wir müssen sie zu den Parametern des `lambda` hinzufügen:

```
(define stromtarif-rechnungsbetrag
  (lambda (grundgebühr pro-kWh kWh)
    (+ grundgebühr (* pro-kWh kWh))))
```

Wir ergänzen noch Kurzbeschreibung und Signatur. Die neuen Parameter sind auch beides rationale Zahlen, das sieht also so aus:

```
; monatlichen Rechnungsbetrag für Stromtarif berechnen
(: stromtarif-rechnungsbetrag (rational rational natural -> rational))
```

Außerdem sollten wir Tests formulieren. Diese können wir aus den Tests für `billig-strom` und `watt-für-wenig` abschreiben. Wir müssen nur jeweils zu den Argumenten noch die Grundgebühr beziehungsweise den Preis pro kWh hinzufügen:

```
(check-expect (stromtarif-rechnungsbetrag 4.90 0.19 0) 4.9)
; Billig-Strom
(check-expect (stromtarif-rechnungsbetrag 4.90 0.19 10) 6.8)
; Billig-Strom
(check-expect (stromtarif-rechnungsbetrag 4.90 0.19 20) 8.7)
; Billig-Strom
(check-expect (stromtarif-rechnungsbetrag 4.90 0.19 30) 10.6)
; Billig-Strom

(check-expect (stromtarif-rechnungsbetrag 8.20 0.16 0) 8.2)
; Watt für wenig
(check-expect (stromtarif-rechnungsbetrag 8.20 0.16 10) 9.8)
; Watt für wenig
(check-expect (stromtarif-rechnungsbetrag 8.20 0.16 20) 11.4)
; Watt für wenig
(check-expect (stromtarif-rechnungsbetrag 8.20 0.16 30) 13.0)
; Watt für wenig
```

Schließlich können wir auch die Definitionen von `billig-strom` und `watt-für-wenig` so ändern, dass sie nicht mehr «selbst» den Rechnungsbetrag ausrechnen, sondern dafür `stromtarif-rechnungsbetrag` benutzen:

```
(define billig-strom
  (lambda (kWh)
    (stromtarif-rechnungsbetrag 4.90 0.19 kWh)))

(define watt-für-wenig
  (lambda (kWh)
    (stromtarif-rechnungsbetrag 8.20 0.16 kWh)))
```

Zu dieser Technik werden wir in diesem Buch noch oft greifen. Sie erspart nicht nur Arbeit und macht unsere Programme leichter zu handhaben, sondern führt manchmal zu ganz neuen Erkenntnissen – zu sehen zum Beispiel in Kapitel 9.

## 1.13 MANTRAS FÜR DIE PROGRAMMIERUNG

Nach den einfachen Programmen bisher werden wir in den folgenden Kapiteln immer interessantere Funktionen schreiben und dabei weitere Konstruktionsanleitungen für spezifische Situationen einführen.

Es gibt aber einige übergreifende Prinzipien des Programmierens, die wir immer wieder aufgreifen werden. Nicht alle sind jetzt schon relevant, aber wir listen sie hier einmal auf, damit sie an einer Stelle gesammelt sind und wir darauf in den folgenden Kapiteln Bezug nehmen können.

Das erste Mantra soll einem Phänomen entgegenwirken, dass wir oft bei unseren Studierenden beobachtet haben, die manchmal bei einer Aufgabe aufgeben, bevor sie alles hingeschrieben haben, was sie wissen.

### MANTRA 1

Achte darauf, dass für alles, was Du weißt, tatsächlich auch etwas im Programm steht.

Das nächste Mantra klingt eigentlich wie Kinderkram:

### MANTRA 2

Lies, was dort steht.

Das genaue Lesen fällt uns aber erstaunlich schwer. Für dieses Buch sind zwei Aspekte des Lesens besonders wichtig. Am wichtigsten ist das Lesen der jeweiligen Aufgabenstellung, und zwar Schritt für Schritt – gerade so, dass wir genug für den jeweils nächsten Schritt verstehen – anstatt alles auf einmal. Außerdem gibt DrRacket (oder andere Programmierumgebungen) oft nützliche Rückmeldungen, die es sich zu lesen lohnt, auch wenn das manchmal mühsam ist.

Die Konstruktionsanleitungen geben Hilfestellung dabei, die Aufgabenstellung strukturiert zu lesen und in Deinem Programm zu schreiben, was Du weißt. Daraus folgt das nächste Mantra:

#### MANTRA 3

Gehe nach den Konstruktionsanleitungen vor. Wenn Du Dich in einer Sackgasse wiederfindest, überprüfe, ob Du alle relevanten Konstruktionsanleitungen angewendet hast.

Die Konstruktionsanleitungen helfen uns, Fortschritte zu machen, auch wenn wir noch nicht alles verstehen, zu lesen, was dort steht und aufzuschreiben, was wir wissen.

Das folgende Mantra haben wir schon in Abschnitt 1.12 auf Seite 34 kennengelernt:

#### MANTRA 4

Wenn Du zwei Programmteile siehst, die sich nur an wenigen Stellen unterscheiden und die inhaltlich verwandt sind, abstrahiere!

Bisher haben wir nur kleine Programme mit ein, zwei oder drei Funktionen geschrieben: Größere Programme bestehen aus vielen Funktionen, und dieses Mantra benennt das Prinzip, nachdem wir es in Funktionen aufteilen:

#### MANTRA 5

Wenn in Deinem Programm verschiedene Größen berechnet werden, schreibe für die Berechnung jeder Größe eine separate Funktion.

Wir werden oft mathematische Einsichten in Form von Gleichungen benutzen, um unsere Programme zu verbessern:

#### MANTRA 6

Nutze Gleichungen aus, um Dein Programm zu vereinfachen.

Manchmal macht es Spaß, ein kompliziertes Programm zusammenzubasteln wie ein Puzzle. Leider führt dies oft zu unerwünschter Komplexität: Komplexe Programme sind fehleranfällig, schwierig zu verstehen und schwierig weiterzuentwickeln. Wir sollten uns davor hüten:

#### MANTRA 7

Wenn Deine Funktion Dir kompliziert scheint, ist es wahrscheinlich, dass sie Fehler enthält: Entweder, weil sie eigentlich einfacher sein sollte, oder weil die Aufgabe kompliziert und damit ihre Lösung fehleranfällig ist.

Das nächste Mantra beschreibt, wie Du besonders gute Repräsentationen entwickeln kannst. Worum es genau geht, steht in Kapitel 5 auf Seite 131.

#### MANTRA 8

Suche nach Kombinatoren für Deine Daten, die aus kleinen Dingen größere Dinge machen – und daraus wieder größere Dinge machen.

## AUFGABEN

### AUFGABE 1.11

Sei  $C$  eine Temperatur in Grad Celsius. Die entsprechende Temperatur  $F$  in Grad Fahrenheit ergibt sich als:

$$F = \frac{9}{5}C + 32$$

1. Schreibe eine Funktion `celsius->fahrenheit`, die als Eingabe eine Temperatur in °C akzeptiert und eine Temperatur in °F zurückgibt. Benutze dazu die Konstruktionsanleitung.
2. Schreibe eine Funktion `fahrenheit->celsius`, die als Eingabe eine Temperatur in °F akzeptiert und eine Temperatur in °C zurückgibt. Benutze dazu ebenfalls die Konstruktionsanleitung!

### AUFGABE 1.12

Vervielfachung von Zeichenketten:

- Schreibe eine Funktion `double-string`, die eine Zeichenkette akzeptiert und diese «verdoppelt», das heißt, für eine Eingabe `"Sperber"` den Rückgabewert `"SperberSperber"` liefert.



- Schreibe eine Funktion `quadruple-string`, die eine Zeichenkette akzeptiert und «vervielfacht».
- Schreibe eine Funktion `octuple-string`, die eine Zeichenkette akzeptiert und «veracht-facht».
- Schreibe eine Funktion `sixteentuple-string`, die eine Zeichenkette akzeptiert und «ver-sechzehnfacht».

### AUFGABE 1.13

Die *Internetzeit* ist ein alternatives Konzept der Zeitmessung, bei dem der Tag nicht in 24 Stunden à 60 Minuten à 60 Sekunden eingeteilt wird, sondern in 1000 sogenannte *.beats*. Ein solcher *.beat* ist damit 1 Minute 26 Sekunden und 4 Zehntelsekunden lang, zur weiteren Unterteilung sind Nachkommawerte erlaubt. Notiert wird die Zeit durch ein @-Symbol vor dem Wert. @o ist gleichbedeutend mit o Uhr mitteleuropäischer Zeit.

Schreibe eine Funktion `time->beats`, die aus einer Uhrzeit aus Stunden, Minuten, Sekunden und Zehntelsekunden die Internetzeit in *.beats* berechnet. Die Funktion bekommt vier Eingaben: Stunden, Minuten, Sekunden und Zehntelsekunden.

Benutze Abstraktion und schreibe Hilfsfunktionen an den Stellen, an denen Sie Teilprobleme lösen!

### AUFGABE 1.14

In den USA und in Europa gibt es unterschiedliche Maße für die Energieeffizienz von Kraftfahrzeugen:

- In Europa ist das gängige Maß der *Verbrauch* in Liter pro 100 km (l/100km);
- in den USA ist das gängige Maß die *Reichweite* in Meilen pro Gallone (mi/gal).

Schreibe Funktionen, die zwischen beiden Maßeinheiten umrechnen. Gehe dazu wie folgt vor:

Halte Dich bei jeder Funktion, die Du schreibst, an den Ablauf: Schreibe zuerst die Kurzbeschreibung und die Signatur. Schreibe als Nächstes einige Testfälle. Leite danach das Gerüst von der Signatur her und vervollständige den Rumpf der Funktion.

1. Schreibe eine Funktion `liters-per-hundred-kilometers`, die eine Menge Benzin in Liter und die Reichweite dieses Benzins in Kilometer akzeptiert und daraus den Verbrauch in Liter pro 100 km berechnet.
2. Schreibe eine Funktion `miles-per-gallon`, die eine Entfernung in Meilen und den Benzinverbrauch auf diese Entfernung in Gallonen akzeptiert und daraus die Reichweite in Meilen pro Gallone berechnet.

3. Definiere eine Konstante `kilometers-per-mile` (eine US-Meile entspricht etwa 1,61 Kilometer) und schreibe zwei Funktionen `kilometers->miles` und `miles->kilometers`, die jeweils eine Entfernung in einer Maßeinheit akzeptieren und die Entfernung in die jeweils andere Maßeinheit umrechnen.
4. Definiere eine Konstante `liters-per-gallon` (eine Gallone entspricht etwa 3,79 Liter) und schreibe zwei Funktionen `liters->gallons` und `gallons->liters`, die jeweils eine Menge in einer Maßeinheit akzeptieren und die Menge in die jeweils andere Maßeinheit umrechnen.
5. Schreibe die Funktion `l/100km->mi/gal`, die einen Verbrauch in Liter pro 100 km akzeptiert und in die Reichweite in Meilen pro Gallone umrechnet. Benutze dafür die Funktionen, die Du in den anderen Teilaufgaben erstellt hast. Solltest Du auf weitere Teilprobleme stoßen, abstrahiere diese Teilprobleme in eigene Funktionen.
6. Schreibe die Funktion `mi/gal->l/100km`, die eine Reichweite in Meilen pro Gallone akzeptiert und in den Verbrauch in Liter pro 100 km umrechnet. Benutze dafür die Funktionen, die Du in den anderen Teilaufgaben erstellt hast. Solltest Du auf weitere Teilprobleme stoßen, abstrahiere diese Teilprobleme in eigene Funktionen.
7. Finde heraus, wie hoch der Benzinverbrauch verschiedener Kraftfahrzeuge ist, die Du täglich im Straßenverkehr in Deutschland siehst. Vergleiche diesen Verbrauch mit den Reichweitenangaben typischer Kraftfahrzeuge für den US-amerikanischen Markt.

#### AUFGABE 1.15

In einem Betrieb wird eine Stechuhr aufgebaut, die bei Arbeitsbeginn und bei Arbeitsende den aktuellen Zeitpunkt auf eine Karte stempelt. Ein Zeitpunkt wird durch Stunden, Minuten und Sekunden beschrieben. Durch geschickte Verhandlungen des Betriebsrats sind Nachtschichten verboten. Arbeitsbeginn ist frühestens um 6 Uhr morgens und Arbeitsende ist spätestens um 22 Uhr. Jeder Arbeiter in diesem Betrieb verdient pro Stunde 72 €.

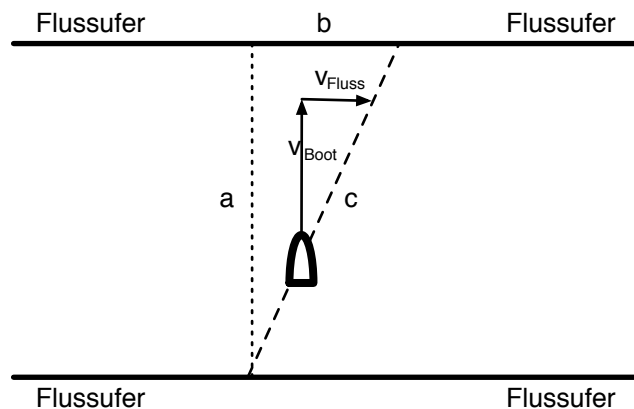
Schreibe eine Funktion `money-earned`, welche aus zwei gegebenen Zeitpunkten (Arbeitsbeginn und Arbeitsende) berechnet, wieviel Geld in dieser Zeit verdient wurde. Es genügt, wenn der Betrag in Cent angegeben wird. Rechne die Uhrzeiten hierfür zunächst in Sekunden um. Die Funktion dafür hat sechs Eingaben: Stunde-Beginn, Minute-Beginn, Sekunde-Beginn, Stunde-Ende, Minute-Ende, Sekunde-Ende.

Benutze Abstraktion und schreibe Hilfsfunktionen an den Stellen, an denen Sie Teilprobleme lösen!

## AUFGABE 1.16

(etwas schwieriger, mit Physik)

Ein Boot überquert einen Fluss mit Strömung und kommt durch die Strömung vom geplanten Kurs ab. Dadurch wird die Strecke, die das Boot tatsächlich zurücklegt, länger.



Gegeben ist die Breite des Flusses  $a$ , die Strömungsgeschwindigkeit des Flusses  $v_{\text{Fluss}}$  und die Geschwindigkeit des Bootes  $v_{\text{Boot}}$ . Berechne die Länge der Strecke, die das Boot tatsächlich zurücklegt. Programmieren Sie dazu Funktionen, die folgende Teilprobleme lösen:

1. Schreiben Sie eine Funktion `crossing-time`, welche die Zeit berechnet, die für das Überqueren des Flusses benötigt wird:

$$t = \frac{a}{v_{\text{Boot}}}$$

2. Schreiben Sie dann eine Funktion `other-shore-offset`, die die Länge der Strecke berechnet, die das Boot abgetrieben wird (also den Versatz am anderen Ufer, im Schaubild die Strecke  $b$ ).
3. Um  $c$  zu berechnen, brauchst Du den *Satz des Pythagoras*:

$$a^2 + b^2 = c^2$$

Schreiben Sie eine Funktion `hypotenuse`, die  $c$  der obigen Gleichung berechnet.

Um die Quadratwurzel einer Zahl zu berechnen, kannst du `sqrt` verwenden, die folgende Signaturdeklaration hat:

```
(: sqrt (number -> number))
```

Solltest Du auf weitere Teilprobleme stoßen, abstrahiere diese Teilprobleme in eigene Funktionen.

4. Schreibe schließlich eine Funktion `boat-travel-distance`, die die tatsächliche Strecke berechnet, die das Boot zurücklegt. Benutze dafür die bisher geschriebenen Funktionen.

## 2 FALLUNTERSCHIEDUNGEN UND VERZWEIGUNGEN

Computerprogramme müssen bei manchen Daten, die sie verarbeiten, zwischen verschiedenen Möglichkeiten differenzieren: Ist die Wassertemperatur warm genug zum Baden? Welche von fünf Tupperschüsseln ist für eine bestimmte Menge Kartoffelsalat groß genug? Welches ist die richtige Abzweigung nach Dortmund? Solche Entscheidungen sind daran festgemacht, dass ein Wert zu einer von mehreren verschiedenen Kategorien gehören kann – es handelt sich dann um eine sogenannte *Fallunterscheidung*; Funktionen operieren auf Daten mit Fallunterscheidung durch *Verzweigungen*.

In diesem Kapitel werden wir zunächst zwei besonders einfache Arten von Fallunterscheidungen behandeln, nämlich *Aufzählungen* und *Zahlenbereiche*. In Kapitel 4 werden wir die Fallunterscheidungen noch einmal aufgreifen und mit sogenannten *gemischten Daten* programmieren.

### 2.1 RECHNEN MIT BOOLESCHEN WERTEN

Für die Programme dieses Kapitels benötigen wir eine neue Art von Daten. Die ergeben sich, wenn wir zum Beispiel zwei Zahlen vergleichen:

`(< 0 5)`

$\mapsto$  `#t`

`(< 0 5)` ist die Schreibweise für  $0 < 5$ . Das `#t` steht für «true» oder «wahr», denn die Aussage «ist 0 kleiner als 5» stimmt ja. Umgekehrt kommt natürlich nicht `#t` heraus:

`(< 5 0)`

$\mapsto$  `#f`

Das `#f` steht für «false» oder «falsch», denn diese Aussage stimmt nicht.

«Wahr» und «falsch» heißen zusammen *boolesche Werte* oder auch *Wahrheitswerte*.<sup>1</sup> Ein Ausdruck, bei dem ein boolescher Wert herauskommt, heißt dementsprechend auch *boolescher Ausdruck*.

---

<sup>1</sup> Die booleschen Werte sind benannt nach George Boole (1815–1864), der als erster einen algebraischen Ansatz für die Behandlung von Logik mit den Werten «wahr» und «falsch» formulierte.

Wir werden boolesche Ausdrücke oft *Bedingungen* nennen. Wenn eine Bedingung **#t** liefert, werden wir auch die Sprachregelung benutzen, dass die Bedingung *gilt* – beziehungsweise, dass, wenn sie **#f** liefert, sie *nicht gilt*.

**<** ist eine eingebaute Funktion, die auf «kleiner als» testet, also dem mathematischen Operator **<** entspricht. Ebenso gibt es auch **>** für «größer als» (Mathematik:  $>$ ), **=** für «gleich» (Mathematik:  $=$ ), **<=** für «kleiner oder gleich» (Mathematik:  $\leq$ ) und **>=** für «größer oder gleich» (Mathematik:  $\geq$ ).

Analog zu **=** für Zahlen können Zeichenketten mit **string=?** verglichen werden:

```
(string=? "Mike" "Mike")
↪ #t
(string=? "Herbert" "Mike")
↪ #f
```

**#t** und **#f** sind Literale, wie Zahlen. Sie können also auch in Programmen stehen:

```
#t
↪ #t
#f
↪ #f
```

Programme können mit booleschen Werten auch rechnen. Ein Ausdruck der Form

```
(and e1 e2 ... en)
```

ergibt immer dann **#t**, wenn alle  $e_i$  **#t** ergeben, sonst **#f**. Bei zwei Operanden  $e_1$  und  $e_2$  ergibt **(and e<sub>1</sub> e<sub>2</sub>)** immer dann **#t**, wenn  $e_1$  und  $e_2$  **#t** ergeben:

```
(and #t #t)
↪ #t
(and #f #t)
↪ #f
(and #t #f)
↪ #f
(and #f #f)
↪ #f
```

Entsprechend gibt es Ausdrücke der Form

```
(or e1 e2 ... en)
```

die immer dann `#t` ergeben, wenn *einer* der  $e_i$  `#t` ergibt, sonst `#f`. Bei zwei Operanden  $e_1$  und  $e_2$  ergibt `(or  $e_1$   $e_2$ )` immer dann `#t`, wenn  $e_1$  *oder*  $e_2$  `#t` ergeben:

```
(or #t #t)
→ #t
(or #f #t)
→ #t
(or #t #f)
→ #t
(or #f #f)
→ #f
```

Das `or` berechnet also ein *inklusive* Oder: `(or  $e_1$   $e_2$ )` bedeutet « $e_1$  oder  $e_2$  oder beides», im Gegensatz zum *exklusiven* Oder «entweder  $e_1$  oder  $e_2$ ».<sup>2</sup>

Des Weiteren gibt es noch eine eingebaute Funktion `not`, die einen booleschen Wert umdreht beziehungsweise dessen *Negation* berechnet, sich also folgendermaßen verhält:

```
(not #f)
→ #t
(not #t)
→ #f
```

Für boolesche Werte gibt es die Signatur `boolean`.

## AUFGABE 2.1

Schreibe eine Funktion, die von zwei booleschen Werten das exklusive Oder berechnet. □

## 2.2 VERZWEIGUNGEN

Um Fallunterscheidungen zu demonstrieren, nehmen wir uns folgende Beispielaufgabe vor: Wir schreiben eine Funktion, die eine Zahl (auf Englisch) «aufsagen» soll. Sie soll sich so verhalten:

```
(say-number 0)
→ "zero"
```

<sup>2</sup> In der Informatik ist mit «oder» ohne Zusatz fast immer das inklusive Oder gemeint. Wer das exklusive Oder meint, sagt auch «exklusives Oder».

```
(say-number 1)
↪ "one"
```

Der Einfachheit halber beschränken wir uns vorläufig auf die Zahlen von Null bis Drei. Die Funktion hat folgende Kurzbeschreibung:

*; Zahl zu Text machen*

Die Funktion macht aus einer natürlichen Zahl eine Zeichenkette und hat folgende Signatur:

```
(: say-number (natural -> string))
```

Die Signatur sagt leider nichts darüber aus, dass die Funktion nur bis Drei funktioniert – später werden wir noch beschreiben, wie die Signatur präziser werden kann. Hier wollen wir uns jedoch erst einmal auf die Funktionsdefinition konzentrieren. Vorher machen wir aber die obigen Beispiele zu Testfällen:

```
(check-expect (say-number 0) "zero")
(check-expect (say-number 1) "one")
```

Das Gerüst ergibt sich direkt aus der Signatur:

```
(define say-number
  (lambda (n)
    ...))
```

Aber wie kommen wir jetzt weiter? Die Eingabe zerfällt ja in vier Fälle – 0, 1, 2 und 3 – sie bildet somit eine *Fallunterscheidung*. Um eine Fallunterscheidung in der Eingabe einer Funktion verarbeiten zu können, benötigen wir ein neues Programmelement, die *Verzweigung*. Verzweigungen beginnen mit dem Wort **cond** und gehören zu den kompliziertesten Programmelementen, die wir in diesem Buch benutzen. Aber keine Sorge, so schlimm wird es nicht. Um eine Verzweigung zu schreiben, müssen wir wissen, *wie viele* Fälle es gibt. In unserer Aufgabe sind das vier. Die Verzweigung dafür hat folgende Form:

```
(define say-number
  (lambda (n)
    (cond
      (... ...)
      (... ...)
      (... ...)
      (... ...))))
```



Jedes der Programmstücke (`... ...`) ist ein sogenannter *Zweig*. Der erste Teil eines Zweigs ist immer eine Bedingung, der für den entsprechenden Fall `#t` liefern sollte und für die anderen Fälle `#f`. In diesem Fall muss die Bedingung jedes Zweiges jeweils eine der Zahlen von Null bis Drei identifizieren:

```
(define say-number
  (lambda (n)
    (cond
      ((= n 0) ...)
      ((= n 1) ...)
      ((= n 2) ...)
      ((= n 3) ...))))
```

Der jeweils zweite Teil des Zweiges ist das gewünschte Ergebnis für den entsprechenden Fall. Wenn wir das ausfüllen, sieht das Resultat so aus:

```
(define say-number
  (lambda (n)
    (cond
      ((= n 0) "zero")
      ((= n 1) "one")
      ((= n 2) "two")
      ((= n 3) "three"))))
```

Fertig!

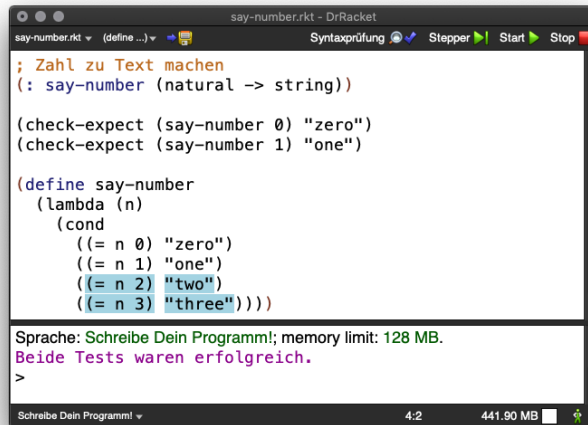
## AUFGABE 2.2

Finde heraus was passiert, wenn Du die Funktion mit einer Zahl aufrufst, für die sie nicht gemacht ist. □

Wie oben schon erwähnt, suggeriert die Signatur von `say-number`, dass sie eine beliebige natürliche Zahl aufsagen kann, auch wenn sie nur bis Drei funktioniert. Glücklicherweise können wir die Signatur präzisieren, und zwar so:

```
(: say-number ((integer-from-to 0 3) -> string))
```

Die Funktion `integer-from-to` liefert also eine Signatur, die natürliche Zahlen zwischen einer Unter- und einer Obergrenze beschreibt.



**Abbildung 2.1:** Anzeige von Abdeckung in DrRacket

## 2.3 ABDECKUNG

Vielleicht ist Dir bei der `say-number`-Funktion aus dem vorigen Abschnitt aufgefallen, dass bei dem Programm unmittelbar, nachdem es gelaufen ist, zwei Zeilen farbig hinterlegt sind, wie in Abbildung 2.1 zu sehen. Das liegt daran, dass diese beiden Zeilen nie ausgewertet wurden, weil die Tests nur die ersten beiden Fälle überprüfen. Theoretisch könnte in den beiden letzten Zweigen für 2 und 3 noch ein Fehler stecken, und die Tests würden es nicht bemerken.

Wenn Du folgende Testfälle hinzufügst, geht die farbige Hinterlegung weg:

```

(check-expect (say-number 2) "two")
(check-expect (say-number 3) "three")

```

Probier es aus – am besten erst nur den einen, dann den anderen!

Die farbige Hinterlegung zeigt die *Abdeckung* an, wie wir es schon in Abschnitt 1.10 auf Seite 29 kurz diskutiert haben: Ein Ausdruck im Programm, der durch die Tests ausgewertet wird, heißt *abgedeckt*. Die farbig hinterlegten Ausdrücke sind also nicht abgedeckt. Der letzte Satz von Konstruktionsanleitung 3 auf Seite 32 sagt genau das: Eine Funktion sollte durch ihre Tests vollständig abgedeckt werden. Die Abdeckung ist also eine Minimalanforderung an unsere Tests.

## 2.4 AUFZÄHLUNGEN

Nächste Beispielaufgabe: Wir wollen Haustiere einteilen in niedliche und nicht so niedliche. Es gibt in der Welt dieser Aufgabe nur drei verschiedene Haustiere: Katzen, Hunde und Schlangen. Haustiere könnten wir in einem Kommentar im Programm so beschreiben:

```
; Ein Haustier ist eins der folgenden:
; - Katze
; - Hund
; - Schlange
```

Solch ein kurzer Kommentar, der die Daten beschreibt, die in unserer Aufgabe vorkommen, heißt *Datendefinition*. Mehr dazu in Abschnitt 2.6. Die Formulierung der Datendefinition «eins der folgenden» deutet daraufhin, dass auch diese Sorte Daten in mehrere Fälle zerfällt – es handelt sich also wieder um eine Fallunterscheidung.

Bevor wir die eigentliche Aufgabe lösen (niedlich oder nicht), müssen wir uns zunächst überlegen, wie wir Haustiere im Programm repräsentieren: Für die Zwecke dieser Aufgabe reicht es zu wissen, mit *welcher* der drei Möglichkeiten wir es zu tun haben. Wir benutzen die einfachste Möglichkeit, Zeichenketten mit dem entsprechenden Text, also "Katze", "Hund" und "Schlange".

Eine solche Auflistung von Werten bildet eine sogenannte *Aufzählung* – ein Spezialfall einer Fallunterscheidung.

Für die Aufzählung der Haustiere gibt es noch keine fertige Signatur, die müssen wir noch definieren. Die *Signatur-Definition* sieht so aus:

```
(define pet
  (signature (enum "Katze" "Hund" "Schlange")))
```

Da sind gleich zwei neue Programmelemente drin:

- Wir schreiben `signature` immer, wenn wir eine neue Signatur erzeugen.
- `enum` (das funktioniert nur innerhalb eines `signature`-Ausdrucks) ist für Aufzählungen zuständig. In einem `enum`-Ausdruck stehen die Werte, die zur Aufzählung gehören.

### AUFGABE 2.3

Die Funktion `say-number` aus dem vorigen Abschnitt hat ja als Signatur-Deklaration die folgende:

```
(: say-number ((integer-from-to 0 3) -> string))
```

Schreibe explizite Datendefinitionen für Ein- und Ausgabe der Funktion und mache ihre Signatur mit Hilfe von Aufzählungen präziser! □

Zurück zu unserer Funktion zur Niedlichkeitsanalyse. Die definierte Signatur `pet` können wir nun in der Signatur-Deklaration unserer Funktion benutzen. Zusammen mit der Kurzbeschreibung sieht das so aus:

```
; Ist Haustier niedlich?
(: cute? (pet -> boolean))
```

Das Fragezeichen gehört zum Namen der Funktion und ist eine Konvention, die wir für die Namen von Funktionen verwenden, die einen booleschen Wert zurückgeben – also solche Funktionen, die eine Ja-/Nein-Frage beantworten.

Da es nur drei Möglichkeiten für die Eingabe gibt, können wir für alle drei jeweils einen Test schreiben:

```
(check-expect (cute? "Katze") #t)
(check-expect (cute? "Hund") #t)
(check-expect (cute? "Schlange") #f)
```

Kommen wir zur Funktion selbst. Zunächst einmal das Gerüst:

```
(define cute?
  (lambda (p)
    ...))
```

Da es sich bei der Eingabe um eine Aufzählung, also eine Fallunterscheidung handelt, brauchen wir eine Verzweigung im Rumpf. Da es drei Fälle in der Aufzählung gibt, braucht die Verzweigung ebenfalls drei Zweige:

```
(define cute?
  (lambda (p)
    (cond
      (... ...)
      (... ...)
      (... ...))))
```

Als Nächstes müssen wir die Bedingungen schreiben, und die sollten `p` jeweils mit `"Katze"`, `"Hund"` und `"Schlange"` vergleichen. Dabei könnte man leicht auf die Idee kommen, `(= p "Katze")` etc. zu schreiben. Dann kommt allerdings eine Fehlermeldung etwa so:

```
=: Zahl als erstes Argument erwartet, "Katze" bekommen
```

Die Funktion `=` fühlt sich also nur für Zahlen zuständig, für Zeichenketten müssen wir die Funktion `string=?` verwenden:

```
(define cute?
  (lambda (p)
    (cond
      ((string=? p "Katze") ...)
      ((string=? p "Hund") ...)
      ((string=? p "Schlange") ...))))
```

Bisher ergibt sich alles rein aus der Definition von `pet`. Schließlich müssen wir noch die Antworten ergänzen. Katzen und Hunde sind niedlich, Schlangen nicht:

```
(define cute?
  (lambda (p)
    (cond
      ((string=? p "Katze") #t)
      ((string=? p "Hund") #t)
      ((string=? p "Schlange") #f))))
```

Fertig!

## 2.5 ZAHLENBEREICHE

Eine weitere häufig vorkommende Art der Fallunterscheidung gibt es bei Zahlenbereichen.

Um das zu demonstrieren, nehmen wir uns folgende Aufgabe vor: Wir schreiben eine Funktion, die «Wasser erhitzt». Das heißt, wir fügen einer bestimmten Menge Wasser bei einer bestimmten Temperatur eine bestimmte Menge Wärme zu. Wir wollen berechnen, wie warm das Wasser danach ist. Der Einfachheit halber messen wir die zugefügte Wärme - genau wie die Temperatur – in Grad Celsius (°C).<sup>3</sup> Wir schreiben dazu drei Versionen der Funktion:

- Die erste, naive Version addiert einfach auf die Anfangstemperatur die hinzugefügte Wärme.
- Die zweite Version berücksichtigt, dass Wasser bei 100°C siedet und nicht heißer werden kann.
- Die dritte Version berücksichtigt zusätzlich, dass gefrorenem Wasser von 0°C eine Wärme von 80°C hinzugefügt werden muss, damit es schmilzt und dann immer noch erst bei 0°C ist.

Wir berücksichtigen also schrittweise, dass die innere Energie des Wassers – die Wärme – nicht dasselbe ist wie dessen Temperatur.

<sup>3</sup> Wir vereinfachen schon sehr stark, aber Temperaturerhöhung und Wärme sind über die Wärmekapazität verwandt.

**Naive Version** | Wir fangen mit der naiven Version an und gehen nach dem Ablauf der Konstruktionsanleitungen vor. Zuerst kommt also eine Kurzbeschreibung:

; Wassertemperatur nach Erhitzen berechnen, naiv

Als Nächstes kommt die Signatur-Deklaration. Die Anfangstemperatur und die hinzugefügte Wärme sind die Eingaben; sie stellen wir als reelle Zahlen dar. Heraus kommt die Endtemperatur, auch eine reelle Zahl. Da wir schon wissen, dass diese Version etwas einfach ist, hängen wir eine `-0` an den Namen:

```
(: heat-water-0 (real real -> real))
```

Nun schreiben wir einige einfache Beispiele als Testfälle hin:

```
(check-expect (heat-water-0 -10 20) 10)
```

```
(check-expect (heat-water-0 10 20) 30)
```

```
(check-expect (heat-water-0 90 20) 110)
```

Als Nächstes kommt das Gerüst an die Reihe:

```
(define heat-water-0
  (lambda (temp heat)
    ...))
```

Der Rumpf ist jetzt kein Hexenwerk, die beiden Eingaben werden einfach addiert:

```
(define heat-water-0
  (lambda (temp heat)
    (+ temp heat)))
```

Noch die Tests laufen lassen und fertig!

**Siedendes Wasser** | Als Nächstes wollten wir berücksichtigen, dass Wasser nicht über  $100^{\circ}\text{C}$  heiß werden kann. Die Kurzbeschreibung passen wir nur leicht an; die Signatur bleibt ebenfalls fast unverändert – wir ändern nur den Namen und machen aus der `0` eine `1`:

; Wassertemperatur nach Erhitzen berechnen, Sieden berücksichtigen

```
(: heat-water-1 (real real -> real))
```

Bei den Tests können wir natürlich die Tests von `heat-water-0` kopieren, aber den letzten müssen wir anpassen:

```
(check-expect (heat-water-1 -10 20) 10)
```

```
(check-expect (heat-water-1 10 20) 30)
```

```
(check-expect (heat-water-1 90 20) 100)
```

Beim Testen ist es immer sinnvoll, auch Grenzfälle zu testen – schaltet die Funktion wirklich um, wenn die 100 erreicht sind.

Dazu dienen folgende zwei Testfälle:

```
(check-expect (heat-water-1 99 1) 100)
(check-expect (heat-water-1 99 2) 100)
```

Da die Signatur die gleiche ist, ist auch das Gerüst identisch:

```
(define heat-water-1
  (lambda (temp heat)
    ...))
```

Ab 100°C muss die Funktion ihr Ergebnis also anders berechnen. Oder, anders gesagt, die Eingaben fallen in zwei verschiedene Klassen: die, bei denen die Summe unter 100 liegt und die, bei denen sie darüber liegt. Wir benötigen also ein `cond` mit zwei Zweigen:

```
(define heat-water-1
  (lambda (temp heat)
    (cond
      (... ...)
      (... ...))))
```

Wir brauchen nun eine Bedingung, die feststellt, ob die Summe aus `temp` und `heat` unterhalb von 100 liegt (genauer: kleiner *oder gleich*, weil 100 gerade so erreichbar ist) und eine dafür, dass die Summe darüber liegt. Die könnten so aussehen:

```
(<= (+ temp heat) 100)
(> (+ temp heat) 100)
```

Wenn wir diese beiden Bedingungen an die entsprechenden Stellen im Rumpf setzen, sieht das so aus:

```
(define heat-water-1
  (lambda (temp heat)
    (cond
      ((<= (+ temp heat) 100) ...)
      (> (+ temp heat) 100) ...))))
```

An die Stellen nach den Bedingungen müssen wir Ausdrücke setzen, die das Ergebnis liefern, das im jeweiligen Fall richtig ist. Das wäre dann:

```
(define heat-water-1
  (lambda (temp heat)
    (cond
      ((<= (+ temp heat) 100) (+ temp heat))
      ((> (+ temp heat) 100) 100))))
```

Fertig!

#### AUFGABE 2.4

Müssen es bei den beiden Bedingungen unbedingt  $\leq$  und  $>$  sein? Was passiert, wenn Du  $\leq$  durch  $<$  ersetzt und das Programm dann laufen lässt? Was passiert, wenn Du dann auch das  $>$  ersetzt – durch  $\geq$ ? Warum funktioniert das Programm dann immer noch? □

**Siendendes Wasser und Eis** | Kommen wir zur «Vollversion». Zur Erinnerung: Da müssen wir noch berücksichtigen, dass gefrorenem Wasser von  $0^\circ\text{C}$  eine Wärme von  $80^\circ\text{C}$  hinzugefügt werden muss, damit es schmilzt und dann immer noch  $0^\circ\text{C}$  hat.

Wir fangen wieder mit der Kurzbeschreibung an:

; Wassertemperatur nach Erhitzen berechnen, mit Eis & Sieden

Da diese Version die letzte ist, hat sie keine Nummer mehr. Ansonsten ist die Signatur unverändert:

```
(: heat-water (real real -> real))
```

Die Tests können wir nicht unverändert übernehmen. Gleich der erste funktioniert nicht mehr:

```
(check-expect (heat-water -10 20) 10)
```

Das Aufwärmen des Wassers von  $-10^\circ\text{C}$  auf  $0^\circ\text{C}$  erfordert nur  $10^\circ\text{C}$  Wärmezufuhr, dann aber müssen erst einmal  $80^\circ$  weitere Wärme zugeführt werden, damit es weiter geht. Den Testfall müssen wir also ändern:

```
(check-expect (heat-water -10 20) 0)
```

Die anderen Testfälle `heat-water-1` können so bleiben:

```
(check-expect (heat-water 10 20) 30)
(check-expect (heat-water 90 20) 100)
(check-expect (heat-water 99 1) 100)
(check-expect (heat-water 99 2) 100)
```



Ein paar weitere Tests sollten aber noch genau klären, was um den Nullpunkt herum so passiert und wo genau er überschritten wird:

```
(check-expect (heat-water -10 5) -5)
(check-expect (heat-water -5 60) 0)
(check-expect (heat-water -5 90) 5)
(check-expect (heat-water -1 81) 0)
(check-expect (heat-water -1 82) 1)
```

Wieder sollten wir darüber nachdenken, in welche Fälle unsere Eingaben zerfallen. Da gibt es drei naheliegende Fälle:

1. Die Anfangstemperatur ist unter 0°C, es wird also Eis erwärmt.
2. Die Erwärmung würde die Wassertemperatur auf über 100°C erhöhen.
3. Alles andere – das Wasser fängt flüssig an und bleibt durch die Erwärmung flüssig.

Der erste Fall hat außerdem noch drei «Unterfälle»:

- Die Erwärmung bleibt unter 0°C.
- Die Erwärmung bleibt bei 0°C «stecken»
- Die Erwärmung erhöht die Temperatur über den Nullpunkt hinaus.

So komplizierte Fallunterscheidungen sind relativ selten. Wenn sie doch einmal auftauchen, ist besondere Sorgfalt gefragt. Deshalb exerzieren wir das hier als Beispiel durch.

Fangen wir wieder mit dem Gerüst für die Funktion an:

```
(define heat-water
  (lambda (temp heat)
    ...))
```

Wir wissen schon aus der Analyse der Fälle, dass es drei Fälle gibt. Deshalb brauchen wir auch wieder ein **cond** mit drei Zweigen.

```
(define heat-water
  (lambda (temp heat)
    (cond
      (... ...)
      (... ...)
      (... ...))))
```

Jetzt ergänzen wir Bedingungen, die den drei Fällen entsprechen. Um in diesem komplizierten Fall die Zuordnung von Fällen und Bedingungen klar erkennbar zu machen, stehen diese jeweils als Kommentar darüber:

```
(define heat-water
  (lambda (temp heat)
    (cond
      ; Die Anfangstemperatur ist unter 0°C, es wird also Eis erwärmt.
      ((< temp 0) ...)
      ; Erwärmung würde Wassertemperatur auf über 100°C erhöhen.
      ((>= (+ temp heat) 100) ...)
      ; Wasser fängt flüssig an und bleibt durch Erwärmung flüssig.
      ((and (>= temp 0) (< (+ temp heat) 100)) ...))))
```

Jetzt können wir uns daran machen, die Zweige auszufüllen. Da der erste Zweig (das Eis) komplizierter ist, schieben wir den erstmal vor uns her. Beim zweiten Zweig kommt einfach 100°C raus. Ebenso einfach ist der dritte Zweig, bei dem das Wasser flüssig bleibt – die Antwort ist dort `(+ temp heat)`. Der Zwischenstand sieht so aus:

```
(define heat-water
  (lambda (temp heat)
    (cond
      ; Die Anfangstemperatur ist unter 0°C, es wird also Eis erwärmt.
      ((< temp 0) ...)
      ; Erwärmung würde Wassertemperatur auf über 100°C erhöhen.
      ((>= (+ temp heat) 100) 100)
      ; Wasser fängt flüssig an und bleibt durch Erwärmung flüssig.
      ((and (>= temp 0) (< (+ temp heat) 100))
       (+ temp heat)))))
```

Dass wir die leichten Fälle zuerst bearbeitet haben, mag wie Faulheit aussehen. Ist es auch – aber es ist auch eine sinnvolle Strategie, die sich aus Mantra 1 ergibt:

MANTRA 1

Achte darauf, dass für alles, was Du weißt, tatsächlich auch etwas im Programm steht.

Auch über den ersten Zweig wissen wir etwas, nämlich dass er selbst eine Fallunterscheidung ist mit drei Fällen. Wir können also das `cond` (wieder einmal) schon hinschreiben:

```
(define heat-water
  (lambda (temp heat)
    (cond
```

```

; Anfangstemperatur ist unter 0°C, es wird also Eis erwärmt.
((< temp 0)
 (cond
  (... ...)
  (... ...)
  (... ...)))
; Erwärmung würde die Wassertemperatur auf über 100°C erhöhen.
(>= (+ temp heat) 100) 100)
; Wasser fängt flüssig an und bleibt durch Erwärmung flüssig.
((and (>= temp 0) (< (+ temp heat) 100))
 (+ temp heat))))

```

Auch hier ist es sinnvoll, die Beschreibungen der Fälle über die Zweige zu schreiben. Danach ergänzen wir die Bedingungen mit folgendem Ergebnis:

```

(define heat-water
 (lambda (temp heat)
  (cond
   ; Erwärmung ist unter 0°C, es wird also Eis erwärmt.
   ((< temp 0)
    (cond
     ; Erwärmung bleibt unter 0°C.
     ((< (+ temp heat) 0) ...)
     ; Erwärmung bleibt bei 0°C "stecken".
     ((and (>= (+ temp heat) 0)
          (< (+ temp heat) 80))
      ...))
   ; Erwärmung erhöht Temperatur über den Nullpunkt hinaus.
   ((and (>= (+ temp heat) 0)
        (>= (+ temp heat) 80))
    ...)))
; Erwärmung würde Wassertemperatur auf über 100°C erhöhen.
(>= (+ temp heat) 100) 100)
; Wasser fängt flüssig an und bleibt durch Erwärmung flüssig.
((and (>= temp 0) (< (+ temp heat) 100))
 (+ temp heat))))

```

Die Bedingungen sind jetzt noch komplizierter, aber erschließen sich hoffentlich durch genauere Betrachtungen:

- «Die Erwärmung bleibt unter 0°C.»  
Das heißt, die Summe von Anfangstemperatur und Erwärmung ist kleiner als 0°C.
- «Die Erwärmung bleibt bei 0°C stecken.»  
Das heißt, die Summe von Temperatur und Erwärmung muss zwischen 0°C und 80°C liegen.
- «Die Erwärmung erhöht die Temperatur über den Nullpunkt hinaus.»  
Das Summe von Temperatur und Erwärmung geht nicht nur über 0°C sondern auch über 80°C hinaus.

Die Antworten unter diesen Bedingungen sind vergleichsweise einfach zu ergänzen:

```
(define heat-water
  (lambda (temp heat)
    (cond
      ; Anfangstemperatur ist unter 0°C, es wird also Eis erwärmt.
      ((< temp 0)
       (cond
         ; Erwärmung bleibt unter 0°C.
         ((< (+ temp heat) 0) (+ temp heat))
         ; Erwärmung bleibt bei 0°C "stecken".
         ((and (>= (+ temp heat) 0)
              (< (+ temp heat) 80))
          0)
         ; Erwärmung erhöht Temperatur über den Nullpunkt hinaus.
         ((and (>= (+ temp heat) 0)
              (>= (+ temp heat) 80))
          (- (+ temp heat) 80))))
      ; Erwärmung würde Wassertemperatur auf über 100°C erhöhen.
      ((>= (+ temp heat) 100) 100)
      ; Wasser fängt flüssig an und bleibt durch Erwärmung flüssig.
      ((and (>= temp 0) (< (+ temp heat) 100))
       (+ temp heat)))))
```

Fertig!<sup>4</sup>

<sup>4</sup> Allerdings noch nicht ganz richtig: Vielleicht siehst Du, dass da noch etwas nicht ganz stimmt. Wir kommen später darauf zurück.

Also *fast* fertig. Wenn wir das Programm näher betrachten, fällt etwas Verbesserungspotenzial auf. Fangen wir an mit der Bedingung:

```
(and (>= (+ temp heat) 0)
      (>= (+ temp heat) 80))
```

Das ist übertrieben: Eine Temperatur über 80°C liegt auch über 0°C. Wir können das also vereinfachen auf:

```
(>= (+ temp heat) 80)
```

Das ist mathematisch einleuchtend, zur Sicherheit sollten wir aber die Tests nochmals laufen lassen: Aber sie laufen alle noch erfolgreich durch.

Bevor wir die Funktion noch weiter vereinfachen, ist es sinnvoll, die Funktionsweise der Auswertung von **cond**-Ausdrücken zu ergründen:

#### AUFGABE 2.5

Lasse die **heat-water**-Funktion im Stepper laufen und beobachte, wie – vor allem in welcher Reihenfolge – die Bedingungen in einem **cond**-Ausdruck ausgewertet werden. □

Wenn Du diese Aufgabe erledigt hast, wirst Du beobachtet haben, dass DrRacket die Bedingungen in einem **cond**-Ausdruck nacheinander auswertet. Sobald eine davon **#t** ergibt, macht DrRacket mit dem Ausdruck dieses Zweiges weiter. Die restlichen Zweige werden gar nicht mehr berücksichtigt, unabhängig davon, ob sie **#t** ergeben könnten oder nicht. Das können wir ausnutzen, um die Funktion weiter zu vereinfachen. Dazu betrachten wir die beiden ersten Bedingungen im «inneren» **cond**:

```
(cond
  ((< (+ temp heat) 0) (+ temp heat))
  ((and (>= (+ temp heat) 0)
        (< (+ temp heat) 80))
   0)
  ...)
```

Wenn die erste Bedingung in diesem Ausschnitt **#f** ergeben hat, also nicht stimmt, dann ist **(+ temp heat)** größer oder gleich 0. Das heißt, der Teilausdruck **(>= (+ temp heat) 0)** in der *nächsten* Bedingung liefert *immer* **#t**. Wir können also vereinfachen:

```
(cond
```

```
((< (+ temp heat) 0) (+ temp heat))
(and #t
    (< (+ temp heat) 80))
0)
...)
```

Ein Ausdruck `(and #t e)` liefert immer das gleiche Ergebnis wie `e` – schau nochmal auf Seite 44, um Dich davon zu überzeugen! Wir können also das innere `cond` weiter vereinfachen zu:

```
(cond
  ((< (+ temp heat) 0) (+ temp heat))
  ((< (+ temp heat) 80) 0)
  ...)
```

Wir benutzen also Mathematik (genauer gesagt *Algebra*, also die Lehre der Gleichungen), um unser Programm zu vereinfachen. Algebra ist ein sehr mächtiges Werkzeug in der Programmierung, und wir werden es noch oft nutzen. Das fällt oft einfacher, wenn wir gar nicht über die konkrete Bedeutung der Ausdrücke nachdenken sondern nur Gleichungen benutzen, wie oft in der Mathematik und entsprechend Mantra 6 auf Seite 37.

Wir sind aber mit dem Vereinfachen noch nicht fertig. Betrachten wir nun die drei «äußeren» Bedingungen:

```
(cond
  ((< temp 0) ...)
  ((>= (+ temp heat) 100) ...)
  ((and (>= temp 0) (< (+ temp heat) 100)) ...))
```

Wenn die erste Bedingungen `#f` ergibt, gilt automatisch `(>= temp 0)`. Wir können die letzte Bedingung also vereinfachen zu:

```
(and #t (< (+ temp heat) 100))
```

Wenn auch die zweite Bedingung `#f` ergibt, gilt auch `(< (+ temp heat) 100)`. Wir können also vereinfachen zu `(and #t #t)` und von da zu `#t`.

Wir können die Funktion mit dieser Einsicht vereinfachen und `#t` als Bedingung hinschreiben. (Probier es aus!) Allerdings finden manche das hässlich: Darum ist es auch möglich, statt `#t` das besondere Wort `else` («andernfalls» auf Englisch) hinzuschreiben. Der `else`-Zweig kommt also zum Zug, wenn alle anderen «durchgefallen» sind. Das Ergebnis sieht dann so aus:

```
(define heat-water
```

In den Lehrsprachen werden Verzweigungen mit der Spezialform `cond` dargestellt. Ein `cond`-Ausdruck hat die folgende Form:

```
(cond
  (b1 a1)
  (b2 a2)
  ...
  (bn-1 an-1)
  (else an))
```

Dabei sind die  $b_i$  und die  $a_i$  ihrerseits Ausdrücke. Der `cond`-Ausdruck wertet nacheinander alle Bedingungen  $b_i$  aus; sobald eine Bedingung  $b_k$  `#t` ergibt, wird der `cond`-Ausdruck durch das entsprechende  $a_k$  ersetzt. Wenn alle Bedingungen fehlschlagen, wird durch  $a_n$  ersetzt. Die Paarungen  $(b_i \ a_i)$  heißen *Zweige* des `cond`-Ausdruckes, und der Zweig mit `else` heißt *else-Zweig*. Der `else`-Zweig kann auch fehlen – dann sollte aber immer eine der Bedingungen `#t` ergeben. Wenn doch einmal bei allen  $b_i$  `#f` herauskommen sollte, bricht DrRacket das Programm ab und gibt eine Fehlermeldung aus.

**Abbildung 2.2:** Verzweigung

```
(lambda (temp heat)
  (cond
    ; Anfangstemperatur ist unter 0°C, es wird also Eis erwärmt.
    ((< temp 0)
     (cond
       ; Erwärmung bleibt unter 0°C.
       ((< (+ temp heat) 0) (+ temp heat))
       ; Erwärmung bleibt bei 0°C "stecken".
       ((< (+ temp heat) 80) 0)
       ; Erwärmung erhöht Temperatur über den Nullpunkt hinaus.
       (else (- (+ temp heat) 80))))
    ; Erwärmung würde Wassertemperatur auf über 100°C erhöhen.
    ((>= (+ temp heat) 100) 100)
    ; Wasser fängt flüssig an und bleibt durch Erwärmung flüssig.
    (else
     (+ temp heat)))))
```

Damit haben wir die komplette Funktionsweise von `cond` angewendet. Abbildung 2.2 fasst sie noch einmal zusammen.

Du kannst von vornherein `else` benutzen, wenn Du Dir sicher bist, dass alle anderen Fälle schon in den vorigen Zweigen abgedeckt sind. Im Zweifelsfall empfehlen wir aber immer den Weg

über die Mathematik, damit die Funktion auch korrekt wird.

A propos korrekt: Ist `heat-water` korrekt? Hier ist ein weiterer Testfall:

```
(check-expect (heat-water -1 191) 100)
```

Von den 191°C wird 1°C benötigt, um auf 0°C zu kommen, dann weitere 80°C, um über 0°C hinauszukommen, bleiben 110°C. Aber über 100°C geht es natürlich trotzdem nicht. Leider sieht die Funktion das anders:

**Check-Fehler:**

Der tatsächliche Wert 110 ist nicht der erwartete Wert 100.

Woran liegt das? Wenn wir die Verzweigungen im Kopf nachvollziehen, sehen wir, dass zunächst der erste Zweig des äußeren `cond` greift, weil die Bedingung (`< temp 0`) als Wert `#t` hat. Im inneren `cond` schließlich ergeben die ersten beiden Bedingungen jeweils `#f`, es bleibt also der `else`-Zweig, und addiert die Wärme «blind» auf die Anfangstemperatur.

Wir müssen also unser Programm korrigieren, weil dieser Fall noch nicht berücksichtigt ist: Die Anfangstemperatur ist unter 0°C, die Erwärmung würde die Temperatur aber über 100°C heben. Es greift also der `else`-Zweig des inneren `cond` des äußeren Zweigs mit (`< temp 0`), und das ist in diesem Fall falsch. Die ersten beiden Zweige sind für Endtemperaturen bis 0°C zuständig, wir müssen also den neuen Zweig unmittelbar vor dem `else`-Zweig einfügen. Der sieht so aus:

```
; Erwärmung würde die Temperatur auf über 100°C bringen
((>= (- (+ temp heat) 80) 100) 100)
```

Jetzt ist das Programm endlich fertig und korrekt!

Doch wie konnte dieser Fehler unbemerkt bleiben, zumindest zunächst von uns? Das liegt daran, dass die Fallunterscheidung, die der Aufgabe zugrundliegt, ziemlich komplex ist. Bei komplexen Programmen ist das Risiko *immer* groß, dass sie Fehler enthalten. Bei komplexen Fallunterscheidungen bedeutet dies, dass fast immer ein Fall im Programm fehlt, zumindest beim ersten Anlauf. Wir sollten uns bei der Gelegenheit an Mantra 7 auf Seite 38 erinnern:

MANTRA 7

Wenn Deine Funktion Dir kompliziert scheint, ist es wahrscheinlich, dass sie Fehler enthält: Entweder, weil sie eigentlich einfacher sein sollte, oder weil die Aufgabe kompliziert und damit ihre Lösung fehleranfällig ist.

Bei komplizierten Fallunterscheidungen, teste deshalb gründlich und gehe davon aus, dass Du Zweige vergessen hast.



**Einfacher Wasser erhitzen** | Aus Mantra 7 kann man auch noch eine und bessere Schlussfolgerung ziehen, als einfach nur sorgfältig zu prüfen und zu testen. Wenn nämlich die Fallunterscheidungen so kompliziert sind, dann solltest Du gründlich prüfen, ob es nicht auch einfacher geht.

Wir haben dieses Prinzip beim Schreiben dieses Buchs leider vernachlässigt. Glücklicherweise hat Nicolas Neuß von der Universität Erlangen eine Vorabversion dieses Buches gelesen. Ihm stieß die Komplexität der Funktion ebenfalls auf, und er machte folgenden Vorschlag:

Das Ganze wird meiner Meinung nach viel hübscher, wenn man das physikalischer angeht. Wenn man daher Konversionsroutinen  $E \rightarrow T$  und  $T \rightarrow E$  schreibt (E steht für innere Energie, also die Wärme, und T für Temperatur), so sieht man erstens die Problematik und zweitens kann man die Temperaturerhöhung einfach berechnen.

Wir versuchen mal, Dr. Neuß' Idee nachzuvollziehen. Wir fangen mit der von ihm erwähnten Funktion  $E \rightarrow T$  an. In unserer Terminologie könnte die Funktion `heat->temperature` heißen, mit folgender Kurzbeschreibung und Signatur:

```
; Aus Wärme Temperatur berechnen
(: heat->temperature (real -> real))
```

Entsprechend den Ausführungen am Anfang des Abschnitts entsprechen unterhalb von  $0^\circ\text{C}$  die Temperatur und die Wärme einander. Bei Wärme zwischen  $0^\circ\text{C}$  und  $80^\circ\text{C}$  ist die Temperatur konstant  $0^\circ\text{C}$ . Ab  $80^\circ\text{C}$  Wärme führt jedes Grad Erhöhung auch zu einer Erhöhung der Temperatur, bis  $100^\circ\text{C}$  Temperatur beziehungsweise  $180^\circ\text{C}$  Wärme erreicht sind. Ab da bleibt die Temperatur bei  $100^\circ\text{C}$ , auch wenn die Wärme steigt.

Die Testfälle dazu könnten also so aussehen:

```
(check-expect (heat->temperature -50) -50)
(check-expect (heat->temperature 0) 0)
(check-expect (heat->temperature 20) 0)
(check-expect (heat->temperature 80) 0)
(check-expect (heat->temperature 81) 1)
(check-expect (heat->temperature 180) 100)
(check-expect (heat->temperature 200) 100)
```

Die Eingabe-Wärme zerfällt entsprechend der Beschreibung oben in vier Fälle, die wir als vier Zweige eines `cond` schreiben:

```
(define heat->temperature
  (lambda (heat)
    (cond
```

```
(... ...)
(... ...)
(... ...)
(... ...)))))
```

Die Bedingungen und Ausdrücke ergeben sich direkt aus der Beschreibung:

```
(define heat->temperature
  (lambda (heat)
    (cond
      ((<= heat 0) heat)
      ((and (< 0 heat)
            (<= heat 80))
       0)
      ((and (< 80 heat)
            (<= heat 180))
       (- heat 80))
      ((< 180 heat) 100))))
```

Wir brauchen außerdem noch eine Funktion in die andere Richtung, wir nennen sie `temperature->heat`. Diese Funktion ist einfacher: Temperaturen bis 0°C entsprechen der Wärme, Temperaturen darüber brauchen 80°C mehr Wärme. Es gibt also nur zwei Fälle und damit zwei Zweige:

**; Aus Temperatur Wärme berechnen**

```
(: temperature->heat (real -> real))
```

```
(check-expect (temperature->heat -20) -20)
(check-expect (temperature->heat -1) -1)
(check-expect (temperature->heat 1) 81)
(check-expect (temperature->heat 100) 180)
```

```
(define temperature->heat
  (lambda (temp)
    (cond
      ((< temp 0) temp)
      ((and (> temp 0)
            (<= temp 100))
       (+ temp 80)))))
```

Dr. Neuß schlug vor, diese beiden Funktionen dazu zu benutzen, um den Rumpf von `heat-water` zu schreiben. Wir versuchen das mal und schreiben sicherheitshalber eine neue Funktion `heat-water-2`, bevor wir die alte wegwerfen:

```
; Wassertemperatur nach Erhitzen berechnen, mit Eis & Sieden
(: heat-water-2 (real real -> real))
```

```
(define heat-water-2
  (lambda (temp heat)
    ...))
```

(Wir nehmen die gleichen Testfälle wie bei `heat-water`.) Wie können wir `heat->temperature` und `temperature->heat` einbringen? Nun, wir können für `temp` die zugehörige Wärme berechnen, `heat` addieren, und das ganze wieder zurück in eine Temperatur wandeln:

```
(define heat-water-2
  (lambda (temp heat)
    (heat->temperature
      (+ (temperature->heat temp) heat))))
```

Und Tatsache, die Funktion besteht alle Testfälle! Besser noch, die beiden Funktionen `heat->temperature` und `temperature->heat` sind viel einfacher zu verstehen als der Moloch `heat-water`.

Bei komplizierten Verzweigungen – generell bei allen komplizierten Programmen – ist es wert zu untersuchen, ob es auch einfacher geht.

## 2.6 DATENANALYSE UND SCHABLONEN

In Abschnitt 2.4 auf Seite 49 ist zum ersten Mal eine *Datendefinition* aufgetaucht:

```
; Ein Haustier ist eins der folgenden:
; - Katze
; - Hund
; - Schlange
```

Diese Datendefinition ist das Ergebnis eines Nachdenkprozesses, der sogenannten *Datenanalyse*. Dieser Begriff ist schon in Abschnitt 1.11 auf Seite 30 aufgetaucht, als zweiter Schritt der Konstruk-

tionsanleitung: Ab sofort werden wir diesen Schritt immer durchführen, weil er zentral ist für die systematische Konstruktion von Programmen.

Die Vorgehensweise bei einer Datenanalyse sieht im allgemeinen so aus:

#### KONSTRUKTIONSANLEITUNG 6 (DATENANALYSE)

Suche in der Aufgabenstellung nach problemrelevanten Größen; Kandidaten sind immer die Substantive. Schreibe für jede dieser Größen eine Datendefinition, es sei denn, diese ist aus dem Kontext offensichtlich.

Wenn es für die Datendefinition noch keine Signatur gibt, schreibe eine Signatur-Definition dazu. Schreibe außerdem Beispiele auf und schreibe jeweils einen Kommentar, der die Beziehung zwischen Daten und Information beschreibt.

Wie Du die Signatur-Definition schreibst, dafür gibt es in diesem Buch eine Reihe spezialisierter Konstruktionsanleitungen. Die Beziehung zwischen Daten und Information geht auf Abschnitt 1.6 auf Seite 23 zurück.

In diesem Fall ist «Haustier» eine problemrelevante Größe, zu der wir eine Datendefinition geschrieben haben. Bei der Erwärmung von Wasser haben wir allerdings nicht extra hingeschrieben, dass die Wassertemperatur eine reelle Zahl ist, weil dies allgemein bekannt ist.

Für die Formulierung der Datendefinition gibt es eine Reihe von typischen Mustern, die sich immer wiederfinden. Eines ist bei der Definition von «Haustier» ersichtlich, nämlich «eins der folgenden», das darauf hinweist, dass es sich bei «Haustier» um eine Fallunterscheidung handelt.

#### KONSTRUKTIONSANLEITUNG 7 (FALLUNTERSCHIEDUNG: DATENANALYSE)

Versuche, für die Datendefinition eine Formulierung «... ist eins der folgenden» zu finden. Wenn das möglich ist, beschreibt Deine Datendefinition eine *Fallunterscheidung*. Schreibe dann eine Auflistung aller Fälle, jeder Fall auf eine separate Zeile:

```
; Ein X ist eins der folgenden:
; - Fall 1
; - Fall 2
; - ...
; - Fall n
```

Für den Sonderfall der Aufzählung haben wir folgende Konstruktionsanleitung:

## KONSTRUKTIONSANLEITUNG 8 (AUFZÄHLUNG: DATENANALYSE)

Falls Deine Datendefinition eine Fallunterscheidung beschreibt und jeder der Fälle nur aus einem einzelnen Wert besteht, handelt es sich um eine *Aufzählung*.

Schreibe für jede Aufzählung eine Signatur-Definition der Form:

```
(define s (signature (enum ...)))
```

Achte darauf, dass die Anzahl der Fälle der Signatur-Definition der Anzahl der Fälle der Datendefinition entspricht.

Die Datenanalyse ist die wichtigste Methode zur Umsetzung der Mantras 1 und 2:

## MANTRA 1

Achte darauf, dass für alles, was Du weißt, tatsächlich auch etwas im Programm steht.

## MANTRA 2

Lies, was dort steht.

Die Datenanalyse lenkt beim Lesen zunächst unseren Blick auf die Daten, *nicht* darauf, was später mit ihnen gemacht wird. (Das kommt später.) Außerdem gibt sie uns Werkzeug an die Hand, wie wir die Erkenntnisse aus dem Lesen strukturieren und aufschreiben können. Insbesondere führt sie immer auch zu Elementen in unserem Programm. Bei der Funktion `cute?` haben wir zum Beispiel aus der Datendefinition zunächst eine Signatur-Definition abgeleitet:

```
(define pet
  (signature (enum "Katze" "Hund" "Schlange")))
```

Bei Aufzählungen stehen alle Beispiele schon in der Signatur-Definition stehen. Wir müssen deshalb nicht noch separat Beispiele aufschreiben.

Da die Datendefinition für «Haustier» drei Fälle hat, muss auch die Signatur-Definition drei Fälle haben. Aus der Signatur-Definition ergibt sich direkt eine Verzweigung im Rumpf, die sich nur aus der Signatur (`pet -> boolean`) beziehungsweise der Signatur-Definition von `pet` als Aufzählung mit drei Fällen ergibt:

```
(define cute?
  (lambda (p)
    (cond
```

```
(... ...)
(... ...)
(... ...)))
```

Da *cute?* eine Eingabe zur Signatur *pet* erwartet und *pet* eine Aufzählungssignatur mit *drei* Fällen ist, *muss* im Rumpf eine Verzweigung mit ihrerseits *drei* Zweigen auftauchen, ganz egal, was die konkrete Aufgabenstellung ist. Wir betonen das Wort *drei* deshalb so, weil viele Fehler dadurch passieren, dass die Anzahl der Zweige im *cond* einer solchen Funktion nicht der Anzahl der Fälle entspricht.

So ein unfertiges Programm, in dem sich einige Elemente aus der Analyse der Daten ergeben haben ohne besondere Berücksichtigung der konkreten Aufgabenstellung, heißt *Schablone*. Zu bestimmten Arten von Datendefinitionen gehören bestimmte Schablonen.

#### KONSTRUKTIONSANLEITUNG 9 (SCHABLONE)

Wenn Du das Gerüst fertiggestellt hast, benutze die Signatur und die dazugehörigen Datendefinitionen, um Konstruktionsanleitungen mit ein oder mehreren Schablonen auszuwählen und übertrage die Elemente der Schablonen in den Rumpf der Funktion.

Die Schablone ist das nächste Werkzeug für die Umsetzung von Mantra 1:

#### MANTRA 1

Achte darauf, dass für alles, was Du weißt, tatsächlich auch etwas im Programm steht.

Für jede Sorte von Datendefinition schreiben wir neben der Konstruktionsanleitung für die Datenanalyse auch eine für die Schablone. Die Schablone für Fallunterscheidungen sieht folgendermaßen aus:

#### KONSTRUKTIONSANLEITUNG 10 (FALLUNTERSCHIEDUNG: SCHABLONE)

Wenn Du eine Funktion schreibst, die eine Fallunterscheidung als Eingabe verarbeitet, schreibe als Schablone in den Rumpf eine Verzweigung mit sovielen Zweigen, wie es in der Fallunterscheidung Fälle gibt, nach folgendem Muster:

```
(define f
  (lambda (a)
    (cond
      (... ...)
      ...
```

(... ...))

Schreibe danach Bedingungen in die Zweige, welche die einzelnen Fälle voneinander unterscheiden.

Diese Konstruktionsanleitung mag Dir ziemlich bürokratisch vorkommen, und das ist sie auch. Es mag Dir langweilig vorkommen, sie stur zu befolgen, und das ist bestimmt ab und zu der Fall. Dafür nehmen die Konstruktionsanleitungen Dir viel Denkarbeit ab, und Du kannst Dein Gehirn dann für die Lösung der eigentlichen Problemstellung verwenden, wenn erst einmal die Schablone steht. Spätestens also wenn Du vor einer vermeintlich schweren Aufgabe stehst, wende erst einmal die Konstruktionsanleitungen an, bevor Du Dir Sorgen machst, ob am Ende alles aufgeht. Was am Ende übrigbleibt, ist oft erstaunlich einfach.

## 2.7 EXKURS: VERZWEIGUNGEN IN DER MATHEMATIK

In der Mathematik gibt es auch Verzweigungen. Diese werden in der Regel mit einer großen geschweiften Klammer geschrieben. In der Mathematik ist es allerdings üblich, alle Bedingungen «auszuschreiben», also ihre Bedeutung nicht von der Reihenfolge der Auswertung abhängig zu machen, wie wir es bei [heat-water](#) getan haben. Die Funktion schreiben wir den mathematischen Gepflogenheiten entsprechend als einbuchstabige Funktion  $h$  mit Parametern  $t$  für [temp](#) und  $h$  für [heat](#):

$$h(t, h) \stackrel{\text{def}}{=} \begin{cases} \begin{cases} t + h & \text{falls } t + h < 0 \\ 0 & \text{falls } t + h \geq 0 \text{ und } t + h < 80 \\ 100 & t + h - 80 \geq 100 \\ t + h - 80 & \text{falls } t + h \geq 80 \end{cases} & \text{falls } t < 0 \\ \begin{cases} 100 \\ t + h \end{cases} & \begin{cases} \text{falls } t + h \geq 100 \\ \text{falls } t \geq 0 \text{ und } t + h < 100 \end{cases} \end{cases}$$

Etwas scher zu erkennen in dieser Schreibweise: Die Bedingung  $t < 0$  ist die Voraussetzung für die ersten vier Fälle.

## 2.8 BOOLESCHE FALLUNTERSCHIEDUNGEN UND BINÄRE VERZWEIGUNGEN

Bei manchen Fallunterscheidungen definiert sich die letzte Kategorie dadurch, dass ein Wert in keine der anderen Kategorien gehört. Dann ist die Benutzung eines `else`-Zweigs im `cond` sinnvoll.

Nehmen wir einmal an, dass Du ein Haustier (wie in Abschnitt 2.4 auf Seite 49) zum Geburtstag geschenkt bekommen hast. Mit dem Geschenk kam das Angebot, das Haustier gegen ein anderes auszutauschen, falls es Dir nicht gefällt. Nehmen wir weiter an, dass Du nur an niedlichen Haustieren interessiert bist. Wir schreiben also eine Funktion, die ein Haustier gegen ein niedliches austauscht, wenn es nicht niedlich ist – und sonst halt nicht austauscht.

Die Funktion könnte folgende Kurzbeschreibung und Signatur haben:

```
(: exchange-for-cute (pet -> pet))
```

Das alte Haustier ist also die Eingabe für die Funktion, und das Haustier nach dem Umtausch kommt heraus. Als Testfälle bieten sich folgende an:

```
(check-expect (exchange-for-cute "Katze") "Katze")
(check-expect (exchange-for-cute "Hund") "Hund")
(check-expect (exchange-for-cute "Schlange") "Katze")
```

Das Gerüst sieht folgendermaßen aus:

```
(define exchange-for-cute
  (lambda (p)
    ...))
```

Als Nächstes ist die Konstruktion der Schablone dran. Wir könnten uns an der Definition von `pet` orientieren und eine Fallunterscheidung mit drei Fällen als Grundlage nehmen. Allerdings ist recht offensichtlich, dass sich das Ergebnis von `exchange-for-cute` danach richtet, ob `cute?` `#t` oder `#f` liefert. `#t` und `#f` bilden hier eine Aufzählung mit nur zwei Fällen. Wir könnten also eine Schablone mit nur zwei Fällen konstruieren:

```
(define exchange-for-cute
  (lambda (p)
    (cond
      ((cute? p) ...)
      (else ...))))
```

Die Lücken sind einfach zu füllen:



Eine binäre Verzweigung hat die folgende Form:

```
(if b
    k
    a)
```

Dabei ist  $b$  die Bedingung und  $k$  und  $a$  sind die beiden Zweige: die *Konsequente*  $k$  und die *Alternative*  $a$ .

Abhängig vom Ausgang der Bedingung ist der Wert der Verzweigung entweder der Wert der Konsequente oder der Wert der Alternative.

**Abbildung 2.3:** Binäre Verzweigung

```
(define exchange-for-cute
  (lambda (p)
    (cond
      ((cute? p) p)
      (else "Katze"))))
```

Diese boolesche Fallunterscheidungen bildet einen häufig auftretenden Spezialfall: Der zweite Fall ist gerade dadurch definiert, dass er nicht der erste ist, der entsprechende `cond`-Ausdruck hat also immer zwei Zeige: einen mit Bedingung und einen mit `else`. Weil solche binären Verzweigungen so häufig vorkommen, gibt es eine etwas kürzere Spezialform für das «binäre `cond` mit `else`» namens `if`. Damit können wir `exchange-for-cute` etwas kürzer schreiben:

```
(define exchange-for-cute
  (lambda (p)
    (if (cute? p)
        p
        "Katze")))
```

Das `if` spart also ein Paar Klammern und das `else`, aber dafür kann es eben nur binäre Verzweigungen. Abbildung 2.3 zeigt die allgemeine Form von `if`-Ausdrücken.

## AUFGABE 2.6

In der Mathematik gibt den «Absolutbetrag» einer Zahl: Für eine negative Zahl liefert es die entsprechende positive Zahl (aus  $-5$  wird also  $5$ ), alle anderen Zahlen bleiben unverändert. Schreibe eine Funktion, die den Absolutbetrag einer Zahl berechnet! □

Für boolesche Fallunterscheidungen gibt es eine separate Konstruktionsanleitung:

### KONSTRUKTIONSANLEITUNG 11 (BOOLESCHER FALLUNTERSCHIEDUNG: SCHABLONE)

Wenn sich das Ergebnis einer Funktion nach einer booleschen Größe richtet, welche die Funktion mit Hilfe der Eingaben berechnen kann, benutze als Schablone im Rumpf eine binäre Verzweigung:

```
(define f
  (lambda (e)
    (if ... ; hier wird die boolesche Größe berechnet
      ...
      ...)))
```

## 2.9 SYNTAKTISCHER ZUCKER

Tatsächlich ist **if** «primitiver» als **cond**: jede **cond**-Form kann in eine äquivalente **if**-Form übersetzt werden, und zwar nach folgendem Schema:

```
(cond (t1 a1) (t2 a2) ... (tn-1 an-1) (else an))
  ↦ (if t1 a1 (if t2 a2 ... (if ... (if tn-1 an-1 an)...)))
```

Die geschachtelte **if**-Form auf der rechten Seite der Übersetzung wertet, genau wie die **cond**-Form, nacheinander alle Bedingungen aus, bis eine **#t** liefert. Die rechte Seite des **cond**-Zweigs ist dann gerade die Konsequente des **ifs**. Erst wenn alle Bedingungen fehlgeschlagen sind, ist die Alternative des letzten **if**-Ausdrucks dran, nämlich  $a_n$  aus dem **else**-Zweig.

Da sich mit Hilfe dieser Übersetzung jede **cond**-Form durch geschachtelte **if**-Formen ersetzen lässt, ist **cond** streng genommen gar nicht notwendig. **Cond** ist deswegen eine sogenannte *abgeleitete Form*. Da **cond** und andere abgeleitete Formen trotzdem praktisch und angenehm zu verwenden sind und damit dem Programmierer die Arbeit versüßen, heißen abgeleitete Formen auch *syntaktischer Zucker*.

### AUFGABE 2.7

Schreibe die Funktion **cute?** aus Abschnitt 2.4 so um, dass sie **if** statt **cond** verwendet. □

Auch **and** und **or** sind eigentlich syntaktischer Zucker: Es ist immer möglich, einen **and**-Ausdruck in **ifs** zu übersetzen. Es gelten folgende Übersetzungsregeln:

```
(and e)  $\mapsto$  e
(and e1 e2 ...)  $\mapsto$  (if e1 (and e2 ...) #f)
```

Ein **and**-Ausdruck mit mehreren Operanden wird so schrittweise in eine Kaskade von **if**-Ausdrücken übersetzt:

```
(and a b c)
 $\mapsto$  (if a (and b c) #f)
 $\mapsto$  (if a (if b (and c) #f) #f)
 $\mapsto$  (if a (if b c #f) #f)
```

Ebenso lassen sich **or**-Ausdrücke immer in **if**-Ausdrücke übersetzen:

```
(or e)  $\mapsto$  e
(or e1 e2 ...)  $\mapsto$  (if e1 #t (or e2 ...))
```

Beispiel:

```
(or a b c)
 $\mapsto$  (if a #t (or b c))
 $\mapsto$  (if a #t (if b #t (or c)))
 $\mapsto$  (if a #t (if b #t c))
```

## AUFGABE 2.8

Beim syntaktischen Zucker für **and** und **or** könnten wir zusätzlich auch definieren, in was (**and**) und (**or**) (also jeweils ohne Operanden) übersetzt wird. Wir könnten dann die Regeln für (**and e**) und (**or e**) weglassen. Was wäre jeweils die korrekte Übersetzung? Du solltest einige Beispiele durchspielen, um dies herauszufinden. (Achtung: Die korrekte Übersetzung überrascht Dich vielleicht.) Übersetze (**and a b c**) und (**or a b c**) jeweils mit Hilfe Deiner neuen Regeln so, als ob die Regeln (**and e**) und (**or e**) nicht da wären.  $\square$

## 2.10 UNSINNIGE DATEN ABFANGEN

Noch einmal zurück zur Funktion **heat-water** in Abschnitt 2.5: Die hat schon ziemlich viele Zweige. Genau genommen decken die aber immer noch nicht alles ab. Bei Anfangstemperaturen über 100°C verhält sich die Funktion merkwürdig:

```
(heat-water 150 0)
 $\hookrightarrow$  100
```

Das liegt daran, dass es gar kein vernünftiges Ergebnis gibt, wenn die Anfangstemperatur über 100°C liegt. Die Funktion berechnet aber trotzdem munter eines. Es wäre gut, wenn wir melden könnten, dass die Funktion unsinnige Eingaben bekommen hat.

Dazu erweitern wir erst einmal die Verzweigung, gleich am Anfang:

```
(define heat-water
  (lambda (temp heat)
    (cond
      ; Anfangstemperatur ist über 100°C, also unzulässig
      ((> temp 100) ...)
      ...)))
```

Nur – was tun im Fehlerfall? Dazu gibt es eine eingebaute Funktion `violation`, die eine Fehlermeldung als Zeichenkette akzeptiert und, wenn sie aufgerufen wird, das Programm abbricht und die Fehlermeldung ausdrückt. Die vollständige fertige Funktion sieht dann so aus:

```
(define heat-water
  (lambda (temp heat)
    (cond
      ; Anfangstemperatur ist über 100°C, also unzulässig
      ((> temp 100)
       (violation "Anfangstemperatur über 100°C"))
      ; Anfangstemperatur ist unter 0°C, es wird also Eis erwärmt.
      ((< temp 0)
       (cond
         ; Erwärmung bleibt unter 0°C.
         ((< (+ temp heat) 0) (+ temp heat))
         ; Erwärmung bleibt bei 0°C "stecken".
         ((< (+ temp heat) 80) 0)
         ; Die Erwärmung würde die Temperatur auf über 100°C bringen
         ((>= (- (+ temp heat) 80) 100) 100)
         ; Erwärmung erhöht Temperatur über Nullpunkt hinaus.
         (else
          (- (+ temp heat) 80))))
      ; Erwärmung würde Wassertemperatur auf über 100°C erhöhen.
      ((>= (+ temp heat) 100) 100)
      ; Wasser fängt flüssig an und bleibt durch Erwärmung flüssig.
      (else
```

```
(+ temp heat))))))
```

Natürlich sollten wir auch den Fehlerfall testen – das geht nicht mit `check-expect`, das ja erwartet, dass ein Testausdruck einen ordnungsgemäßen Wert liefert. Für Fehlerfälle gibt es `check-error`, das Testfälle erzeugt, die dann bestanden sind, wenn die Auswertung einen Fehler liefert:

```
(check-error (heat-water 150 0)) ; Anfangstemperatur über 100°C
```

Die einfachere Funktion `heat-water-2` besteht diesen Test übrigens auch, ganz ohne Verwendung von `violation`. Warum ist das so? Hier ist die Fehlermeldung:

```
cond: alle Bedingungen ergaben #f
```

Außerdem markiert DrRacket die Funktion `temperature->heat` als Übeltäter. Deren Bedingungen schließen Temperaturen über 100°C explizit aus. Auch hier ist also die einfachere Version klar im Vorteil.

#### AUFGABE 2.9

Probiere die Ausdrücke `(heat-water 0 0)` und `(heat-water-2 0 0)` aus. Sie liefern unterschiedliche Ergebnisse. Warum? Welche Version hat recht? ☐

## AUFGABEN

#### AUFGABE 2.10

Zu den «Flensburg»-Punkten, die es bei Verstößen gegen die Straßenverkehrsordnung gibt, hat eine Seite im Internet folgendes zu sagen:

0 Punkte	Keine Sanktionen
1 bis 3 Punkte	Vormerkung
4 bis 5 Punkte	Ermahnung
6 bis 7 Punkte	Verwarnung
8 Punkte	Entzug

Schreibe eine Funktion, die zu einer gegebenen Punktezahl die daraus folgende Konsequenz berechnet!

#### AUFGABE 2.11

Schreibe eine Funktion `card-type`, die den Umsatz einer Kreditkarte akzeptiert und die eine entsprechende Kategorie als Zeichenkette zurückgibt. Verwende die Konstruktionsanleitungen: Schreibe die Kurzbeschreibung auf, führe eine Datenanalyse durch und schreibe die Signatur auf. Erstelle dann die Testfälle und das Gerüst. Vervollständige danach den Rumpf der Funktion und vergewissere Dich, dass die Tests erfolgreich laufen.

		Umsatz	<	15.000	⇒	Weiß
15.000	≤	Umsatz	<	50.000	⇒	Gold
50.000	≤	Umsatz	≤	150.000	⇒	Platin
150.000	<	Umsatz			⇒	Schwarz

#### AUFGABE 2.12

1. Schreibe eine Funktion `min-2`, die als Argumente zwei Zahlen nimmt und die kleinere der beiden Zahlen zurückgibt. Schreibe außerdem eine Funktion `min-3`, die als Argumente drei Zahlen nimmt und die kleinste der drei Zahlen zurückgibt. Verwende die Konstruktionsanleitung: Schreibe explizit Kurzbeschreibung und Signatur auf, erstelle dann das Gerüst und die Testfälle. Vervollständige danach den Rumpf der Funktion und vergewissere Dich, dass die Tests erfolgreich laufen.
2. Schreibe analog eine Funktion `max-2` und `max-3`.
3. Schreibe eine Funktion `min-3`, welche die kleinste von drei gegebenen Zahlen ausgibt.
4. Schreibe eine Funktion `min-3?`, die überprüft ob die erste von drei gegebenen Zahlen das Minimum ist.
5. Schreibe eine Funktion `valid-value?`, die überprüft ob die erste von drei gegebenen Zahlen zwischen den beiden anderen liegt. Mit «zwischen» ist gemeint, dass die erste Zahl auch gleich der zweiten oder der dritten sein kann. Gehe davon aus, dass der Aufruf immer `(valid-value? value min max)` lautet.
6. `clamp`, die wie folgt definiert ist:

$$\text{clamp}(x, \text{min}, \text{max}) = \begin{cases} x & \text{min} \leq x \leq \text{max} \\ \text{min} & x < \text{min} \\ \text{max} & x > \text{max} \end{cases}$$

#### AUFGABE 2.13



Die Bußgelder sind wie folgt definiert:

- Bei Rot über die Ampel innerhalb der ersten Sekunde €50 und 3 Punkte.
- Bei Rot über die Ampel innerhalb der ersten Sekunde mit Gefährdung oder Sachbeschädigung €125, 4 Punkte und 1 Monat Fahrverbot.
- Bei Rot über die Ampel nach der ersten Sekunde €125, 4 Punkte und 1 Monat Fahrverbot.
- Bei Rot über die Ampel nach der ersten Sekunde mit Gefährdung oder Sachbeschädigung €200, 4 Punkte und 1 Monat Fahrverbot.



### 3 ZUSAMMENGESETZTE DATEN

Viele Informationen, die in Programmen repräsentiert werden, bestehen aus mehreren Bestandteilen:

- Ein Festessen besteht aus Vorspeise, Hauptgang und Nachspeise.
- Eine Uhrzeit besteht aus Stunden und Minuten.
- Eine Tür besteht aus Türblatt und Türgriff.

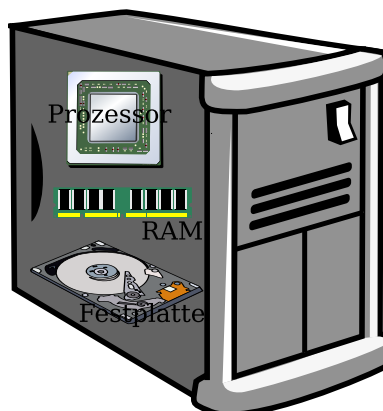
Es werden also mehrere Dinge zu einem *zusammengesetzt*. Eine andere Betrachtungsweise ist, dass ein einzelnes Ding *mehrere Eigenschaften* hat:

- Ein Filzstift hat Dicke und Farbe.
- Eine Katze hat Alter und Geschlecht.
- Ein Lautsprecher hat Minimal- und Maximalfrequenz.

Um solche Informationen abzubilden, führen wir in diesem Kapitel eine neue Art Daten ein, die *zusammengesetzten Daten*.

#### 3.1 COMPUTER KONFIGURIEREN

Viele Computerhändler erlauben ihren Kunden, bestimmte Komponenten eines neuen Computers selbst auszuwählen, zum Beispiel den Prozessor, die Festplatte oder die Größe des RAM-Hauptspeichers:



Anders gesagt, ein Computer *besteht aus*:

- Prozessor
- RAM
- Festplatte

Natürlich besteht ein Computer auch noch aus anderen Teilen, die aber (zumindest in diesem Beispiel) immer gleich oder irrelevant sind. In einer Bestellung müssen also nur diese drei Bestandteile enthalten sein. Wir nehmen an, dass es beim Prozessor nur auf den Namen («Athlon», «Xeon», «Cell», ...) ankommt, beim RAM nur auf die Größe in Gigabyte, und auch bei der Festplatte nur auf die Größe in Gigabyte.

Wir können daraus eine amtliche Datendefinition machen:

```
; Ein Computer besteht aus:
; - Prozessor
; - Hauptspeicher-Kapazität in Gbyte
; - Festplatten-Kapazität in Gbyte
```

Wichtig ist hier die Formulierung «besteht aus», die auf zusammengesetzte Daten hindeutet. Die Daten, die aus so einer Datendefinition für zusammengesetzte Daten entstehen, können als Tabelle dargestellt werden:

	<b>Feld</b>	<b>Komponente</b>
Computer	Prozessor	"Cell"
	RAM	8
	Festplatte	250

Diese Tabelle steht demnach für einen Computer mit Cell-Prozessor, 8 Gigabyte RAM und einer 250-Gigabyte-Festplatte. Sie hat also mehrere Bestandteile und ist damit zusammengesetzt. Die Computerfirma wird viele Computer nach diesem Schema ausliefern, bei denen allesamt jeweils Prozessor, RAM und Festplatte in der Bestellung stehen wird. Solche Informationen, die alle dem gleichen Schema folgen, können also nach «Typ» sortiert werden, wobei der Typ festlegt, um was für eine Art Information es geht und aus welchen Teilen sie zusammengesetzt ist. In der obigen Tabelle ist das *Feld* die Allgemeinbezeichnung für ein Bestandteil, das alle Computer haben. Die *Komponente* ist das konkrete Bestandteil eines einzelnen Computers.

Zusammengesetzte Daten bilden die Lehrsprachen durch sogenannte *Records* ab. Jeder Record gehört zu einem bestimmten *Record-Typ*, der festlegt, was für eine Sorte Information repräsentiert wird und welche Felder die Records des Record-Typs haben.

Der Record-Typ für Computer sieht feste Felder vor («Prozessor», «RAM» und «Festplatte»). Ein einzelner Record dieses Typs besteht aus Komponenten, eine pro Feld. (In diesem Fall

"Cell", 8 und 250.)

Record-Typen müssen in einem Programm explizit mit Hilfe einer speziellen *Record-Definition* definiert werden, die mit `define-record` beginnt. Hier ist die Record-Definition für Computer:

```
(define-record computer
  make-computer
  (computer-processor string)
  (computer-ram        natural)
  (computer-hard-drive natural))
```

Diese etwas komplizierte Form erläutern wir Schritt für Schritt. Weil sie gleich mehrere Funktionen definiert, die mit Records zu tun haben, heißt sie `define-record`. Als erstes nach dem `define-record` steht der Name der Signatur für die Werte des Record-Typs, `computer`. Diese Signatur nennen wir ab sofort die *Record-Signatur*.

Als Nächstes in der Record-Definition kommt `make-computer`, der Name einer Funktion, die für die *Konstruktion* eines Computer-Records zuständig ist. Analog zum Zusammenbauen eines Computers mit Cell-Prozessor, 8 Gigabyte RAM und 250 Gigabyte Festplatte wird der dazugehörige Record mit der Funktion `make-computer` folgendermaßen hergestellt:

```
(make-computer "Cell" 8 250)
↪ #<record:computer "Cell" 8 250>
```

`Make-computer` hat folgende Signatur:

```
(: make-computer (string natural natural -> computer))
```

Die drei Eingaben sind der Reihe nach Prozessor, RAM und Festplatte. `Make-computer` macht daraus einen Wert der Sorte `computer` der Computer-Records. Da `make-computer` einen Computer-Record «konstruiert», heißt die Funktion auch *Konstruktor*.

Mit der Schreibweise

```
#<record:... ...>
```

werden Record-Typ und ihre Komponenten in der REPL sichtbar.

Hier sind einige Beispiele für Computer-Records, mit Kommentaren, welche die Beziehung zwischen Daten und Information (siehe Abschnitt 1.6 auf Seite 23) herstellt:

```
; Cell, 4 Gbyte RAM, 1000 Gbyte Festplatte
(define gamer (make-computer "Cell" 4 1000))
```

```
gamer
```

```

↪ #<record:computer "Cell" 4 1000>

; Xeon, 2 Gbyte RAM, 500 Gbyte Festplatte
(define workstation (make-computer "Xeon" 2 500))

workstation
↪ #<record:computer "Xeon" 2 500>

```

Umgekehrt zur Konstruktion nehmen manche Bastler aus dem Computer die Einzelteile wieder heraus, zum Beispiel, um sie in einem anderen Computer zu verbauen. Für dieses Herausnehmen sind die Funktionen `computer-processor`, `computer-ram` und `computer-hard-drive` zuständig, die von der Record-Definition definiert werden:

```

(computer-processor gamer)
↪ "Cell"
(computer-ram gamer)
↪ 4
(computer-hard-drive gamer)
↪ 1000

```

Diese drei Funktionen heißen *Selektoren*. Sie haben folgende Signaturen:

```

(: computer-processor (computer -> string))
(: computer-ram (computer -> natural))
(: computer-hard-drive (computer -> natural))

```

Genau genommen sind diese Signaturen redundant: In der Record-Definition steht ja schon, dass die Felder die Signaturen `string`, `natural` und `natural` sind. Jeder Selektor hat `computer` als Eingabe und die jeweilige Feld-Signatur als Ausgabe. Auch die Signatur-Deklaration für den Konstruktor ist redundant.

### AUFGABE 3.1

Wie ist der Zusammenhang zwischen der Record-Definition und der Signatur des Konstruktors?

□

Trotzdem ist es zumindest am Anfang hilfreich, sich die Arbeitsweise von Konstruktor und Selektoren anhand ihrer Signaturen klarzumachen. Wir schreiben sie darum in diesem Kapitel noch hin, danach nicht mehr.

Mit Hilfe des Konstruktors und der Selektoren können wir weitergehende Funktionen definieren. Für den Anfang könnte das eine Funktion sein, die den Gesamtspeicher eines Computers berechnet, also Hauptspeicher und Festplattenspeicher zusammen. Eine solche Funktion müsste Kurzbeschreibung und Signatur wie folgt haben:

```
; Gesamtspeicher berechnen
(: total-memory (computer -> natural))
```

Hier sind unsere Erwartungen an `total-memory`, als Testfälle formuliert:

```
(check-expect (total-memory workstation) 502)
(check-expect (total-memory gamer) 1004)
```

Das Gerüst ist wie folgt:

```
(define total-memory
  (lambda (c)
    ...))
```

Um etwas aus dem Record zu berechnen, muss `total-memory` (und so gut wie jede andere Funktion auch) die Bestandteile betrachten. Es ist deshalb sinnvoll, die Schablone mit den Aufrufen der Selektoren zu bestücken.

```
(define total-memory
  (lambda (c)
    ... (computer-processor c) ...
    ... (computer-ram c) ...
    ... (computer-hard-drive c) ...))
```

Jetzt wo die Schablone fertig ist, können wir uns mit dem Inhalt der Aufgabe beschäftigen: Der Prozessor hat nichts mit der Speichermenge zu tun, wir können den entsprechenden Selektoraufruf also wieder löschen:

```
(define total-memory
  (lambda (c)
    ... (computer-ram c) ...
    ... (computer-hard-drive c) ...))
```

Der Gesamtspeicher ist die Summe der beiden Komponenten:

```
(define total-memory
  (lambda (c)
```

```
(+ (computer-ram c)
   (computer-hard-drive c)))
```

Fertig!

`Total-memory` ist ein Beispiel für eine Funktion, die einen Record akzeptiert. Umgekehrt gibt es auch Funktionen, die Records produzieren. Angenommen, unser Computerhändler bietet neben der Einzelkonfiguration von Prozessor, Hauptspeicher und Festplatte einige Standardmodelle an – sagen wir, ein Billigmodell, ein Modell für Profis (was immer eine «Profi» sein mag) und ein Modell für Computerspieler. Je nachdem, welches der Modelle der Kunde auswählt, muss die entsprechende Konfiguration zusammengesetzt werden. Für die Standardkonfiguration gibt es drei feste Möglichkeiten, es handelt sich hier also um eine Aufzählung.

Für die Aufzählung machen wir erst einmal – nach Konstruktionsanleitung 8 auf Seite 67 – eine Daten- und eine Signatur-Definition:

```
; Ein Modell ist eins der folgenden:
; - Billigmodell
; - Profi-Modell
; - Gamer-Modell
(define model
  (signature
    (enum "cheap" "professional" "gamer")))
```

Eine Funktion, die zu einer Standardkonfiguration den passenden Computer fertigt, könnte damit folgende Kurzbeschreibung und Signatur haben:

```
; Standard-Computer zusammenstellen
(: standard-computer (model -> computer))
```

Die Testfälle sollten alle drei Standardkonfigurationen abdecken:

```
(check-expect (standard-computer "cheap")
               (make-computer "Sempron" 2 500))
(check-expect (standard-computer "professional")
               (make-computer "Xeon" 4 1000))
(check-expect (standard-computer "gamer")
               (make-computer "Quad" 4 750))
```

Hier ist das Gerüst:

```
(define standard-computer
```

```
(lambda (k)
  ...))
```

Bei der Schablone gehen wir wieder nach Konstruktionsanleitung 8 auf Seite 67 vor. Da es sich beim Argument um eine Fallunterscheidung – eine Aufzählung mit *drei* Alternativen – handelt, können wir die dazu passende Schablone – eine Verzweigung mit *drei* Zweigen – zum Einsatz bringen:

```
(define standard-computer
  (lambda (k)
    (cond
      (... ...)
      (... ...)
      (... ...))))
```

Bei den Tests der Zweige müssen wir `k` mit den Elementen der Aufzählung vergleichen. Da es sich um Zeichenketten handelt, nehmen wir dazu `string=?`:

```
(define standard-computer
  (lambda (k)
    (cond
      ((string=? k "cheap") ...)
      ((string=? k "professional") ...)
      ((string=? k "gamer") ...))))
```

In jedem Zweig müssen wir nun dafür sorgen, dass der entsprechende Computer hergestellt wird. Für das Herstellen von Computer-Records ist der Konstruktor `make-computer` zuständig. Dementsprechend müssen wir in jedem Zweig einen Aufruf von `make-computer` platzieren, jeweils mit drei Argumenten:

```
(define standard-computer
  (lambda (k)
    (cond
      ((string=? k "cheap")
       (make-computer ... ..))
      ((string=? k "professional")
       (make-computer ... ..))
      ((string=? k "gamer")
       (make-computer ... ..)))))
```

Jetzt müssen wir die Argumente für die Aufrufe von `make-computer` zur Verfügung stellen. Für jeden Aufruf sind das der Prozessor, die Größe des Hauptspeichers und die Größe der Festplatte. Die entsprechenden Angaben können wir zum Beispiel den Testfällen entnehmen. Folgendes kommt dabei heraus:

```
(define standard-computer
  (lambda (k)
    (cond
      ((string=? k "cheap")
       (make-computer "Sempron" 2 500))
      ((string=? k "professional")
       (make-computer "Xeon" 4 1000))
      ((string=? k "gamer")
       (make-computer "Quad" 4 750)))))
```

Fertig!

### AUFGABE 3.2

Schreibe eine Funktion, die einen Computer klassifiziert als «fett», «Durchschnitt» oder «müde». Definiere die Kriterien dafür selbst. □

## 3.2 ZUSAMMENGESETZTE DATEN SELBST KONSTRUIEREN

Für ein weiteres Beispiel greifen wir auf folgenden Satz aus der Einleitung zurück, den wir schon als Datendefinition auslegen können:

```
; Eine Uhrzeit besteht aus Stunde und Minute.
```

Für die Entwicklung der dazu passenden Record-Definition müssen wir uns einen Namen für den Record-Typ ausdenken. Dann können wir bereits ein karges Gerüst hinschreiben:

```
(define-record wallclock-time
  make-wallclock-time
  ...)
```

Als Nächstes müssen wir festlegen, *wie viele* Bestandteile die Records haben sollen. In diesem Fall («Stunde und Minute») sind es zwei. Wir können das Gerüst entsprechend erweitern:



```
(define-record wallclock-time
  make-wallclock-time
  (... ...)
  (... ...))
```

Auf die Anzahl der Bestandteile zu achten, hilft uns dabei, Mantra 1 auf Seite 36 umzusetzen:

#### MANTRA 1

Achte darauf, dass für alles, was Du weißt, tatsächlich auch etwas im Programm steht.

Als Nächstes kommen die Namen der Selektoren. Dabei befolgen wir eine Konvention, die Selektoren alle mit `wallclock-time-` anfangen zu lassen. Bei der Benennung des Konstruktor haben wir ebenfalls eine Konvention angewendet, dessen Name sich aus `make-` und dem Namen des Record-Typs ergibt. Also:

```
(define-record wallclock-time
  make-wallclock-time
  (wallclock-time-hour ...)
  (wallclock-time-minute ...))
```

Dass wir die Selektoren untereinander schreiben, dient lediglich der Übersichtlichkeit, ist also ebenfalls eine Konvention. Es fehlen noch die Signaturen bei `wallclock-time-hour` und `wallclock-time-minute`. Nicht nur die: Wir brauchen überhaupt erstmal Datendefinitionen für die beiden Felder.

```
; Eine Stunde ist eine ganze Zahl zwischen 0 und 23.
; Eine Minute ist eine ganze Zahl zwischen 0 und 59.
```

Da es um ganze Zahlen ab 0 geht, könnten wir `natural` verwenden, präziser ist es aber, wenn wir `integer-from-to` aus Abschnitt 2.2 auf Seite 47 benutzen und eigene Signatur-Definitionen einführen:

```
; Eine Stunde ist eine ganze Zahl zwischen 0 und 23.
(define hour (signature (integer-from-to 0 23)))
; Eine Minute ist eine ganze Zahl zwischen 0 und 59.
(define minute (signature (integer-from-to 0 59)))
```

Diese Signaturene können wir jetzt in der Record-Definition für `wallclock-time` verwenden:

Eine Record-Definition hat folgende allgemeine Gestalt:

```
(define-record t
  c
  (sel1 sig1)
  ...
  (seln sign))
```

Diese Form definiert einen Record-Typ mit  $n$  Feldern. Dabei sind  $t$ ,  $c$ ,  $sel_1 \dots sel_n$  allesamt Variablen, für die `define-record` Definitionen anlegt:

- $t$  ist der Name der Record-Signatur.
- $c$  ist der Name des Konstruktors, den `define-record` anlegt. Der Konstruktor hat folgende Signatur:

```
(: c (sig1 ... sign -> t))
```

- $sel_1, \dots, sel_n$  sind die Namen der Selektoren für die Felder des Record-Typs. Der Selektor  $s_i$  hat folgende Signatur:

```
(: seli (t -> sigi))
```

**Abbildung 3.1:** `define-record` (einfach)

```
(define-record wallclock-time
  make-wallclock-time
  (wallclock-time-hour hour)
  (wallclock-time-minute minute))
```

Wir schreiben die Signaturen der definierten Funktionen aus. Diese ergeben sich direkt aus der Record-Definition:

```
(: make-wallclock-time (hour minute -> wallclock-time))
(: wallclock-time-hour (wallclock-time -> hour))
(: wallclock-time-minute (wallclock-time -> minute))
```

Der Konstruktor akzeptiert für jedes Feld ein Argument – entsprechend stehen die Signaturen der Felder vor dem Pfeil. Heraus kommt beim Konstruktor immer ein Record, da steht also der Name des Record-Typs.

Bei den Selektoren ist es umgekehrt: Da steht immer die Record-Signatur vorn (sie akzeptieren ja jeweils einen Record) und nach dem Pfeil steht die Signatur des jeweiligen Feldes.

Abbildung 3.1 fasst die Form von Record-Definitionen zusammen.

Hier sind drei Beispiele für Uhrzeiten als Daten, mit Kommentaren, welche die repräsentierte Information beschreiben:

```
(define wt1 (make-wallclock-time 11 55)) ; fünf vor zwölf
(define wt2 (make-wallclock-time 0 0)) ; Mitternacht
(define wt3 (make-wallclock-time 1 1)) ; 1 Uhr 1
```

Zuerst berechnen wir für eine Uhrzeit die Anzahl der Minuten seit Mitternacht. Hier sind Kurzbeschreibung, Signatur, Testfälle und Gerüst:

```
; Minuten seit Mitternacht berechnen
(: minutes-since-midnight (wallclock-time -> natural))

(check-expect (minutes-since-midnight wt1) (+ (* 11 60) 55))
(check-expect (minutes-since-midnight wt2) 0)
(check-expect (minutes-since-midnight wt3) 61)
```

```
(define minutes-since-midnight
  (lambda (wt)
    ...))
```

`Minutes-since-midnight` soll eine Funktion sein, die Uhrzeiten als Eingabe akzeptiert, also zusammengesetzte Daten. Eine Funktion, die aus zusammengesetzten Daten etwas berechnet, muss meist deren Bestandteile verwenden, auf die sie mit den Selektoren zugreifen kann. Wir fügen als nächsten Schritt Aufrufe beider Selektoren ein:

```
(define minutes-since-midnight
  (lambda (wt)
    ... (wallclock-time-hour wt) ...
    ... (wallclock-time-minute wt) ...))
```

Jetzt setzen wir noch etwas Wissen über Uhrzeiten ein und vervollständigen damit den Rumpf:

```
(define minutes-since-midnight
  (lambda (wt)
    (+ (* 60 (wallclock-time-hour wt))
      (wallclock-time-minute wt))))
```

### AUFGABE 3.3

Schreibe eine Funktion, die für eine Uhrzeit zurückliefert, ob sich diese auf den Vormittag oder den Nachmittag (also vor oder nach 12 Uhr mittags) bezieht. □

Die Funktion `minutes-since-midnight` können wir auch umdrehen:

```
; Aus Minuten seit Mitternacht die Uhrzeit berechnen
(: minutes-since-midnight->wallclock-time (natural -> wallclock-time))
```

Der Pfeil in `minutes-since-midnight->wallclock-time` gehört zum Namen dazu und steht für die Umwandlung einer Größe in eine andere. Die Testfälle sind gegenüber `minutes-since-midnight` umgedreht.

```
(check-expect (minutes-since-midnight->wallclock-time (+ (* 11 60) 55
  ))
  wt1)
(check-expect (minutes-since-midnight->wallclock-time 0)
  wt2)
(check-expect (minutes-since-midnight->wallclock-time 61)
  wt3)
```

Hier ist das Gerüst:

```
(define minutes-since-midnight->wallclock-time
  (lambda (msm)
    ...))
```

Dies ist eine Funktion, die eine Uhrzeit produziert – sie muss also den Konstruktor für `wallclock-time` aufrufen. Daraus ergibt sich folgende Schablone:

```
(define minutes-since-midnight->wallclock-time
  (lambda (msm)
    (make-wallclock-time ... ...)))
```

Um die Schablone zum Rumpf zu vervollständigen, müssen wir aus den Minuten seit Mitternacht `msm` zunächst die Stunde berechnen. Dazu brauchen wir eine Funktion, die ganzzahlig teilt. Die eingebaute Funktion `/` macht das leider nicht:

```
(/ 61 60)
↪ 1.016̄
```

Aber die Funktion `quotient` hilft uns weiter:

```
(quotient 61 60)
↪ 1
```

Das können wir in der Schablone benutzen:

```
(define minutes-since-midnight->wallclock-time
  (lambda (msm)
    (make-wallclock-time (quotient msm 60) ...)))
```

Es fehlt noch die Minute – dafür brauchen wir den Divisions*rest*. Den berechnet die eingebaute Funktion `remainder`:

```
(remainder 67 60)
↪ 7
(remainder 125 60)
↪ 5
```

Damit können wir den Rumpf vervollständigen:

```
(define minutes-since-midnight->wallclock-time
  (lambda (msm)
    (make-wallclock-time (quotient msm 60) (remainder msm 60))))
```

#### AUFGABE 3.4

Schreibe eine Funktion `make-wallclock-time-12h`, die eine Uhrzeit aus einer Zeitangabe im 12-Stunden-Format konstruiert. Also zum Beispiel:

```
(make-wallclock-time-12h 6 30 "AM")
↪ #<record:wallclock-time 6 30>
(make-wallclock-time-12h 6 30 "PM")
↪ #<record:wallclock-time 18 30>
(make-wallclock-time-12h 12 0 "PM")
↪ #<record:wallclock-time 12 0>
(make-wallclock-time-12h 12 00 "AM")
↪ #<record:wallclock-time 0 0>
(make-wallclock-time-12h 12 30 "PM")
↪ #<record:wallclock-time 12 30>
```

(Für die beiden Fälle 12:00AM und 12:00PM gibt es keine eindeutige Zuordnung, wir haben das willkürlich festgelegt.) □

### 3.3 KONSTRUKTIONSANLEITUNGEN FÜR ZUSAMMENGESetzte DATEN

Dieser Abschnitt fasst die Erkenntnisse aus den Beispielen zu zusammengesetzten Daten in Form von Konstruktionsanleitungen zusammen. Wir fangen an mit der Datenanalyse und der zugehörigen Record-Definition.

#### KONSTRUKTIONSANLEITUNG 12 (ZUSAMMENGESetzte DATEN: DATENANALYSE)

Zusammengesetzte Daten kannst Du an Formulierungen wie «ein  $X$  besteht aus ...», «ein  $X$  ist charakterisiert durch ...» oder «ein  $X$  hat ...» erkennen. Manchmal lautet die Formulierung etwas anders. Die daraus resultierende Datendefinition ist ein Kommentar im Programm in folgender Form:

```
; Ein X hat / besteht aus / ist charakterisiert durch:
; - Bestandteil / Eigenschaft 1
; - Bestandteil / Eigenschaft 2
; ...
; - Bestandteil / Eigenschaft n
```

Auf die Datendefinition folgt eine entsprechende Record-Definition. Dafür überlege Dir Namen für den Record-Typ  $T$  und für die Felder,  $f_1 \dots f_n$ . Für jedes Feld solltest Du außerdem die dazu passende Signatur  $sig_i$  angeben. Die Record-Definition hat dann folgende Form:

```
(define-record T
  make-T
  (T-f1 sig1)
  ...
  (T-fn sign))
```

Der Name des Record-Typs  $T$  ist die Record-Signatur, `make-T` ist der Konstruktor und  $T$ - $f_i$  sind die Selektoren.

Dass der Konstruktorname mit `make-` anfängt und dass die Selektornamen sich aus dem Namen des Typs und der Felder zusammensetzt, ist reine Konvention. Von ihr solltest Du aber nur aus guten Gründen abweichen.

Unter die Record-Definition gehören die Signaturen für den Konstruktor und die Selektoren:

```
(: make-T (sig1 ... sign -> T))
```

```
(: T-f1 (T -> sig1))
...
(: T-fn (T -> sign))
```

Wenn Du genügend Übung in der Verwendung von Konstruktoren und Selektoren hast, kannst Du die Signaturen (die ja redundant sind) auch weglassen: Die relevanten Signaturen für die Felder stehen ja schon in der Record-Definition.

Wenn Du die Datenanalyse und die Record-Definition für zusammengesetzte Daten abgeschlossen hast, solltest Du anhand der Signatur der Funktion feststellen, ob die zusammengesetzten Daten als Ein- oder als Ausgabe verwendet werden. Abhängig davon kannst Du die entsprechende Schablone aus den folgenden beiden Konstruktionsanleitung auswählen.

#### KONSTRUKTIONSANLEITUNG 13 (ZUSAMMENGESETZTE DATEN ALS EINGABE: SCHABLONE)

Wenn Deine Funktion zusammengesetzte Daten als Eingabe akzeptiert (das ergibt sich aus der Signatur), gehe nach Schreiben des Gerüsts folgendermaßen vor:

1. Für jede Komponente, schreibe (*sel r*) in die Schablone, wobei *sel* der Selektor der Komponente und *r* der Name des Record-Parameters ist, also zum Beispiel:

```
(wallclock-time-hour wt)
```

2. Vervollständige die Schablone, indem Du einen Ausdruck konstruierst, in dem die Selektor-Anwendungen vorkommen.
3. Es ist möglich, dass nicht alle Selektor-Anwendungen im Rumpf verwendet werden: In diesem Fall lösche die Selektor-Anwendung wieder.

Mit etwas Übung kannst Du nicht benötigte Selektor-Anwendung auch von vornherein weglassen. Gelegentlich deutet es aber auf einen Fehler hin, wenn eine fehlt: Darum ist es oft sinnvoll, sie zunächst hinzuschreiben, ganz im Sinne von Mantra 1 auf Seite 36:

#### MANTRA 1

Achte darauf, dass für alles, was Du weißt, tatsächlich auch etwas im Programm steht.

Funktionen, die zusammengesetzte Daten als Ausgabe haben, müssen einen entsprechenden Record konstruieren und deshalb den Konstruktor aufrufen. Die Schablone ergibt sich folgendermaßen:

#### KONSTRUKTIONSANLEITUNG 14 (ZUSAMMENGESetzte DATEN ALS AUSGABE: SCHABLO-NE)

Wenn Deine Funktion zusammengesetzte Daten als Ausgabe hat, schreibe einen Aufruf des passenden Record-Konstruktors in den Rumpf, zunächst mit einer Ellipse für jedes Feld des Records, also zum Beispiel:

```
(make-wallclock-time ... ...)
```

### 3.4 EIN- UND AUSGABE ZUSAMMENGESETZTER DATEN

In diesem Abschnitt kombinieren wir Ein- und Ausgabe zusammengesetzter Daten in einer einzigen Funktion.

Im Beispiel dafür geht es um Gürteltiere in Texas: Die überqueren insbesondere die Highways und werden dabei leider oft überfahren – am Straßenrand sind entsprechend viele Gürteltiere zu sehen. Außerdem füttern freundliche Autofahrer gelegentlich die Gürteltiere. Mit diesen beiden Aspekten wollen wir uns beschäftigen: Was passiert, wenn ein Gürteltier überfahren wird? Was passiert, wenn ein Gürteltier gefüttert wird? Entsprechend interessiert uns, ob ein Gürteltier am Leben ist und welches Gewicht es hat. Das können wir nach Konstruktionsanleitung 12 auf Seite 92 direkt in eine Datendefinition übersetzen:

```
; Ein Gürteltier hat folgende Eigenschaften:
; - Gewicht (in g)
; - lebendig oder tot
```

Wiederum handelt es sich um zusammengesetzte Daten, wie aus der Formulierung «hat» ersichtlich ist. Wir beschränken uns hier auf die beiden Eigenschaften, die für die Aufgabenstellung relevant sind. Aus der Datendefinition können wir direkt eine passende Record-Definition machen:

```
(define-record dillo
  make-dillo
  (dillo-weight natural)
  (dillo-alive? boolean))
```

(«Dillo» steht kurz für «Armadillo», englisch für Gürteltier.)

Für das Feld `alive?` könnten wir unterschiedliche Repräsentationen wählen: Eine Aufzählung wäre möglich; wir haben uns für einen booleschen Wert entschieden, der die Frage «Lebt das



Gürteltier?» beantwortet. Hier sind die Signaturen für die Record-Funktionen:

```
(: make-dillo (natural boolean -> dillo))
(: dillo-weight (dillo -> natural))
(: dillo-alive? (dillo -> boolean))
```

Hier sind einige Exemplare als Daten plus Information:

```
(define dillo1 (make-dillo 55000 #t)) ; 55 kg, lebendig
(define dillo2 (make-dillo 58000 #f)) ; 58 kg, tot
(define dillo3 (make-dillo 60000 #t)) ; 60 kg, lebendig
(define dillo4 (make-dillo 63000 #f)) ; 63 kg, tot
```

Fangen wir damit an, Gürteltiere zu füttern. Die Standard-Futter-Portion ist dabei 500g, und das Gürteltier nimmt durch die Fütterung um das entsprechende Gewicht zu. Hier Kurzbeschreibung und Signatur:

```
; Gürteltier mit 500g Futter füttern
(: feed-dillo (dillo -> dillo))
```

Hier der erste, naheliegende Testfall:

```
(check-expect (feed-dillo dillo1) (make-dillo 55500 #t))
```

Bei `feed-dillo` ist relevant, was es mit toten Gürteltieren macht: Tote Gürteltiere fressen nicht, entsprechend nehmen sie auch nicht zu, wenn man ihnen Futter anbietet:

```
(check-expect (feed-dillo dillo2) dillo2)
```

Hier das Gerüst der Funktion:

```
(define feed-dillo
  (lambda (dillo)
    ...))
```

Für den Namen des Parameters verwenden wir auch `dillo`, nicht zu verwechseln mit der Signatur, die ebenfalls `dillo` heißt. Das `lambda` sorgt dafür, dass `dillo` innerhalb seines Rumpfes sich auf den Parameter bezieht, nicht auf die weiter außen stehende Signatur.

### AUFGABE 3.5

Um Dir klarzumachen, welches `dillo` zu welchem `lambda` beziehungsweise zu welcher Definition gehört, kannst Du in DrRacket den Knopf [Syntaxprüfung](#) drücken und danach den Mauszeiger über die verschiedenen Vorkommen von `dillo` bewegen. □

`Feed-dillo` hat zusammengesetzte Daten sowohl als Eingabe als auch als Ausgabe. Entsprechend kommen die Schablonen für beide Situationen zum Einsatz. Zunächst die Schablone für zusammengesetzte Daten als Eingabe aus Konstruktionsanleitung 13 auf Seite 93; wir schreiben die Aufrufe der Selektoren auf:

```
(define feed-dillo
  (lambda (dillo)
    ... (dillo-weight dillo) ...
    ... (dillo-alive? dillo) ...))
```

Dazu kommt die Schablone für zusammengesetzte Daten als Ausgabe aus Konstruktionsanleitung 14 auf Seite 94, also der Aufruf des Konstruktors:

```
(define feed-dillo
  (lambda (dillo)
    (make-dillo ... ...)
    ... (dillo-weight dillo) ...
    ... (dillo-alive? dillo) ...))
```

Der zweite Testfall zeigt, dass, was `feed-dillo` betrifft, die Gürteltiere in zwei verschiedene Gruppen fallen: `Feed-dillo` verhält sich bei lebenden Gürteltieren anders als bei toten: eine Fallunterscheidung.

Entsprechend brauchen wir eine Verzweigung im Rumpf, und aufgrund des Wertes von `(dillo-alive? dillo)`, der glücklicherweise schon in der Schablone steht. Da `dillo-alive?` einen booleschen Wert liefert, handelt es sich um eine boolesche Fallunterscheidung. Deshalb können wir Konstruktionsanleitung 11 auf Seite 72 anwenden und eine binäre Verzweigung benutzen:

```
(define feed-dillo
  (lambda (dillo)
    (if (dillo-alive? dillo)
        ...
        ...)
    (make-dillo ... ...)
    ... (dillo-weight dillo) ...))
```

Nun müssen wir noch die beiden Zweige ergänzen. Am einfachsten ist «Gürteltier tot», dann nämlich kommt das gleiche Gürteltier aus der Funktion, das hineingegangen ist. Wir setzen also `dillo` als Alternative der Verzweigung ein:

```
(define feed-dillo
```

```
(lambda (dillo)
  (if (dillo-alive? dillo)
      ...
      dillo)
  (make-dillo ... ...)
  ... (dillo-weight dillo) ...))
```

Im ersten Zweig müssen wir schließlich einen neuen Gürteltier-Wert berechnen, der die Zunahme berücksichtigt. Dabei werden der Konstruktor-Aufruf und der zweite Selektor-Aufruf aus der Schablone verbraucht:

```
(define feed-dillo
  (lambda (dillo)
    (if (dillo-alive? dillo)
        (make-dillo ... ...)
        dillo)))
```

Wir müssen beim Aufruf des Konstruktors `make-dillo` angeben, welches Gewicht das frisch gefütterte Gürteltier haben soll und ob es noch am Leben ist. Das Gewicht erhöht sich um das Gewicht des Futters. Außerdem ist das Gürteltier noch am Leben, weil der Konstruktoraufruf in dem Zweig steht, in dem das so ist:

```
(define feed-dillo
  (lambda (dillo)
    (if (dillo-alive? dillo)
        (make-dillo (+ (dillo-weight dillo) 500)
                    #t)
        dillo)))
```

Wir kommen nun zum unangenehmen Teil, dem Überfahren, das aus einem lebenden Gürteltier ein totes macht. Hier Kurzbeschreibung und Signatur:

```
; Gürteltier überfahren
(: run-over-dillo (dillo -> dillo))
```

Aus dem Beispiel `dillo1` können wir den ersten Testfall machen:

```
(check-expect (run-over-dillo dillo1) (make-dillo 55000 #f))
```

Wir sollten aber auch berücksichtigen, was `run-over-dillo` mit toten Gürteltieren anstellt. Diese bleiben auch nach dem Überfahren tot:

```
(check-expect (run-over-dillo dillo2) dillo2)
```

Hier das Gerüst der Funktion:

```
(define run-over-dillo
  (lambda (dillo)
    ...))
```

`Run-over-dillo` hat zusammengesetzte Daten sowohl als Eingabe als auch als Ausgabe. Entsprechend kommen ein weiteres Mal die Schablonen für beide Situationen zum Einsatz. Zunächst die Schablone für zusammengesetzte Daten als Eingabe; wir schreiben die Aufrufe der Selektoren auf:

```
(define run-over-dillo
  (lambda (dillo)
    ... (dillo-weight dillo) ...
    ... (dillo-alive? dillo) ...))
```

Dazu kommt die Schablone für zusammengesetzte Daten als Ausgabe, also der Aufruf des Konstruktors:

```
(define run-over-dillo
  (lambda (dillo)
    (make-dillo ... ...)
    ... (dillo-weight dillo) ...
    ... (dillo-alive? dillo) ...))
```

Da das Überfahren das Gewicht nicht ändert, übernimmt der Ausdruck für das Gewicht das Gewicht des Eingabe-Gürteltiers aus der Schablone:

```
(define run-over-dillo
  (lambda (dillo)
    (make-dillo (dillo-weight dillo) ...)
    ... (dillo-alive? dillo) ...))
```

Das Gürteltier ist nach dem Überfahren auf jeden Fall tot. Da es keine Rolle spielt, ob das Gürteltier vorher lebendig war oder nicht, können wir den Selektoraufruf `(dillo-alive? dillo)` verwerfen:

```
(define run-over-dillo
  (lambda (dillo)
    (make-dillo (dillo-weight dillo)
                #f)))
```

Fertig!

### 3.5 ALTERNATIVEN BEI DEN KONSTRUKTIONSANLEITUNGEN

Vielleicht ist Dir bei der Folge von Schablonen für `feed-dillo` aufgefallen, dass wir die Anordnung der Schablonen für die Konstruktion zusammengesetzter Daten und die Schablone für binäre Fallunterscheidungen recht willkürlich angeordnet haben, indem wir die Verzweigung «außen» um den Konstruktoraufruf gestellt haben.

Die Konstruktionsanleitungen hätten wir genauso gut andersherum anwenden können, indem wir die Verzweigung «innen» in den Konstruktoraufruf von `make-dillo` gestellt hätten:

```
(define feed-dillo
  (lambda (dillo)
    (make-dillo (if (dillo-alive? dillo)
                    ...
                    ...))
    ...))
... (dillo-alive? dillo) ...
... (dillo-weight dillo) ...))
```

Bei dieser Vorgehensweise füllen wir zunächst die Ellipsen für die beiden möglichen Gewichte aus:

```
(define feed-dillo
  (lambda (dillo)
    (make-dillo (if (dillo-alive? dillo)
                    (+ (dillo-weight dillo) 500)
                    (dillo-weight dillo))
    ...))
... (dillo-alive? dillo) ...))
```

Was das zweite Argument von `make-dillo` betrifft, also ob das Gürteltier lebendig oder tot ist, so ist der Wert dort so wie vorher, also entsprechend dem Ergebnis von `dillo-alive?`:

```
(define feed-dillo
  (lambda (dillo)
    (make-dillo (if (dillo-alive? dillo)
                    (+ (dillo-weight dillo) 500)
```

```
(dillo-weight dillo))
(dillo-alive? dillo)))
```

Diese Version ist genauso korrekt wie die erste, und keine ist offensichtlich «besser» als die andere.<sup>1</sup>

Bei der Kombination von Konstruktionsanleitungen ist es also oft möglich, mehrere unterschiedliche Wege zu beschreiten. Meist funktionieren alle davon, sie unterscheiden sich aber gelegentlich in Länge und subjektiver Eleganz.

Eine andere Alternative bei der Konstruktion von `feed-dillo` wäre an dieser Stelle möglich gewesen:

```
(define feed-dillo
  (lambda (dillo)
    (if (dillo-alive? dillo)
        (make-dillo ... ...)
        dillo)))
```

Hier haben wir ziemlich vorschnell `dillo` in der Alternative des `if`-Ausdrucks geschrieben, weil es uns «offensichtlich» erschien, dass der Zustand eines toten Gürteltiers nach dem Füttern genauso wie vorher ist. Es wäre konsequenter gewesen, erst einmal die Schablonen vollständig anzuwenden:

```
(define feed-dillo
  (lambda (dillo)
    (if (dillo-alive? dillo)
        (make-dillo ... ...)
        (make-dillo ... ...))))
```

Dann hätten wir auch für den zweiten Zweig («Gürteltier tot») die Fragen beantwortet, welches Gewicht das Gürteltier hat und ob es noch lebt:

```
(define feed-dillo
  (lambda (dillo)
    (if (dillo-alive? dillo)
        (make-dillo (+ (dillo-weight dillo) 500) #t)
        (make-dillo (dillo-weight dillo) #f))))
```

Auch diese Version ist richtig. Wir könnten diese Version noch etwas umformen, in dem wir die Beobachtung ausnutzen, dass der Wert des `dillo-alive?`-Felds im Ergebnis dem Ergebnis von `(dillo-alive? dillo)` entspricht und das im zweiten `make-dillo`-Aufruf einsetzen:

<sup>1</sup> Die erste Version im Fall «totes Gürteltier» vermeidet, ein neues Gürteltier-Objekt zu erzeugen. Dafür ist die zweite Version kürzer. In der Praxis ist der Unterschied unwichtig.

```
(define feed-dillo
  (lambda (dillo)
    (if (dillo-alive? dillo)
        (make-dillo (+ (dillo-weight dillo) 500) #t)
        (make-dillo (dillo-weight dillo) (dillo-alive? dillo)))))
```

Das Programm ist zwar länger geworden, es gibt aber auch eine Einsicht frei, wenn wir den zweiten `make-dillo`-Aufruf näher anschauen. Dieser Aufruf stellt ein `dillo`-Record her, bei dem jedes Feld aus dem entsprechenden Feld von `dillo` bestückt wird. Es ist deshalb gleich zu `dillo`. Eine Gleichung bringt das zum Ausdruck:

$$(\text{make-dillo } (\text{dillo-alive? dillo}) \text{ (dillo-weight dillo)}) = \text{dillo}$$

Wir können also den `make-dillo`-Aufruf durch `dillo` ersetzen und so das Programm mit Hilfe der Gleichung vereinfachen:

```
(define feed-dillo
  (lambda (dillo)
    (if (dillo-alive? dillo)
        (make-dillo (+ (dillo-weight dillo) 500) #t)
        dillo)))
```

Jetzt ist die erste Version von `feed-dillo` entstanden, ohne dass wir uns vorschnell überlegen mussten, dass im Fall «Gürteltier tot» die Eingabe gleich der Ausgabe ist. In diesem Beispiel mögen die resultierende Einsicht und Vereinfachung banal sein. Wir können aber Gleichungen oft benutzen, um erstaunliche Einsichten zu erzielen oder unsere Programme kürzer, schneller oder eleganter zu machen. Das entspricht Mantra 6 von Seite 37:

#### MANTRA 6

Nutze Gleichungen aus, um Dein Programm zu vereinfachen.

#### AUFGABE 3.6

Formuliere Gleichungen entsprechend der für `dillo` für die Record-Typen `computer` und `wallclock-time` aus diesem Kapitel. □

## AUFGABEN

### AUFGABE 3.7

Schreibe eine Daten- und eine Record-Definition für *Brüche* und verschiedene Funktionen für das Bruchrechnen:

- Kürzen eines Bruchs
- Test auf Gleichheit der durch zwei Brüche repräsentierten rationalen Zahlen
- Addition, Subtraktion, Multiplikation und Division von Brüchen

**Hinweis:** Zur Lösung der Aufgabe ist die folgende eingebaute Funktion hilfreich:

```
(: gcd (natural natural -> natural)),
```

die den größten gemeinsamen Teiler («greatest common divisor») von zwei natürlichen Zahlen berechnet.

### AUFGABE 3.8

Jedes Qux hat einen Namen. Außerdem interessiert Experten, wieviele Bas ein Qux hat. Es wird außerdem zwischen Arg-Quxen, Foo-Quxen und Bla-Quxen unterschieden.

1. Schreibe eine Daten-Definition für Quxe sowie eine dazu passende Record-Definition. Notiere dazu auch die Signaturen der Selektoren.
2. Schreibe Signatur, Gerüst und Schablone für eine Funktion, die ein Qux akzeptiert und eine Zeichenkette zurückgibt. Identifiziere die dazu benutzten Konstruktionsanleitungen. Achte darauf, auch die Konstruktionsanleitungen für die Komponenten von Qux-Records anzuwenden.
3. Nimm an, Du hättest für eine zu schreibende Funktion `quxop2` die folgende Signatur festgelegt:

```
(: quxop2 (natural (enum "Hx" "Bx" "Px") -> qux))
```

(Dabei ist angenommen, dass die Record-Definition für ein Qux den Namen `qux` hat.) Entwickle daraus Gerüst und Schablone der zu schreibenden Funktion mit Hilfe der passenden Konstruktionsanleitungen.

### AUFGABE 3.9

Schreibe ein Programm zur Verwaltung von wöchentlichen Raumreservierungen an der Uni!



1. Entwerfe eine Daten- und Record-Definition für einen Eintrag eines Verwaltungssystems für Vorlesungs- und Seminarräume. Jeder Eintrag beinhaltet folgende Informationen: der Name des Raums (als Zeichenkette), der Wochentag, die Uhrzeit (es wird nur in Stunden gerechnet) und der Name des Dozenten, der den Raum belegt.
2. Schreibe eine Funktion `reserve`, die als Argumente einen Eintrag und einen Dozentennamen akzeptiert und einen Eintrag zurückgibt. Falls der Raum noch nicht belegt wurde (das heißt im Eintrag ist der Dozentennamenname «»), soll der Raum reserviert werden und damit ein neuer Eintrag zurückgegeben werden, bei dem der Dozentennamenname gesetzt ist. Andernfalls wird der Eintrag unverändert zurückgegeben.

#### AUFGABE 3.10

Schreibe weitere Funktionen für die Computer aus Abschnitt 3.1:

- Überlege Dir, wie Du für einen Computer einen geeigneten Preis abhängig von der Konfiguration berechnen würdest. Schreibe eine Funktion, welche Deine Methode realisiert.
- Schreibe eine Funktion, die den Speicher eines Computers erweitert. Sie akzeptiert einen Computer und eine Zahl und liefert einen neuen Computer, bei dem der Hauptspeicher um die Zahl erhöht ist.

#### AUFGABE 3.11

Es geht ums Backen von Kuchen.

1. Erstelle eine Datendefinition `dough` für den Teig. Jeder Teig besteht aus Eiern, Mehl, Zucker und Wasser und hat ein Gesamtgewicht. Überlege Dir geeignete Einheiten für die Zutaten.
2. Erstelle eine Datendefinition `cake` für Kuchen. Diese enthält einen Teig, eine Backdauer in Minuten und das Endgewicht des Kuchens.
3. Schreibe eine Funktion `ingredients->dough` welche eine Anzahl an Eiern, eine Menge Mehl in Gramm, eine Menge Zucker in Gramm und eine Menge Wasser in Milliliter erhält und daraus einen Teig herstellt. Gehe davon aus, dass jedes Ei 64g wiegt.
4. Schreibe eine Funktion `bake-cake`. Diese erhält einen Teig, eine Backdauer in Minuten und erstellt einen Kuchen. Gehe davon aus, dass nach dem Backen noch 80% des Wassers im Kuchen sind.

#### AUFGABE 3.12

Schreibe eine Datendefinition `appointment` für Termine, bestehend aus Datum, Uhrzeit, Dauer (in Minuten) und Ort. Verwende für das Datum und die Uhrzeit weitere Datendefinitionen bestehend aus Tag, Monat und Jahr beziehungsweise Stunde und Minute.

1. Schreibe eine Funktion `date-ok?`, die feststellt, ob ein Datums-Objekt einem tatsächlichen Kalenderdatum entspricht, also korrekte Daten wie 1.1.1970 von unsinnigen wie 34.17.2006 unterscheidet. Lasse dazu Schaltjahre außer Acht. Beachte die Monate mit 28, 30 und 31 Tagen.
2. Schreibe eine Funktion `date-equal?`, die vergleicht, ob zwei Datums-Objekte gleich sind.
3. Schreibe eine Funktion `time-ok?`, die feststellt, ob ein Zeit-Objekt einer tatsächlichen Uhrzeit entspricht.
4. Schreibe eine Funktion `time-overlap?`, die überprüft, ob sich zwei Zeiten mit einer jeweils gegebenen Dauer (in Minuten) überschneiden. Gehe davon aus, dass es sich um Zeiten desselben Tages handelt.
5. Schreibe eine Funktion `overlap?`, die prüft, ob sich zwei gegebene Termine überschneiden. Beachte die Dauer der Termine. Gehe davon aus, dass die Termine nicht über Mitternacht liegen.

**Hinweis:** Zur Lösung der Aufgabe kann die eingebaute Funktion

```
(: remainder (natural natural -> natural))
```

hilfreich sein. Sie berechnet den Rest einer ganzzahligen Division.

### AUFGABE 3.13

Erstelle eine Daten- und eine Record-Definition für einen Fahrzeugschein (siehe Abbildung 3.2). Gliedere die Felder des Fahrzeugscheins sinnvoll in Untergruppen und erstelle für diese Untergruppen eigene Daten- und Record-Definitionen. Benutze sprechende Bezeichner für Records und Felder! Gib ein Beispiel an, indem Du ein Fahrzeugschein-Objekt mit allen Einträgen erzeugst.

### AUFGABE 3.14

Schreibe für den Tübinger Stadtverkehr ein Programm, welches überprüft, ob ein Fahrzeug in den Umweltzonen fahren darf.

1. Definiere einen Datentyp für Fahrzeuge. Dieser Datentyp soll den Typ, das Nummernschild und die Schadstoffklasse des Fahrzeuges beinhalten.
2. Erstelle die Beispielfahrzeuge für die Fahrzeugtypen «Stadtbus», «Reisebus», «Dieselauto» und «Benzinauto». Gehe davon aus, dass die Busse der Schadstoffklasse 2, das Dieselauto der

Schadstoffklasse 3 und das Benzinauto der Schadstoffklasse 4 angehören.

3. Schreibe eine Funktion `fahrverbot?`, welche überprüft, ob ein gegebenes Fahrzeug bei einer gegebenen Mindest-Schadstoffklasse noch fahren darf. Gestalte die Signatur so, dass er nur Mindest-Schadstoffklassen von 1 bis 4 akzeptiert.
4. Die Mühlstraße in Tübingen ist in einer Richtung für alle Fahrzeuge außer Stadtbusse gesperrt. Schreibe eine Funktion `sonderrecht?`, die überprüft, ob ein gegebenes Fahrzeug die Mühlstraße in der gesperrten Richtung befahren darf.
5. Der Stadtrat hat die Idee, den Tourismus dadurch anzukurbeln, dass Sonntags auch Reisebusse die Mühlstraße in der gesperrten Richtung befahren dürfen. Erweitere hierfür die Funktion `sonderrecht?` um den Wochentag und lasse sonntags auch Reisebusse zu.

Verwende beim Schreiben der Funktion die Konstruktionsanleitungen für Funktionen und für Fallunterscheidungen. Schreibe Testfälle, die alle Möglichkeiten der Fallunterscheidung abdecken.

### AUFGABE 3.15

Schreibe ein Programm für einen Paketdienst, das den Preis für ein Paket berechnet!

1. Schreibe eine Daten- und eine Record-Definition für *Adressen*. Zu einer Adresse gehören der Name, die Straße mit Hausnummer, die Postleitzahl, der Ort und das Land.
2. Der Paketdienst verlangt einen Zuschlag für Sendungen, die international verschickt werden. Schreibe eine Funktion `international?`, die als Argument eine *Adresse* bekommt und feststellt, ob die Adresse im Ausland liegt.
3. Der Paketdienst hat einen Sondertarif für Sendungen, die innerhalb der gleichen Postleitzahl verschickt werden. Schreibe eine Funktion `same-zip-code?`, die als Argumente zwei *Adressen* bekommt und feststellt, ob die Postleitzahlen und die Länder der beiden Adressen gleich sind.
4. Ein Paket wird klassifiziert nach seinen Abmessungen. Schreibe eine Daten- und eine Record-Definition für *Abmessungen*. Abmessungen bestehen aus Länge, Breite und Höhe.
5. Die Paketpreise richten sich nach der Größe des zu verschickenden Pakets. Der Paketdienst verwendet die drei Größenklassen *Small*, *Medium* und *Large*, um die Kosten für das Paket zu berechnen. Ausschlaggebendes Kriterium für die Paketgröße ist die Summe der längsten und der kürzesten Seite des Pakets. Schreibe eine Funktion `add-longest-and-shortest-side`, die als Argument eine *Abmessung* bekommt. Der Rückgabewert von `add-longest-and-shortest-side` soll die Summe der längsten und der kürzesten Seite der Abmessung sein. Lagere die Teilprobleme in zwei Hilfsfunktionen aus: `longest-side` und `shortest-side`.
6. Schreibe eine Daten- und eine Record-Definition für *Pakete*. Ein Paket besteht aus einer Absender-

und einer Empfängeradresse. Benutze für die Adressen die bereits erstellte Record-Definition. Außerdem hat ein Paket noch weitere Eigenschaften: Die Abmessungen (benutze dafür die bereits erstellte Record-Definition), das Gewicht, die Beförderungsdauer und eine Zusatzoption Nachnahme. Die Beförderungsdauer soll *normal*, *next-day* oder *next-morning* sein.

7. Schreibe eine Funktion `parcel-size-class`, die als Argument ein *Paket* bekommt und die Größenklasse zurückgibt. Ausschlaggebend für die Paketgröße ist die Abmessung (siehe oben). Folgende Tabelle enthält die Zuordnung von Paketgröße und Abmessung:

Paketgröße	Abmessung
S	0–50 cm
M	>50–100 cm
L	>100 cm

8. Schreibe eine Funktion `calculate-base-postage`, die als Argument ein *Paket* bekommt und die Basis-Portokosten für dieses Paket berechnet. Lege Sie dabei folgende Grundtarifabelle des Paketdienstes zugrunde:

Paketgröße	Gewicht		
	0–5 kg	>5–10 kg	>10 kg
S	3,00	6,00	9,00
M	6,00	10,00	14,00
L	9,00	15,00	21,00

9. Schreibe eine Funktion `transportation-time-factor`, die als Argument ein *Paket* bekommt und den Aufschlagsfaktor für die Beförderungsdauer zurückliefert. Lege dabei folgende Aufschlagsfaktoren zugrunde:

Beförderungsdistanz	Beförderungsdauer		
	normal	next-day	next-morning
gleiche PLZ	-25%	+0%	+25%
Inland	+0%	+50%	+100%
Ausland	+100%	+200%	+300%

10. Schreibe eine Funktion `cash-on-delivery-surcharge`, die als Argument ein *Paket* bekommt und den Aufschlag für die Nachnahme zurückliefert. Lege dabei folgende Aufschläge zugrunde:

Beförderungsdistanz	Nachnahmegebühr
Inland	+3,00
Ausland	+9,00

- ii. Schreibe eine Funktion `calculate-postage`, die als Argument ein *Paket* bekommt und die Portokosten berechnet. Benutze dafür die bereits programmierten Lösungen der verschiedenen Teilprobleme.

### AUFGABE 3.16

Auf seinen Reisen um die Welt trifft Dr. Sperber viele interessante und skurrile Zeitgenossen. Unter ihnen Dr. Knaubichler, ein Experte auf dem Gebiet der Kreuzung von mystischen Kreaturen. Es gibt drei klassische Grundkreaturen:

- Der Garnolaf, der Stärke besitzt.
- Das Ronugor, das Wissen besitzt.
- Der Tschipotol, der Risikobereitschaft besitzt.

Die Merkmale der Kreaturen sind unterschiedlich ausgeprägt. Die Knaubichler-Kreaturen-merkmal-Skala geht von 0 bis 100. Die Grundkreaturen haben jeweils nur ein Merkmal, keine besitzt ein Merkmal einer anderen Grundkreatur.

Nun kreuzt Dr. Knaubichler die Grundkreaturen miteinander. Es entstehen also

- Das Ronulaf, mit Wissen und Stärke.
- Der Tschigor, mit Wissen und Risikobereitschaft.
- Das Lapotol, mit Stärke und Risikobereitschaft.
- Der Tschirgaronu, Wissen, Stärke und Risikobereitschaft.

Leider erben die Kreuzungen nicht die vollen Merkmale beider Grundkreaturen.

Bei der Kreuzung unterschiedlicher Grundkreaturen gilt folgendes: Bei der Kreuzung von Ronugor und Tschipotol wird das übernommene Wissen um 10% verringert; bei Tschipotol und Garnolaf hingegen nimmt die Risikobereitschaft um 5% ab, aber die Stärke legt um 8% zu. Garnolaf und Ronugor lassen die Stärke um 5% zulegen. Kreuzt man alle drei Grundkreaturen, nimmt jede Eigenschaft um 3% ab.

Werden zwei Kreaturen der gleichen Sorte gekreuzt, so entsteht eine neue Kreatur mit  $\frac{2}{3}$  der addierten Eigenschaften der beiden Grundkreaturen. Man kann nur Grundkreaturen kreuzen.

In jedem Fall kann eine Eigenschaft maximal den Wert 100 haben.

Dr. Knaubichler braucht ein Programm, welches die neue Kreatur berechnet, bevor er die Kreuzungen durchführt. Hilf ihm dabei!

1. Mache eine Datenanalyse und erstelle passende Daten- und Record-Definitionen. Gib alle Signaturen der Record-Funktionen an!
2. Schreibe für jede der oben aufgelisteten Kreuzungen von zwei Grundkreaturen eine Funktion, die die Kreuzung vornimmt und eine neue Kreatur zurückgibt.



## 4 GEMISCHTE DATEN

Manchmal kommen an einer Stelle in unserem Problem verschiedene Klassen der gleichen Sorte Daten vor:

- Ein Tier kann ein Gürteltier oder ein Papagei sein.
- Eine Koordinate kann eine kartesische Koordinate oder eine Polarkoordinate sein.
- Ein Essen kann ein Frühstück, Mittagessen oder Abendessen sein.

Solche Daten heißen *gemischte Daten*, und es handelt sich um eine weitere Form von Fallunterscheidungen, neben den aus Kapitel 2 bekannten Aufzählungen und Zahlenbereichen.

Obwohl die Daten verschiedenartig sind, unterstützen sie doch gemeinsame Operationen: Das Gewicht eines Tiers kann sowohl für Gürteltiere als auch Papageien berechnet werden, der Abstand vom Ursprung kann für beide Koordinatendarstellungen berechnet werden, die Anzahl der Gänge kann für jede Art Essen bestimmt werden.

### 4.1 GEMISCHTE DATEN

In der Einleitung war die Rede von Papageien: die benutzen wir, um gemischte Daten einzuführen. Vorher müssen wir jedoch Papageien mit den bekannten Mitteln definieren. Wir erweitern dafür die Datei mit dem Gürteltier aus dem vorigen Kapitel.

Genau wie bei Gürteltieren interessiert uns bei Papageien das Gewicht, aber wir nehmen an, da Papageien in der Regel nicht auf texanischen Highways überfahren werden, dass sie immer lebendig sind. Außerdem betrachten wir ausschließlich sprechende Papageien, die jeweils einen einzelnen Satz sagen können. Hier die Datendefinition:

```
; Ein Papagei hat folgende Eigenschaften:  
; - Gewicht in Gramm  
; - Satz, den er sagt
```

Hier die dazu passende Record-Definition:

```
(define-record parrot  
  make-parrot  
  (parrot-weight natural)  
  (parrot-sentence string))
```

... und die passenden Signaturen:

```
(: make-parrot (natural string -> parrot))
(: parrot-weight (parrot -> natural))
(: parrot-sentence (parrot -> string))
```

Hier zwei Beispiele für Papageien mit Kommentaren, die die Beziehung zwischen Daten und Information beschreiben:

```
(define parrot1 (make-parrot 10000 "Der Gärtner war's.")) ; 10kg, Miss
    Marple
(define parrot2 (make-parrot 5000 "Ich liebe Dich.")) ; 5kg,
    Romantiker
```

Du kannst Dir vielleicht vorstellen, dass Papageien und Gürteltiere sich in einem Programm begegnen, also *gemischt* vorkommen. Papageien und Gürteltiere gehören zum gemeinsamen Oberbegriff *Tier*. Dafür könnte eine Beschreibung so aussehen:

Ein *Tier* ist eins der folgenden:

- ein Gürteltier
- ein Papagei

Die Formulierung «eins der folgenden» ist Dir schon aus Abschnitt 2.4 und Konstruktionsanleitung 7 auf Seite 66 bekannt: Sie deutet auf eine Fallunterscheidung hin.

Da aber Gürteltiere und Papageien nicht als nur jeweils ein Wert repräsentiert sind, handelt es sich nicht um eine Aufzählung: Die beiden Fälle der Fallunterscheidung beschreiben unterschiedliche *Klassen* von Tieren, jede mit ihrer eigenen Signatur. Damit liegt eine neue Organisation von Daten vor: *gemischte Daten*. Entsprechend ist es durchaus sinnvoll, nach dem «Gewicht eines Tiers» zu fragen oder «ein Tier zu füttern», was wir im folgenden auch vorhaben.

Die Beschreibung des Begriffs «Tier» ist bereits als Datendefinition geeignet, und muss für Inklusion im Programm nur als Kommentar umformatiert werden:

```
; Ein Tier ist eins der folgenden:
; - Gürteltier
; - Papagei
```

Bei zusammengesetzten Daten kann die Datendefinition in eine Record-Definition überführt werden. In diesem Fall ist für jede einzelne Klasse *Tier* jeweils schon eine Record-Definition da. Wenn wir Tiere im allgemeinen in Funktionen verarbeiten wollen, brauchen wir allerdings eine Signatur für Tiere. Zum Beispiel wollen wir eine Funktion schreiben, die für jedes Tier das Gewicht ermittelt. Diese Funktion könnte folgende Signatur haben:

```
(: animal-weight (animal -> natural))
```



Wir brauchen also eine Definition für die Signatur `animal`. Diese sieht folgendermaßen aus:

```
(define animal
  (signature
    (mixed dillo parrot)))
```

Das `signature` kennen wir von den Fallunterscheidungen aus Abschnitt 2.4 auf Seite 49. Das `mixed` ist neu und steht für «gemischte Daten», also Fallunterscheidungen, bei denen jeder Fall seine eigene Signatur hat. Du kannst die obige Definition lesen als «Tiere sind gemischt aus Gürteltieren und Papageien»; das klingt aber auf Deutsch hölzern, weshalb wir für die Datendefinition bei der Formulierung «eins der folgenden» bleiben. Mit der Definition steht die Signatur `animal` zur Verfügung. Wir haben schon mit der Signatur für `animal-weight` vorgegriffen. Hier ist sie noch einmal zusammen mit einer Kurzbeschreibung:

```
; Gewicht eines Tiers feststellen
(: animal-weight (animal -> natural))
```

Diese Funktion sollte entsprechend für Gürteltiere *und* Papageien funktionieren, wir brauchen also Testfälle für beide:

```
(check-expect (animal-weight dillo1) 55000)
(check-expect (animal-weight dillo2) 58000)
(check-expect (animal-weight parrot1) 10000)
(check-expect (animal-weight parrot2) 5000)
```

Das Gerüst sieht so aus:

```
(define animal-weight
  (lambda (animal)
    ...))
```

Tiere bilden auch eine Fallunterscheidung in den Daten, mit zwei Fällen: Gürteltiere und Papageien. Im Rumpf der Funktion brauchen wir also eine Verzweigung mit zwei Zweigen:

```
(define animal-weight
  (lambda (animal)
    (cond
      (... ...)
      (... ...))))
```

Wir brauchen als Nächstes zwei Bedingungen – eine, die Gürteltiere und eine, die Papageien identifiziert. Dafür erweitern wir die Record-Definitionen um ein neues Element, das *Prädikat*. Die Prädikate werden uns erlauben, die Bedingungen zu schreiben. Die Record-Definitionen sehen dann so aus:

```
(define-record dillo
  make-dillo
  dillo? ;  $\Leftarrow$  Prädikat
  (dillo-weight natural)
  (dillo-alive? boolean))
```

```
(define-record parrot
  make-parrot
  parrot? ;  $\Leftarrow$  Prädikat
  (parrot-weight natural)
  (parrot-sentence string))
```

Es sind zwei neue Namen hinzugekommen, `dillo?` und `parrot?`. Das sind die Prädikate, und sie haben folgenden Signaturen:

```
(: dillo? (any -> boolean))
(: parrot? (any -> boolean))
```

Das `any` ist auch neu und ist die Signatur für einen beliebigen Wert: Ein Prädikat kann also auf absolut alles angewendet werden. Die beiden Prädikate unterscheiden Gürteltiere beziehungsweise Papageien von anderen Werten:

```
(dillo? dillo1)
 $\hookrightarrow$  #t
(dillo? parrot1)
 $\hookrightarrow$  #f
(parrot? dillo1)
 $\hookrightarrow$  #f
(parrot? parrot1)
 $\hookrightarrow$  #t
(dillo? 5)
 $\hookrightarrow$  #f
(parrot? "foo")
 $\hookrightarrow$  #f
```

Im allgemeinen ist ein Prädikat eine Funktion mit der Signatur `(any -> boolean)`, die für eine bestimmte Sorte von Werten `#t` liefert, für alle andere Werte aber `#f`. In diesem speziellen Fall identifizieren die Prädikate `dillo?` und `parrot?` jeweils alle Werte, die zu einem bestimmten Record-Typ gehören. Prädikate müssen aber nicht zu einem Record-Typ gehören, wir können ein Prädikat auch selber schreiben. Außerdem sind eine Reihe von Prädikaten vordefiniert – dazu mehr später in diesem Kapitel auf 121.

#### AUFGABE 4.1

Schreibe ein Prädikat namens `animal?` für alle Tiere, also alle Werte, die ein `dillo`-Record oder ein `parrot`-Record sind. □

Die Prädikate `dillo?` und `parrot?` können wir benutzen, um die Bedingungen aus der Schablone zu bestücken:

```
(define animal-weight
  (lambda (animal)
    (cond
      ((dillo? animal) ...)
      ((parrot? animal) ...))))
```

Im ersten Zweig – dem Zweig für Gürteltiere – kommt nun `dillo-weight` zum Einsatz, im zweiten Zweig – für Papageien – ist `parrot-weight` zuständig:

```
(define animal-weight
  (lambda (animal)
    (cond
      ((dillo? animal) (dillo-weight animal))
      ((parrot? animal) (parrot-weight animal)))))
```

Fertig!

Abbildung 4.1 ist eine aktualisierte Beschreibung der Form von `define-record`, bei der auch Prädikate berücksichtigt sind

Aus diesem Beispiel ergibt sich eine Konstruktionsanleitung für gemischte Daten. Zunächst die Datenanalyse:

#### KONSTRUKTIONSANLEITUNG 15 (GEMISCHTE DATEN: DATENANALYSE)

Gemischte Daten sind Fallunterscheidungen, bei denen jeder Fall eine eigene Klasse von Da-

Eine `define-record`-Form hat folgende allgemeine Gestalt:

```
(define-record t
  c
  p
  (sel1 sig1)
  ...
  (seln sign))
```

Diese Form definiert einen Record-Typ mit  $n$  Feldern. Dabei sind  $t$ ,  $c$ ,  $sel_1 \dots sel_n$  allesamt Variablen, für die `define-record` Definitionen anlegt:

- $t$  ist der Name der Record-Signatur.
- $c$  ist der Name des Konstruktors mit folgender Signatur:

```
(: c (sig1 ... sign -> t))
```

- $p$  ist der Name des Prädikats, der Records diesen Typs von anderen Werten unterscheidet. Das Prädikat kann auch weggelassen werden. Das Prädikat hat folgende Signatur:

```
(: p (any -> boolean))
```

- $sel_1, \dots, sel_n$  sind die Namen der Selektoren für die Felder des Record-Typs. Der Selektor  $sel_i$  hat folgende Signatur:

```
(: seli (t -> sigi))
```

**Abbildung 4.1:** `define-record` (mit Prädikaten)

ten mit eigener Signatur ist. Schreibe bei gemischten Daten eine Signatur-Definition der folgenden Form unter die Datendefinition:

```
(define sig
  (signature
    (mixed sig1 sig2 ... sign)))
```

$Sig$  ist die Signatur für die neue Datensorte;  $sig_1$  bis  $sig_n$  sind die Signaturen, aus denen die neue Datensorte zusammengemischt ist.

Die Konstruktionsanleitung für die Schablone ist der aus der Konstruktionsanleitung 10 für Fallunterscheidungen auf Seite 68 ähnlich:

#### KONSTRUKTIONSANLEITUNG 16 (GEMISCHTE DATEN ALS EINGABE: SCHABLONE)

Eine Schablone für eine Funktion und deren Testfälle, die gemischte Daten akzeptiert, kannst

Du folgendermaßen konstruieren:

- Schreibe Tests für jeden der Fälle.
- Schreibe eine `cond`-Verzweigung als Rumpf in die Schablone, die genau  $n$  Zweige hat – also genau so viele Zweige, wie es Fälle in der Datendefinition beziehungsweise der Signatur gibt.
- Schreibe für jeden Zweig eine Bedingung, der den entsprechenden Fall identifiziert.
- Vervollständige die Zweige, indem Du eine Datenanalyse für jeden einzelnen Fall vornimmst und entsprechende Hilfsfunktionen und Konstruktionsanleitungen benutzt. Die übersichtlichsten Programme entstehen meist, wenn für jeden Fall separate Hilfsfunktionen definiert sind.

Eine Konstruktionsanleitung oder Schablone für gemischte Daten *als Ausgabe* ist unnötig – Du benutzt einfach die Schablone des entsprechenden Falls.

Beachte den Unterschied zwischen `enum` und `mixed`, die leicht zu verwechseln sind: `enum` steht für «einer der folgenden *Werte*», während `mixed` für «gehörend zu einer der folgenden *Signaturen*» steht.

#### AUFGABE 4.2

Denk Dir sinnvolle Datendefinitionen für Frühstück, Mittagessen und Abendessen aus (ein ... essen besteht aus ...), so dass es Sinn ergibt, für jede Sorte Essen die Frage «Was für ein Gemüse ist enthalten?» zu stellen. Schreibe zunächst für jede Sorte Essen eine Funktion, die das enthaltene Gemüse liefert. Schreibe dann eine Datendefinition für den Oberbegriff «Essen»; ein Essen kann ein Frühstück, ein Mittagessen oder ein Abendessen sein. Schreibe eine Funktion, welche das enthaltene Gemüse für ein Essen liefert. □

## 4.2 GEMISCHTE DATEN ALS AUSGABE

Wir können einen Papagei ähnlich wie ein Gürteltier füttern – nur die Portion ist kleiner, wir nehmen 50 g an. Kurzbeschreibung und Signatur:

```
; Papagei mit 50 g Futter füttern
(: feed-parrot (parrot -> parrot))
```

Testfälle:

```
(check-expect (feed-parrot parrot1) (make-parrot 10050 "Der Gärtner
war's."))
```

```
(check-expect (feed-parrot parrot2) (make-parrot 5050 "Ich liebe Dich.
"))
```

Gerüst:

```
(define feed-parrot
  (lambda (parrot)
    ...))
```

Die Schablone entsteht aus der Kombination der Schablonen für zusammengesetzte Daten als Eingabe (Konstruktionsanleitung 13 auf Seite 93) und als Ausgabe (Konstruktionsanleitung 14 auf Seite 94):

```
(define feed-parrot
  (lambda (parrot)
    (make-parrot ... ...)
    ... (parrot-weight parrot) ...
    ... (parrot-sentence parrot) ...))
```

... und schließlich der vollständige Rumpf:

```
(define feed-parrot
  (lambda (parrot)
    (make-parrot (+ (parrot-weight parrot) 50)
                  (parrot-sentence parrot))))
```

Nun können wir eine Funktion schreiben, die ein beliebiges Tier akzeptiert und es füttert. Hier Kurzbeschreibung und Signatur:

#### AUFGABE 4.3

```
; Tier füttern (: feed-animal (animal -> animal))
```

Die Funktion soll sich genauso verhalten wie `feed-dillo` beziehungsweise `feed-parrot`. Das können wir direkt als Tests ausdrücken:

```
(check-expect (feed-animal parrot1) (feed-parrot parrot1))
(check-expect (feed-animal parrot2) (feed-parrot parrot2))
(check-expect (feed-animal dillo1) (feed-dillo dillo1))
(check-expect (feed-animal dillo2) (feed-dillo dillo2))
```

Das Gerüst sieht so aus:

```
(define feed-animal
  (lambda (animal)
    ...))
```

Da die Eingabe gemischt ist, brauchen wir eine Verzweigung mit soviel Zweigen wie `animal` Fälle hat, also zwei, die wir gleich mit den passenden Bedingungen versehen:

```
(define feed-animal
  (lambda (animal)
    (cond
      ((dillo? animal) ...)
      ((parrot? animal) ...))))
```

Schließlich müssen wir noch die Antworten ergänzen, für die wir die Funktionen benutzen, die wir schon geschrieben haben:

```
(define feed-animal
  (lambda (animal)
    (cond
      ((dillo? animal) (feed-dillo animal))
      ((parrot? animal) (feed-parrot animal)))))
```

Das Beispiel zeigt, dass wir keine spezielle Konstruktionsanleitung für gemischte Daten als Ausgabe brauchen: Wir müssen nur darauf achten, den jeweils richtigen Fall zu liefern.

#### AUFGABE 4.4

Schreibe eine Funktion `run-over-animal`, die ein Tier überfährt!

□

### 4.3 DIE LEBENSMITTELAMPEL

In diesem Abschnitt nehmen wir uns noch ein weiteres Beispiel für gemischte Daten vor, diesmal von vorneherein unter Benutzung der Konstruktionsanleitung aus dem vorigen Abschnitt.

Die Lebensmittelampel ist eine Kennzeichnung auf Lebensmittelverpackungen, die den Gehalt von gesundheitsrelevanten Nährstoffen in den Ampfelfarben ausweist: Grün für niedrigen Gehalt, Gelb für mittleren Gehalt und Rot für hohen Gehalt. Bei Zucker zum Beispiel sieht die Ampel so aus, bezogen auf 100 g eines Lebensmittels:

grün	niedriger Gehalt	weniger als 5 g
gelb	mittlerer Gehalt	zwischen 5 g und 22,5 g
rot	hoher Gehalt	mehr als 22,5 g

Trotz der Bemühungen der Europäischen Union sind die Bezeichnungen uneinheitlich. Technisch gesehen ist die Ampel redundant, wenn der Zuckergehalt in Gramm angegeben ist. Manchmal ist allerdings der Zuckergehalt auch separat für Fruktose und Glukose angegeben.

Ein Computerprogramm könnte aber den Umgang erleichtern, indem es jede Angabe auf einer Lebensmittelpackung – Zucker in Gramm insgesamt, Fruktose und Glukose separat sowie die Ampel – in die einheitliche Ampel-Form bringt. Schreiben wir also ein solches Programm. Zunächst die Datenanalyse:

- Die Zuckermenge in Gramm ist eine (rationale) Zahl.
- Zuckeranteile bestehen aus der Menge von Fruktose und Glukose.
- Eine Zuckerampel ist rot, gelb oder grün.
- Der Zuckergehalt kann entweder als Zuckermenge, Zuckeranteile oder Zuckerampel angegeben werden.

Hier ist die Datendefinition für die zusammengesetzten Zuckeranteile:

```
; Zuckeranteile bestehen aus:
; - Fruktose-Menge (in g)
; - Glukose-Menge (in g)
```

Daraus ergibt sich direkt die Record-Definition mit zwei Komponenten, jeweils rationale Zahlen:

```
(define-record sugars
  make-sugars
  sugars?
  (sugars-fructose-g rational)
  (sugars-glucose-g rational))
```

Hier einige Beispiele, zusammen mit Kommentaren, die die Beziehung zwischen Daten und Information beschreiben:

```
(define s1 (make-sugars 1 1)) ; 1 g Fruktose, 1 g Glukose
(define s2 (make-sugars 2 3)) ; 2 g Fruktose, 3 g Glukose
(define s3 (make-sugars 5 5)) ; 5 g Fruktose, 5 g Glukose
(define s4 (make-sugars 10 2.5)) ; 10 g Fruktose, 2.5 g Glukose
(define s5 (make-sugars 10 13)) ; 10 g Fruktose, 13 g Glukose
(define s6 (make-sugars 15 10)) ; 15 g Fruktose, 10 g Glukose
```



Bei der Ampel selbst handelt es sich um eine einfache Fallunterscheidung:

```
; Eine Ampel ist einer der folgenden Werte:  
; - rot  
; - gelb  
; - grün
```

Das ist eine Aufzählung, wir geben entsprechend der dazu passenden Signatur einen Namen:

```
(define traffic-light  
  (signature  
    (enum "red" "yellow" "green")))
```

Die Angabe über den Zuckergehalt kann jede der drei oben genannten Formen annehmen:

```
; Ein Zuckergehalt ist eins der folgenden:  
; - Gewicht in Gramm  
; - Zuckeranteile  
; - Ampelbezeichnung
```

Wieder ist an der Formulierung erkennbar, dass es sich um gemischte Daten handelt, und zwar mit drei Fällen. Das übersetzen wir in ein Signatur-Definition mit `mixed` und ebenfalls drei Fällen:

```
(define sugar-content  
  (signature  
    (mixed rational  
      sugars  
      traffic-light)))
```

Das Beispiel zeigt, dass die Fälle einer Definition für gemischte Daten nicht allesamt Records sein müssen. Es ist allerdings wichtig, dass die Fälle *disjunkt* sind, also jeder Wert eindeutig einem der Fälle zugeordnet werden kann: Sonst wäre es nicht möglich, eine sinnvolle Verzweigung zu schreiben, welche die Fälle unterscheidet.

Nun zu unserer Funktion zur Ermittlung der Ampelbezeichnung für den Zuckergehalt. Hier Kurzbeschreibung und Signatur:

```
; Ampelbezeichnung für Zuckergehalt ermitteln  
(: sugar-traffic-light (sugar-content -> traffic-light))
```

Wir brauchen ziemlich viele Testfälle, um alle Fälle von Zuckergehalt abzudecken sowie die Eckfälle der Tabelle von oben.

```

(check-expect (sugar-traffic-light 2) "green")
(check-expect (sugar-traffic-light 5) "yellow")
(check-expect (sugar-traffic-light 10) "yellow")
(check-expect (sugar-traffic-light 12.5) "yellow")
(check-expect (sugar-traffic-light 23) "red")

(check-expect (sugar-traffic-light s1) "green")
(check-expect (sugar-traffic-light s2) "yellow")
(check-expect (sugar-traffic-light s3) "yellow")
(check-expect (sugar-traffic-light s4) "yellow")
(check-expect (sugar-traffic-light s5) "red")
(check-expect (sugar-traffic-light s6) "red")

(check-expect (sugar-traffic-light "green") "green")
(check-expect (sugar-traffic-light "yellow") "yellow")
(check-expect (sugar-traffic-light "red") "red")

```

Als Nächstes ist, wie immer, das Gerüst dran:

```

(define sugar-traffic-light
  (lambda (sugar-content)
    ...))

```

Als Nächstes wenden wir die Schablone für Funktionen an, die gemischte Daten akzeptieren. Wir brauchen eine Verzweigung mit sovielen Zweigen wie `sugar-content` Fälle hat, also drei:

```

(define sugar-traffic-light
  (lambda (sugar-content)
    (cond
      (... ...)
      (... ...)
      (... ...))))

```

Als Nächstes brauchen wir Tests für die drei Fälle. Für den zweiten Fall ist das einfach, da es sich um `sugars`-Records handelt: da gibt es das Prädikat `sugars?`. Beim ersten Fall handelt es sich aber um eine rationale Zahl, beim dritten um eine Zeichenkette – beides eingebaute Datensorten. Für diese gibt es die eingebauten Prädikate `rational?` und `string?` – Abbildung 4.2 zählt noch mehr eingebaute Prädikate auf. Mit dieser Information gewappnet können wir die Tests ergänzen:

Folgende Prädikate sind eingebaut:

- `number?` testet, ob ein Wert eine Zahl ist.
- `real?` testet, ob ein Wert eine reelle Zahl ist.
- `rational?` testet, ob ein Wert eine rationale Zahl ist.
- `natural?` testet, ob ein Wert eine natürliche Zahl ist.
- `string?` testet, ob ein Wert eine Zeichenkette ist.
- `boolean?` testet, ob ein Wert ein boolescher Wert ist.

**Abbildung 4.2:** Eingebaute Prädikate

```
(define sugar-traffic-light
  (lambda (sugar-content)
    (cond
      ((rational? sugar-content) ...)
      ((sugars? sugar-content) ...)
      ((string? sugar-content) ...))))
```

Im ersten Zweig handelt es sich nicht nur um eine rationale Zahl, sondern auch um eine Fallunterscheidung mit drei Fällen entsprechend der Tabelle vom Anfang:

```
(define sugar-traffic-light
  (lambda (sugar-content)
    (cond
      ((rational? sugar-content)
        (cond
          (... ...)
          (... ...)
          (... ...)))
      ((sugars? sugar-content) ...)
      ((string? sugar-content) ...))))
```

Als Nächstes ergänzen wir Tests entsprechend der Tabelle:

```
(define sugar-traffic-light
  (lambda (sugar-content)
    (cond
      ((rational? sugar-content)
        (cond
          ((< sugar-content 5) ...)
          ((and (>= sugar-content 5) (<= sugar-content 12.5)) ...)
          ((> sugar-content 12.5) ...))))
```

```
((sugars? sugar-content) ...)
((string? sugar-content) ...)))
```

Schließlich müssen wir noch die Antworten eintragen:

```
(define sugar-traffic-light
  (lambda (sugar-content)
    (cond
      ((rational? sugar-content)
       (cond
         ((< sugar-content 5) "green")
         ((and (>= sugar-content 5) (<= sugar-content 12.5)) "yellow")
         ((> sugar-content 12.5) "red"))))
      ((sugars? sugar-content) ...)
      ((string? sugar-content) ...))))
```

Das verschachtelte `cond` ist etwas unübersichtlich: In Konstruktionsanleitung 16 auf Seite 115 ist bereits vermerkt, dass es sinnvoll ist, diesen Zweig in eine separate Hilfsfunktion auszulagern. Hier sind Kurzbeschreibung und Signatur für diese Hilfsfunktion:

```
; Zuckeranteil in g in Ampel umwandeln
(: sugar-weight->traffic-light (rational -> traffic-light))
```

Die Testfälle lassen sich aus den Testfällen für `sugar-traffic-light` durch einfaches Kopieren und Umbenennen gewinnen:

```
(check-expect (sugar-weight->traffic-light 2) "green")
(check-expect (sugar-weight->traffic-light 5) "yellow")
(check-expect (sugar-weight->traffic-light 10) "yellow")
(check-expect (sugar-weight->traffic-light 12.5) "yellow")
(check-expect (sugar-weight->traffic-light 23) "red")
```

Gerüst:

```
(define sugar-weight->traffic-light
  (lambda (sugar-weight)
    ...))
```

Den Rumpf haben wir ja schon geschrieben, wir müssen ihn nur noch hierher bewegen und `sugar-content` in `sugar-weight` umbenennen. Das DrRacket-System bietet dazu nach einem Klick

auf «Syntaxprüfung» die Möglichkeit, mit einem Rechtsklick auf den Parameter `sugar-content` ein Menü aufzuklappen, das unter anderem die Auswahl «`sugar-content` umbenennen» anbietet. DrRacket sorgt dann dafür, dass alle zugehörigen Vorkommen von `sugar-content` in gleicher Weise umbenannt werden.

```
(define sugar-weight->traffic-light
  (lambda (sugar-weight)
    (cond
      ((< sugar-weight 5) "green")
      ((and (>= sugar-weight 5) (<= sugar-weight 22.5)) "yellow")
      ((> sugar-weight 12.5) "red"))))
```

Zurück zu `sugar-traffic-light`: Dort benutzen wir zunächst die neu definierte Hilfsfunktion:

```
(define sugar-traffic-light
  (lambda (sugar-content)
    (cond
      ((rational? sugar-content)
       (sugar-weight->traffic-light sugar-content))
      ((sugars? sugar-content) ...)
      ((string? sugar-content) ...))))
```

Beim nächsten Zweig geht es um den Fall `sugars` – zusammengesetzte Daten. Wir können also die entsprechende Schablone anwenden:

```
(define sugar-traffic-light
  (lambda (sugar-content)
    (cond
      ((rational? sugar-content)
       (sugar-weight->traffic-light sugar-content))
      ((sugars? sugar-content) ... (sugars-fructose-g sugar-content)
       ...
       ... (sugars-glucose-g sugar-content) ...)
      ((string? sugar-content) ...))))
```

Wir müssen Fruktose- und den Glukose-Anteil addieren und die Summe dann ebenfalls entsprechend der Tabelle vom Anfang in eine Ampelfarbe umwandeln. Aber halt! – Genau für das Umwandeln der Zahl aus der Tabelle in eine Ampelfarbe haben wir ja gerade die Hilfsfunktion `sugar-weight->traffic-light` geschrieben, und diese können wir erneut zum Einsatz bringen:

```
(define sugar-traffic-light
  (lambda (sugar-content)
    (cond
      ((rational? sugar-content)
       (sugar-weight->traffic-light sugar-content))
      ((sugars? sugar-content)
       (sugar-weight->traffic-light
        (+ (sugars-fructose-g sugar-content)
           (sugars-glucose-g sugar-content))))
      ((string? sugar-content) ...))))
```

Bleibt der letzte Fall – der ist zum Glück trivial, da es sich schon um eine Farbe handelt, Die muss `sugar-traffic-light` nur zurückgeben:

```
(define sugar-traffic-light
  (lambda (sugar-content)
    (cond
      ((rational? sugar-content)
       (sugar-weight->traffic-light sugar-content))
      ((sugars? sugar-content)
       (sugar-weight->traffic-light
        (+ (sugars-fructose-g sugar-content)
           (sugars-glucose-g sugar-content))))
      ((string? sugar-content) sugar-content))))
```

Fertig!

## 4.4 FEHLER REPRÄSENTIEREN

Dieser Abschnitt beschreibt eine weitere häufig vorkommende Anwendung für gemischte Daten. Es geht um Situationen, wo eine Funktion keine sinnvolle Antwort weiß.

Als Beispiel schreiben wir eine Funktion, welche die Steigung einer Geraden auf der Ebene berechnet anhand der Koordinaten zweier Punkte darauf. Die Funktion akzeptiert vier Argumente: Die X- und die Y-Koordinate des ersten Punkts und danach die X- und die Y-Koordinate des zweiten Punkts. Wir fangen mit Kurzbeschreibung, Signatur, einigen Tests und einem Gerüst an:

```
; Steigung einer Gerade berechnen
(: slope (real real real real -> real))
```

```
(check-expect (slope 0 0 1 1) 1)
(check-expect (slope 0 0 2 1) 1/2)
(check-expect (slope 1 2 3 5) 3/2)
```

```
(define slope
  (lambda (x1 y1 x2 y2)
    ...))
```

Die Formelsammlung sagt, dass wir zur Berechnung der Steigung die Differenz der Y-Koordinaten durch die Differenz der X-Koordinaten teilen müssen:

```
(define slope
  (lambda (x1 y1 x2 y2)
    (/ (- y2 y1)
       (- x2 x1))))
```

Fertig, könnte man meinen. Es gibt aber Linien, die keine Steigung haben. In der REPL sieht das so aus:

```
> (slope 0 0 0 1)
/: durch 0 geteilt
```

Die Meldung ist ein bisschen unfreundlich und birgt mehrere Probleme:

- Die Meldung besagt zwar, was das Programm gemacht hat (durch 0 geteilt), aber benennt nicht die Ursache des Problems: dass die Gerade keine Steigung hat.
- Das Programm bricht einfach ab, ohne dass es eine Möglichkeit bekommt, mit dem Problem umzugehen.
- Dass das Programm einfach abbrechen kann, ist aus der Signatur nicht ersichtlich, die suggeriert, dass die Funktion immer eine reelle Zahl zurückliefert.

Alle drei Probleme können wir auf einmal mit gemischten Daten lösen. Dazu müssen wir zunächst dafür sorgen, dass die Funktion einen sinnvollen Wert zurückliefert, bevor sie durch 0 teilen würde. Dadurch werden die Eingaben in zwei Klassen aufgeteilt: Es gibt eine Steigung oder eben nicht. Es bietet sich eine binäre Verzweigung an, um diese beiden Klassen zu unterscheiden:

```
(define slope
  (lambda (x1 y1 x2 y2)
```

```
(if (= (- x2 x1) 0)
    ...
    (/ (- y2 y1)
        (- x2 x1))))
```

Nur müssen wir anstelle der Ellipse irgendetwas ins Programm schreiben. Wir können nicht einfach irgendeine reelle Zahl nehmen, da diese ja auch eine legitime Steigung sein könnte. Stattdessen brauchen wir einen Wert, den wir von den legitimen Zahlen unterscheiden können. Wir könnten eine spezielle Zeichenkette oder `#f` benutzen. (Siehe dazu Aufgabe 4.10 auf Seite 130.)

Wir empfehlen aber, stattdessen einen eigenen Record-Typ zu definieren, und zwar so:

```
; Es gibt keine Steigung
(define-record no-slope
  make-no-slope
  no-slope?)
```

Dieser Record-Typ hat keine Felder, das heißt, er kann nur benutzt werden, um einen einzigen Wert herzustellen, nämlich `(make-no-slope)`. (Es handelt sich also um eine degenerierte Version von zusammengesetzten Daten ohne Bestandteile.) Diesen Wert können wir aber durch das Prädikat `no-slope?` von allen anderen Werten unterscheiden und benutzen ihn, wenn es keine Steigung gibt:

```
(define slope
  (lambda (x1 y1 x2 y2)
    (if (= (- x2 x1) 0)
        (make-no-slope)
        (/ (- y2 y1)
            (- x2 x1)))))
```

Jetzt können wir einen Testfall definieren:

```
(check-expect (no-slope? (slope 0 0 0 1)) #t)
```

Allerdings liefert DrRacket für diesen Testfall eine Signaturverletzung: Die Signatur tut ja immer noch so, als würde immer eine reelle Zahl herauskommen. Wir erweitern sie also folgendermaßen:

```
(: slope (real real real real -> (mixed real no-slope)))
```

Dieser Signatur sieht man unmittelbar an, was sie liefert und kann aus den Namen auch schließen, was die beiden Fälle bedeuten. Natürlich könnten wir eine eigene Signatur definieren wie sonst bei gemischten Daten, das lohnt aber hier kaum.



## AUFGABEN

### AUFGABE 4.5

Vereinfache in der Funktion `sugar-weight->traffic-light` auf Seite 123 die Bedingungen, wie in Abschnitt 2.5 auf Seite 60 beschrieben.

### AUFGABE 4.6

Ein Supermarkt möchte seine Waren in einem Programm verwalten. Es gibt drei Warenklassen:

**Essen** beschrieben durch einen Namen, den Stückpreis, das Mindesthaltbarkeitsdatum und den aktuellen Bestand im Supermarkt

**Getränke** beschrieben durch einen Namen, den Stückpreis, das Mindesthaltbarkeitsdatum und den Bestand. Zusätzlich muss hier noch festgehalten werden, ob Pfand verlangt wird.

**Sonstige** beschrieben durch einen Namen, den Stückpreis und den Bestand.

1. Führe eine Datenanalyse durch und erstelle Daten- und Record-Definitionen.
2. Schreibe eine Funktion `stückpreis`, die eine Warenklasse akzeptiert und den Stückpreis zurückgibt.
3. Schreibe eine Funktion `buchen`, die eine Warenklasse und die Anzahl der abzubuchenden Exemplare akzeptiert, den Bestand der Warenklasse reduziert und die Warenklasse zurückgibt. Falls mehr Exemplare gefordert werden, als in der Warenklasse vorhanden sind, soll der Bestand auf 0 gesetzt werden.
4. Schreibe eine Funktion `haltbar?`, die eine Warenklasse und ein Datum akzeptiert und `#t` zurückgibt, falls das Mindesthaltbarkeitsdatum (MHD) nicht überschritten wurde. Falls kein MHD bekannt ist, soll `#t` zurückgegeben werden.

### AUFGABE 4.7

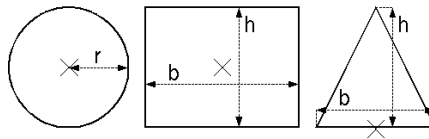
Erweitere die Lösung von Aufgabe 3.16 aus dem vorigen Kapitel (Seite 107):

1. Schreibe nun eine Funktion, die zwei Grundkreaturen in beliebiger Reihenfolge akzeptiert, die Kreuzung vornimmt und eine neue Kreatur zurückgibt.
2. Schreibe nun eine Funktion, die drei Grundkreaturen in beliebiger Reihenfolge akzeptiert, die Kreuzung vornimmt und eine neue Kreatur zurückgibt.

### AUFGABE 4.8

Das Spiele-Entwicklungsteam von *Exciting Games, Inc.* ist in Personalnot: Für die Implementierung des bahnbrechenden 2D-Spiels *The Adventures of DrRacket* wird dringend eine Repräsentation der grafischen Formen Kreis, Rechteck und Dreieck benötigt. Alle Hacker sind leider wahnsinnig damit beschäftigt, bunte Pixel in der Gegend herumzuschieben und können sich daher nicht dieser grundlegenden Aufgabe widmen. Deswegen fällt Dir diese Aufgabe zu! Dir geht nur ein mit Pizzabelag verschmierter Zettel zu, der offenbar als Arbeitsanweisung gedacht ist. Darauf steht:

- Es gibt drei Klassen von Formen: Kreise, Rechtecke und gleichschenklige Dreiecke. Alle Formen haben eine Ursprungscoordinate und enthalten eine Information über die Größe (Radius, Höhe, Breite, etc), wie Du auf der Abbildung siehst. Benutze für die Ursprungscoordinate die vorgegebenen kartesischen Koordinaten!



- Es soll Funktionen geben, die auf allen Formen funktionieren und Folgendes leisten:
  - kartesische Koordinate einer Form zurückgeben
  - $x$ -Koordinate einer Form zurückgeben
  - $y$ -Koordinate einer Form zurückgeben
  - Abstand des Ursprungs einer Form zum Ursprung des Koordinatensystems zurückgeben
  - Flächeninhalt einer Form zurückgeben
  - Form um  $\Delta_x$  in  $x$ -Richtung und um  $\Delta_y$  in  $y$ -Richtung verschieben, also eine Form zurückgeben, deren Ursprungscoordinate entsprechend verschoben ist
  - Flächeninhalt zweier Formen vergleichen: Für gleich große Formen liefert die Funktion "equal" zurück, sonst "smaller", wenn die erste Form kleiner ist, beziehungsweise "bigger", wenn die erste Form größer ist.

Zeige den Hackern, was richtiges Software-Engineering ist und benutze die passenden Konstruktionsanleitungen, um Repräsentationen für die Formen und die zugehörigen Funktionen zu schreiben!

#### AUFGABE 4.9

Es gibt verschiedene Verstöße gegen die Straßenverkehrsordnung:

- Falschparken mit Ort und Zeitpunkt des Verstoßes
- Überfahren einer roten Ampel mit Ort und Zeitpunkt des Verstoßes sowie der Dauer in Sekunden, wie lange die Ampel bereits rot war

- Überhöhte Geschwindigkeit mit Ort und Zeitpunkt des Verstoßes sowie der Höhe der Geschwindigkeitsübertretung

Neben den angegebenen Bestandteilen muss später für jeden Verstoß vermerkt werden, ob zusätzlich eine «Gefährdung des Straßenverkehrs» vorliegt.

1. Schreibe eine Datendefinition für Verstöße gegen die Straßenverkehrsordnung. Schreibe Daten- und Record-Definitionen für die verschiedenen Verstöße. Die Zeitpunkte kannst Du durch Zeichenketten wie "11.4.1971 22:00" repräsentieren. Gib alle Signaturen an!
2. Schreibe eine Funktion, die einen beliebigen Verstoß entgegennimmt und den Ort des Verstoßes zurückgibt.  
Schreibe analog eine Funktion, die einen beliebigen Verstoß entgegennimmt und den Zeitpunkt des Verstoßes zurückgibt.
3. Schreibe eine Funktion, die einen beliebigen Verstoß entgegennimmt, diesen als Gefährdung des Straßenverkehrs betrachtet und wieder zurückgibt (das heißt jeder Verstoß ist ein Verstoß mit Gefährdung).

Die Verstöße haben unterschiedliche Tatbestände zur Folge:

- Einfaches Vergehen mit Bußgeld
  - Ordnungswidrigkeit mit Bußgeld, Punkte für die Verkehrssünderdatei und Fahrverbot in Monaten
  - Straftat mit Punkten für die Verkehrssünderdatei und Freiheitsstrafe in Monaten
1. Schreibe eine Datendefinition für Tatbestände auf Verstöße gegen die Straßenverkehrsordnung. Schreibe Daten- und Record-Definitionen für die verschiedenen Tatbestände. Gib alle Signaturen an!
  2. Schreibe eine Funktion, die einen beliebigen Tatbestand akzeptiert und die Anzahl der anfallenden Punkte zurückliefert.  
Schreibe nun Funktionen, die Verstöße akzeptieren und die Tatbestände berechnen:
    1. Schreibe eine Funktion, die Falschparken akzeptiert und ein einfaches Vergehen mit 20 Euro Bußgeld zurückgibt. Wenn das Falschparken eine Gefährdung des Straßenverkehrs darstellt (z.B. bei Behinderung von Rettungsfahrzeugen), soll die Funktion eine Ordnungswidrigkeit mit 40 Euro Bußgeld, einem Punkt und keinem Fahrverbot zurückgeben.
    2. Schreibe eine Funktion, die Überfahren einer roten Ampel akzeptiert und eine Ordnungswidrigkeit zurückgibt. Dabei gilt:
      - wenn die Ampel kürzer als eine Sekunde rot war und keine Gefährdung vorlag: 50 Euro, 3 Punkte, kein Fahrverbot
      - wenn die Ampel mindestens eine Sekunde rot war oder eine Gefährdung vorlag: 125 Euro, 4 Punkte, 1 Monat Fahrverbot

3. Schreibe eine Funktion, die überhöhte Geschwindigkeit akzeptiert und den Tatbestand zurückgibt. Dabei gilt:
  - Geschwindigkeitsübertretung weniger als 20 km/h ohne Gefährdung: einfaches Vergehen mit 35 Euro
  - Geschwindigkeitsübertretung zwischen 20 und 40 km/h (inklusive) ohne Gefährdung: Ordnungswidrigkeit mit 75 Euro, 3 Punkten und 1 Monat Fahrverbot
  - Geschwindigkeitsübertretung von mehr als 40 km/h ohne Gefährdung: Ordnungswidrigkeit mit 200 Euro, 4 Punkte und 3 Monate Fahrverbot
  - bei gleichzeitiger Gefährdung des Straßenverkehrs: Straftat mit 3 Punkte mehr als ohne Gefährdung angegeben und Freiheitsstrafe, die doppelt so lang ist wie das Fahrverbot, das ohne Gefährdung gilt (wenn es kein Fahrverbot gibt, dann gibt es auch keine Freiheitsstrafe)
4. Schreibe eine Funktion, die einen Verstoß akzeptiert und dessen Folge zurückgibt. Benutze die Funktionen aus den vorherigen Teilaufgaben.

#### AUFGABE 4.10

Experimentiere für die Funktion `slope` in Abschnitt 4.4 auf Seite 124 mit anderen Repräsentationen als einem eigenen Record-Typ für «keine Steigung»:

- Liefere `#f` zurück.
- Liefere eine Zeichenkette `"no slope"` zurück.

Wie muss jeweils die Signatur von `slope` lauten?

Wie könnte ein Programm dann nach einem Aufruf feststellen, ob es keine Steigung gab? Schreibe für beide Varianten einen Ersatz für das Prädikat `no-slope?` diese Situation feststellt.

Was haben die insgesamt drei Varianten für die Repräsentation von «keine Steigung» jeweils für Vor- und Nachteile?

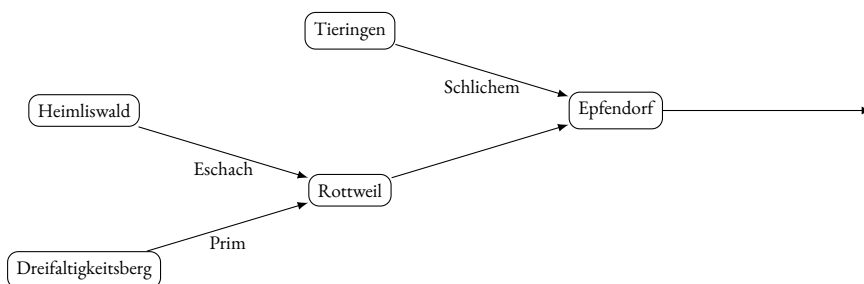
## 5 PROGRAMMIEREN MIT SELBSTBEZÜGEN UND KOMBINATOREN

Bisher haben wir nur Daten vorgestellt, die feste Größe haben und damit im Wortsinn beschränkt sind. Viele Informationen haben aber variable Größe: Die Bücher im Regal werden immer mehr, Bauwerke bestehen aus variabel vielen Bauteilen. Um solche Informationen als Daten zu repräsentieren, stellen wir in diesem Kapitel einen weiteren Aspekt der Datenanalyse vor, den *Selbstbezug*. Hier sind einfache Beispiele:

- Ein Fluss besteht aus Hauptfluss und Nebenfluss – beides wieder Flüsse.
- Ein Dateiverzeichnis auf dem Computer kann Unterverzeichnisse enthalten.
- Ein großer Bücherstapel besteht aus einem etwas kleineren Bücherstapel und einem weiteren Buch.

Bei all diesen Beispielen werden aus kleineren Dingen größere gemacht: Aus zwei Flüssen, die zusammenfließen, wird ein großer Fluss. Aus mehreren Dateien und Verzeichnissen wird ein großes Verzeichnis. Aus einem kleinen Bücherstapel wird ein großer. In der Programmierung nennen wir das Bauen von großen Dingen aus kleinen Dingen derselben Sorte *kombinieren*. Die Funktionen, die das erledigen, heißen *Kombinatoren*: Um die geht es in diesem Kapitel.

### 5.1 FLÜSSE ABBILDEN



**Abbildung 5.1:** Der Neckar nahe den Quellen

Abbildung 5.1 ist ein Diagramm, das die Struktur des Neckar nahe den Quellen beschreibt: Er fließt zunächst aus den Bächen Eschach und Prim zusammen, und danach fließen immer weitere Bäche und Flüsse hinzu – in der Abbildung noch die Schlichem.

Die Struktur eines Flusses werden wir durch Daten abbilden, um danach eine Funktion zu

schreiben, die feststellt, ob ein Fluss durch einen bestimmten Ort fließt. Unser erster Versuch einer Datendefinition sieht so aus:

```
; Ein Fluss kommt entweder aus:
; - einer Quelle
; - einem Hauptfluss und einem Nebenfluss
```

Man kann zwar schon sehen, dass es sich um eine Fallunterscheidung handelt, aber wir können die Formulierung schärfen, um klar zu machen, dass es sich um gemischte Daten handelt:

```
; Ein Fluss ist eins der folgenden:
; - ein Bach aus einer Quelle
; - ein Zusammentreffen von einem Hauptfluss und einem Nebenfluss
```

Daraus können wir direkt eine Signatur-Definition machen:

```
(define river
  (signature (mixed creek confluence)))
```

Die Datendefinitionen für «Bach» und «Zusammentreffen» fehlen allerdings noch. In Abbildung 5.1 steht an den Bächen jeweils noch der Name. Eine passende Datendefinition sieht so aus:

```
; Ein Bach hat folgende Eigenschaften:
; - Ursprungsort
```

Das klingt ein bisschen komisch, weil der Bach ja nur *eine* Eigenschaft hat, aber die Datendefinition trotzdem zusammengesetzte Daten beschreibt. Trotzdem ist das legitim und korrekt – niemand wird sagen wollen, dass der Ursprungsort der Bach *ist*. Außerdem kannst Du Dir vorstellen, dass ein Bach später noch mehr Eigenschaften bekommt, die wir in den Daten festhalten wollen. (Wasserqualität oder Fließgeschwindigkeit zum Beispiel.) Entsprechend ist es auch sinnvoll, dafür eine Record-Definition zu schreiben:

```
(define-record creek
  make-creek
  creek?
  (creek-origin string))
```

Wenn später noch weitere Eigenschaften hinzukommen, können wir die Record-Definition erweitern, ohne anderen Code zu beeinträchtigen.

Kommen wir zu den Zusammentreffen. Auch hier ist ein Ort relevant – dort, wo sie zusammentreffen. Außerdem ist wichtig, *welche* Flüsse oder Bäche da zusammentreffen. Die Datendefinition könnte so aussehen:

```
; Ein Zusammentreffen besteht aus:
; - Ort
; - Hauptfluss
; - Nebenfluss
```

Die Formulierung zeigt eindeutig, dass es sich um zusammengesetzte Daten handelt. Aber diese Datendefinition weist eine Auffälligkeit auf, die erst sichtbar wird, wenn wir sie im Zusammenhang der Gesamt-Datendefinition für Flüsse betrachten:

```
; Ein Fluss ist eins der folgenden:
; - ...
; - ein Zusammentreffen ...
...
; Ein Zusammentreffen besteht aus:
; - Ort
; - Hauptfluss
; - Nebenfluss
```

Die Datendefinition für «Fluss» benutzt selbst den Begriff «Fluss». Noch offensichtlicher wird das, wenn wir die dazu passende Record-Definition erstellen:

```
(define-record confluence
  make-confluence
  confluence?
  (confluence-location string)
  (confluence-main-stem river)
  (confluence-tributary river))
```

In der Definition von `confluence` wird `river` benutzt und umgekehrt. Das ist nichts schlimmes – im Gegenteil, diese beiden *Selbstbezüge* erlauben uns, ganz unterschiedliche Flüsse zu beschreiben, insbesondere solche unterschiedlicher Größe. Den Abschnitt des Neckar in 5.1 können wir so abbilden:

```
(define eschach (make-creek "Heimliswald")) ; Quelle des Neckar
(define prim (make-creek "Dreifaltigkeitsberg")) ; Quelle des Neckar
; erster Zusammenfluss des Neckar:
(define neckar-1 (make-confluence "Rottweil" eschach prim))
; Zufluss des Neckar:
(define schlichem (make-creek "Tieringen"))
```

; zweiter Zusammenfluss des Neckar:

```
(define neckar-2 (make-confluence "Epfendorf" neckar-1 schlichem))
```

An diesen Beispielen sieht man sehr schön, dass `make-confluence` zwei Flüsse zu einem kombiniert – es handelt sich um einen *Kombinator*.

Wir hatten angekündigt, eine Funktion zu schreiben, die feststellt, ob Wasser von einem bestimmten Ort in einen Fluss fließt. Wir machen das wieder nach Vorschrift, also zuerst die Kurzbeschreibung:

; Fließt Wasser von einem Ort in Fluss?

In der Kurzbeschreibung sind schon die Substantive «Fluss» und «Ort» als Eingaben aufgeführt, die übertragen wir in die Signatur. Da die Funktion eine Ja-/Nein-Frage beantwortet, liefert sie einen booleschen Wert und hat ein Fragezeichen hinten am Namen:

```
(: flows-from? (string river -> boolean))
```

Bei den Tests achten wir darauf, dass wir sowohl Bäche als auch Flüsse testen, und sowohl Fälle, bei denen `#t` als auch solche, bei denen `#f` herauskommt:

```
(check-expect (flows-from? "Heimliswald" eschach) #t)
(check-expect (flows-from? "Tübingen" eschach) #f)
(check-expect (flows-from? "Heimliswald" neckar-2) #t)
(check-expect (flows-from? "Rottweil" neckar-2) #t)
(check-expect (flows-from? "Berlin" neckar-2) #f)
```

Das Gerüst ergibt sich wie immer direkt aus der Signatur:

```
(define flows-from?
  (lambda (location river)
    ...))
```

Da es sich bei `river` um gemischte Daten handelt, können wir die Schablone dafür ergänzen. Da `river` zwei Fälle hat (Bäche und Zusammenflüsse), muss die Schablone zwei Zweige haben. Die Bedingungen sind Aufrufe der jeweiligen Prädikate für `creek` und `confluence`:

```
(define flows-from?
  (lambda (location river)
    (cond
      ((creek? river) ...)
      ((confluence? river) ...))))
```



Nun können wir Code für die beiden Zweige ergänzen. Fangen wir mit den Bächen an. Da es sich bei `creek` um zusammengesetzte Daten handelt (mit nur einem Bestandteil), sollte der Aufruf des Selektors im Rumpf vorkommen:

```
(define flows-from?
  (lambda (location river)
    (cond
      ((creek? river)
       ... (creek-origin river) ...)
      ((confluence? river)
       ...))))
```

Der Selektor `creek-origin` hat einen Ursprungsort. Wenn dieser dem gesuchten Ort entspricht, so fließt der Fluss durch diesen Ort, sonst nicht. Im ersten Fall sollte `#t` herauskommen, im zweiten `#f`. Das sieht so aus:

```
(define flows-from?
  (lambda (location river)
    (cond
      ((creek? river)
       (if (string=? (creek-origin river) location)
           #t
           #f))
      ((confluence? river)
       ...))))
```

Dieser Code kann noch etwas vereinfacht werden: Der `if`-Ausdruck sagt ja salopp formuliert:

Wenn `(string=? (creek-origin river) location)` als Resultat `#t` liefert,  
dann `#t`, und wenn es `#f` liefert, dann `#f`.

Da reicht es auch, `(string=? (creek-origin river) location)` hinzuschreiben.

#### AUFGABE 5.1

Kannst Du diese Vereinfachung auf `if`-Ausdrücken verallgemeinern und als Gleichung schreiben?



Als Nächstes können wir uns um den anderen Fall kümmern, `confluence`. Auch dort handelt es sich um zusammengesetzte Daten, wir können also schon einmal die Selektoraufrufe hinschreiben:

```
(define flows-from?
  (lambda (location river)
    (cond
      ((creek? river)
       (string=? (creek-origin river) location))
      ((confluence? river)
       ...
       (confluence-location river)
       (confluence-main-stem river)
       (confluence-tributary river)
       ...))))
```

Soweit die Schablone, wir können uns also Gedanken zur eigentlichen Aufgabe machen.

Als erstes steht da `(confluence-main-stem river)` der Ort des Zusammenflusses. Wenn es sich dabei um den gesuchten Ort handelt, können wir die Frage, ob der Wasser von diesem Ort in den Fluss fließt, bereits mit «ja» beziehungsweise `#t` beantworten:

```
(define flows-from?
  (lambda (location river)
    (cond
      ((creek? river)
       (string=? (creek-origin river) location))
      ((confluence? river)
       (if (string=? (confluence-location river) location)
           #t
           ...
           (confluence-main-stem river)
           (confluence-tributary river)
           ...))))
```

Aber was wenn nicht? Dann müssen wir uns an die beiden anderen Selektor-Aufrufe halten, die uns den Haupt- und den Nebenfluss des Zusammenflusses liefern. Wenn der eine oder der andere Wasser aus dem gesuchten Ort enthält, dann könnten wir die Frage unserer Funktion ebenfalls mit «ja» beantworten. Dazu müssten wir also feststellen:

1. Fließt Wasser vom Ort `origin` in den Hauptfluss?
2. Fließt Wasser vom Ort `origin` in den Nebenfluss?

Nun, wir schreiben ja gerade eine Funktion, die feststellt, ob Wasser von einem bestimmten Ort in einen bestimmten Fluss fließt. Können wir die benutzen? Wir schreiben das mal hin:

```

(define flows-from?
  (lambda (location river)
    (cond
      ((creek? river)
       (string=? (creek-origin river) location))
      ((confluence? river)
       (if (string=? (confluence-location river) location)
           #t
           ...
           (flows-from? location (confluence-main-stem river))
           (flows-from? location (confluence-tributary river))
           ...))))))

```

Die beiden Aufrufe von `flows-from?` entsprechen gerade den beiden Fragen von oben. Wir können ihre Ergebnisse kombinieren, um unsere Antwort zu errechnen: Wenn Wasser aus `location` durch den Hauptfluss fließt *oder* der Wasser aus `location` in den Nebenfluss fließt, so fließt auch Wasser von `location` in den «Gesamtfluss». So sieht das aus:

```

(define flows-from?
  (lambda (location river)
    (cond
      ((creek? river)
       (string=? (creek-origin river) location))
      ((confluence? river)
       (if (string=? (confluence-location river) location)
           #t
           (or
            (flows-from? location (confluence-main-stem river))
            (flows-from? location (confluence-tributary river)))))))

```

Fertig!

In dieser Funktion passiert etwas besonderes, das es in den Funktionen davor noch nicht gab: Sie enthält einen Aufruf von sich selbst, einen sogenannten *rekursiven Aufruf*. Diese rekursiven Aufrufe sind genau an den Stellen, wo die Datendefinition von `river` Selbstbezüge enthält.

Das ist kein Zufall: Nahezu alle Funktionen, die Daten mit Selbstbezügen akzeptieren, enthalten an den Stellen der Selbstbezüge rekursive Aufrufe. Daraus machen wir eine Schablone:

### KONSTRUKTIONSANLEITUNG 17 (SELBSTBEZÜGE ALS EINGABE: SCHABLONE)

Wenn Du eine Funktion schreibst, die Daten akzeptiert, in denen Selbstbezüge enthalten sind, dann schreibe an die Stellen der Selbstbezüge jeweils einen rekursiven Aufruf.

Ein Nachtrag noch: Der `if`-Ausdruck in der `flows-from?`-Funktion passt *fast* auf das Muster aus Aufgabe 5.1. Der einzige Unterschied ist, dass in der Alternative der Verzweigung nicht `#f` steht. Allgemein wir ein solcher `if`-Ausdruck folgendermaßen ausgewertet:

```
(if b #t a)
↪ #t ; falls b = #t
↪ a  ; falls b = #f
```

Das können wir mit dem Verhalten von `or` vergleichen:

```
(or b a)
↪ #t ; falls b ↪ #t
↪ a  ; falls b ↪ #f
```

Wir können den `if`-Ausdruck aus `flows-from?` also durch ein `or` ersetzen:

```
(or (string=? (confluence-location river) location)
    (or
     (flows-from? location (confluence-main-stem river))
     (flows-from? location (confluence-tributary river))))
```

Das können wir sogar noch weiter vereinfachen, weil `or` auch mit drei Operanden funktioniert:

```
(or (string=? (confluence-location river) location)
    (flows-from? location (confluence-main-stem river))
    (flows-from? location (confluence-tributary river)))
```

Wenn wir das Resultat auf seine Bedeutung hin lesen, steht da folgendes:

Wasser fließt aus `location` in den Zusammenfluss `river`, wenn

- der Zusammenfluss gerade bei `location` stattfindet

**oder**

- Wasser von `location` in den Hauptfluss fließt

**oder**

- Wasser von `location` in den Nebenfluss fließt.

So erzählt ist die Funktionsweise von `flows-from?` (hoffentlich) direkt verständlich.

Wir können so aus diesem Beispiel und Aufgabe 5.1 zwei Gleichungen entsprechend Mantra 6 auf Seite 37 machen:

```
(if b #t #f) = b
(if b #t a)  = (or b a)
```

## 5.2 BILDER MODELLIEREN

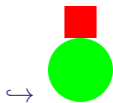
Erinnerst Du Dich noch an Abschnitt 1.4.2 auf Seite 15? Da haben wir mit Hilfe des `image.rkt`-Teachpacks die Funktionen `beside` und `above` verwendet, um Bilder zusammenzusetzen:

```
(define s1 (square 40 "solid" "red"))
(define c1 (circle 40 "solid" "green"))
(define p1 (star-polygon 20 10 3 "solid" "blue"))
(beside s1 p1)
```



↪

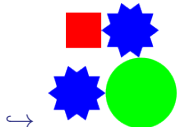
```
(above s1 c1)
```



↪

Bei beiden Funktionen war es möglich, die Ergebnisse wieder als Bilder zu verwenden:

```
(above (beside s1 p1) (beside p1 c1))
```



↪

Für Bilder ist die Signatur `image` zuständig; mit ihrer Hilfe können wir Signaturen für `above` und `beside` formulieren:

```
(: beside (image image -> image))
(: above (image image -> image))
```

An den Signaturen lässt sich klar erkennen, dass es sich um Kombinatoren handelt. Indirekt können wir daraus schließen, dass es in der Datendefinition von Bildern ebenfalls Selbstbezüge gibt. Die könnte so aussehen:

```
; Ein Bild ist eins der folgenden:
; ...
; - eine Nebeneinanderstellung von einem Bild und noch einem Bild
; - eine Übereinanderstellung von einem Bild und noch einem Bild
; ...
```

Das heißt, bei den Bildern ist das gleiche Konstruktionsprinzip am Werk wie bei den Flüssen: Aus «kleinen» Bildern werden «größere», und daraus gegebenenfalls noch größere undsoweiter.

An den Signaturen von `beside` und `above` ist leicht erkennbar, dass es sich um Kombinatoren handelt: Es gehen `image` rein und `image` kommt auch wieder heraus. Um das bei den Flüssen aus dem letzten Abschnitt klar anzuzeigen, wäre es sinnvoll, für `make-confluence` ebenfalls eine explizite Signatur anzugeben:

```
(: make-confluence (string river river -> river))
```

Kombinatoren verstecken sich in vielen Problemstellungen, Du musst Sie nur suchen: Dann findest Du erstaunlich oft auch welche. Darum geht es in Mantra 8:

#### MANTRA 8

Suche nach Kombinatoren für Deine Daten, die aus kleinen Dingen größere Dinge machen – und daraus wieder größere Dinge machen.

## 5.3 FINANZVERTRÄGE ABBILDEN

Es ist Zeit für ein etwas größeres Beispiel. Wir bilden Finanzverträge ab und programmieren dazu eine stark vereinfachte Version eines Klassikers der Informatik-Forschung, das zu mehreren Veröffentlichungen und zur Gründung einer erfolgreichen Firma geführt hat.<sup>1</sup> Selbstbezüge spielen eine wichtige Rolle.

<sup>1</sup> Wer mehr wissen möchte, sollte das Paper von Peyton Jones und Eber [PEYTON JONES and EBER, 2003] lesen, das es auch im Internet kostenlos zum Download gibt.

Wichtig bei diesem Beispiel: Für die Repräsentation von Finanzverträgen aus diesem Abschnitt haben renommierte Informatik-Forscher mehrere Monate gebraucht. Wir können sie mit den bisher präsentierten Programmiermitteln nachbauen, aber Du solltest nicht von Dir erwarten, dass Du das Beispiel in wenigen Stunden auch selbst hättest entwickeln können.

In der Finanzbranche werden oft Verträge geschlossen, die zur strukturierten Zahlung von Geldbeträgen führen. Wer eine Bankerin fragt, was der einfachste solche Vertrag ist, bekommt oft die Antwort *Zero-Bond*. Hier ist ein Beispiel für einen Zero-Bond:

Ich bekomme am 31. Juni 2030 1000 €.

So ein Vertrag wird immer zwischen zwei Vertragspartnern geschlossen und der eine ist immer «ich». Daher kommt die Formulierung «Ich bekomme». Der Begriff «Zero-Bond» setzt sich zusammen aus «keine Zinsen» (Zero) und «Recht auf spätere Zahlung» (Bond).

Wir können daraus eine einfache Datendefinition machen:

```
; Ein Zero-Bond hat folgende Eigenschaften:  
; - Datum  
; - Betrag  
; - Währung
```

An der Formulierung «hat folgende Eigenschaften» sehen wir sofort, dass es sich um zusammengesetzte Daten handelt. Daraus folgt direkt diese Record-Definition:

```
(define-record zero-bond  
  make-zero-bond  
  (zero-bond-date      date)  
  (zero-bond-amount    rational)  
  (zero-bond-currency  currency))
```

Wir müssten noch Definitionen für `date` und `currency` beisteuern, schieben das aber auf die lange Bank. Wir wollen stattdessen zunächst noch einmal grundsätzlich über die richtige Datendefinition für solche Verträge nachdenken.

Es gibt ja noch viel mehr solche Verträge: Futures, Forwards, Swaps, Swaptions ... Jeder Vertrag hat mehrere Eigenschaften, das würde also darauf hinauslaufen, dass wir sehr viele Datendefinitionen für zusammengesetzte Daten und die zugehörigen Record-Definitionen schreiben müssten – und am Ende eine Datendefinition zu «Vertrag», die auf eine große Signatur für gemischte Daten hinausläuft. Das ist eine Menge Arbeit, und sie wäre auch nie fertig, weil sich die Banker ständig neue Verträge ausdenken.

Wer aber genug solche Verträge studiert, bemerkt, dass diese aus Bausteinen bestehen, die immer wiederkehren. Das suggeriert einen anderen Ansatz, nämlich diese Bausteine zu identifizieren und dafür die Datenanalyse durchzuführen. Wir vollziehen stattdessen die Arbeit der Autoren des Papers [PEYTON JONES and EBER, 2003] nach. Die haben (über mehrere Monate) einen Satz bestechend einfacher Bausteine zusammengestellt, indem sie bekannte Verträge in möglichst kleine und einfache Bestandteile zerlegten. Jeder dieser Bestandteile ist für sich genommen ein «kleiner Vertrag», es gibt aber auch Kombinatoren, die aus kleinen Verträgen größere machen, insbesondere auch noch unbekannte Klassen von Verträgen. Mit diesen Bestandteilen lassen sich durch den Einsatz von Kombinatoren sehr viele Finanzverträge (auch noch unbekannte) abbilden.

Da es mehrere Klassen von Bausteinen und Kombinatoren gibt, werden wir Verträge als gemischte Daten repräsentieren und dafür schrittweise eine Datendefinition folgender Form aufbauen:

```
; Ein Vertrag ist eins der folgenden:
; - ...
```

Dazu gehört natürlich eine passende Signatur-Definition, die wir ebenfalls schrittweise aufbauen werden:

```
(define contract
  (signature (mixed ...)))
```

Um die Bausteine und Kombinatoren zu identifizieren, schauen wir uns zunächst den Zero-Bond näher an. Auch wenn der einfach aussieht, gibt es mehrere Aspekte, die auch unabhängig voneinander Sinn ergeben: die Auszahlung eines bestimmten *Betrags* in einer bestimmten *Währung* und die Möglichkeit, eine Zahlung *später* zu veranlassen. Das sind drei separate Ideen, und es ist sinnvoll, diese getrennt als Daten zu repräsentieren.

**Zahlung einer bestimmten Währung** | Fangen wir an mit der Zahlung einer bestimmten Währung. Der Einfachheit halber nehmen wir an, dass alle Zahlungen in Euro sind. Um die richtige Daten-Definition zu finden, versuchen wir, die *einfachste* Euro-Zahlung zu finden, und die besteht aus *einem* Euro. Die Daten-Definition dazu ist ernüchternd einfach:

```
; Ein Euro hat keine Eigenschaften
```

Wir könnten den Euro einfach als die Zeichenkette "EUR" abbilden. Das wäre sinnvoll, wenn der Euro Teil einer Aufzählung ist. Dies ist allerdings hier nicht der Fall, andere Vertragsbausteine haben durchaus Eigenschaften. (Siehe dazu auch Aufgabe 5.13 auf Seite 163.) Wir nehmen die Datendefinition stattdessen beim Wort «hat ... Eigenschaften» und bilden den Euro als zusammengesetzte Daten ab. Dann wird daraus folgende Record-Definition:



```
(define-record one-euro
  make-one-euro
  one-euro?)
```

Wir fügen als Nächstes `one-euro` zur Daten- und zur Signatur-Definition von `contract` hinzu:

```
; Ein Vertrag ist eins der folgenden:
; - ein Euro
; - ...
(define contract
  (signature (mixed one-euro ...)))
```

**Beliebige Beträge** | Als Nächstes bilden wir die Möglichkeit ab, einen Betrag zu wählen. Eine naheliegende erste Idee wäre, die Idee «*n* Euros» abzubilden, etwa so:

```
; Euros haben folgende Eigenschaft:
; - wieviele
(define-record euros
  make-euros
  euros?
  (euros-amount rational))
```

Wenn wir das hätten, bräuchten wir aber `one-euro` nicht mehr. Außerdem wäre das Resultat dann nicht mehr so einfach wie möglich: Besser wäre es, wenn wir die Idee «wieviele» trennen könnten von «Euro». Dann könnten wir sagen «mehrere von einem Euro». Das klingt erstmal komisch, auch in der Datendefinition:

```
; Mehrere haben folgende Eigenschaften:
; - wieviele
; - wovon
```

Aber die richtige Form für zusammengesetzte Daten hat das schon einmal. Es reicht für eine erste Skizze einer Record-Definition:

```
(define-record multiple
  make-multiple
  multiple?
  (multiple-number ...)
  (multiple-of      ...))
```

Wir müssen aber noch klären, was die Signatur von `multiple-number` respektive `multiple-of` ist. Bei `multiple-number` müssen wir uns nicht auf ganze Zahlen beschränken – ein halber Euro geht schließlich auch. Da ist `rational` sinnvoll.

Bei `multiple-of` liegt es nahe, `one-euro` hinzuschreiben: Dann könnten wir den Zero-Bond schon hinschreiben. Allerdings wäre `multiple` ziemlich eingeschränkt und nicht besonders zukunftssicher: Was, wenn andere Währungen dazukommen, `one-gbp` oder so? Vielleicht sind wir versucht, eine Signatur `currency` für Währungen zu definieren und für `multiple-of` zu verwenden. Aber das wäre immer noch zu eingeschränkt: Man kann ja auch mehrere Aktien (zum Beispiel) im Rahmen eines Vertrags übergeben – oder auch mehrere Zero-Bonds oder andere Verträge.

Und da ist sie ganz plötzlich, die zentrale Idee: Dass ein Vertrag den Abschluss *anderer* Verträge verfügen kann. Wir schreiben deshalb als Signatur von `multiple-of` einfach `contract` und schärfen bei der Gelegenheit die Datendefinition:

```
; Ein Vielfaches besteht aus:
; - Anzahl
; - Vertrag
(define-record multiple
  make-multiple
  multiple?
  (multiple-number rational)
  (multiple-of      contract))
```

Nun gibt es schon zwei Klassen Verträge – `one-euro` und `multiple`. Wir erweitern die Definition von `contract` entsprechend:

```
(define contract
  (signature (mixed one-euro multiple)))
```

Mit Hilfe von `one-euro` und `multiple` können wir einen Vertrag definieren, der festlegt, dass wir 100 € bekommen – sofort:

```
(define euro100 (make-multiple 100 (make-one-euro))) ; 100 Euros
```

## AUFGABE 5.2

Definiere eine Funktion `make-euros` mit Kurzbeschreibung, Signatur und Test wie folgt:

```
; Euro-Betrag auszahlen
(: make-euros (rational -> contract))

(check-expect (make-euros 100) euro100)
```



Mit Hilfe der Funktion aus der Aufgabe können wir leicht auch 200 € definieren:

```
(define euro200 (make-euros 200)) ; 200 Euros
```

**Verzögerungen** | Für die Zero-Bonds benötigen wir noch die Möglichkeit, eine Zahlung zu verzögern. Dazu benutzen wir den gleichen Trick wie bei `multiple` und definieren einen weiteren Kombinator, der einen ganzen Vertrag verzögert: Wir schließen also einen Vertrag ab, später einen Vertrag abzuschließen, und der kann dann die Zahlung auslösen. Um so einen Vertrag zu beschreiben, benötigen wir das Datum, an dem der «spätere» Vertrag aktiv wird und den späteren Vertrag selbst:

```
; Eine Verzögerung besteht aus:
; - Datum
; - Vertrag, der zu dem Datum gültig wird
```

Diese Datendefinition beschreibt wieder zusammengesetzte Daten. Die dazugehörige Record-Definition sieht so aus:

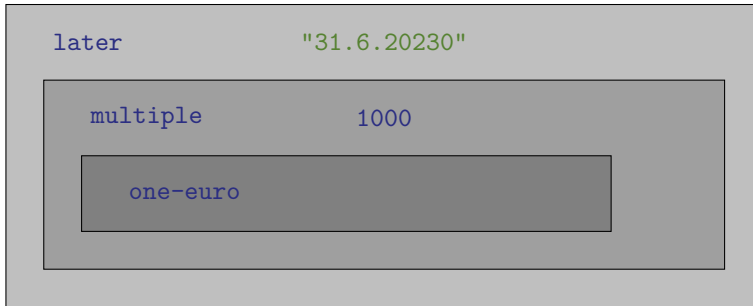
```
(define-record later
  make-later
  later?
  (later-date    date)
  (later-contract contract))
```

Mit `later` können wir die Signatur-Definition von `contract` erweitern:

```
(define contract
  (signature (mixed one-euro multiple later)))
```

Aber es fehlt noch etwas: Wir haben für `later-date` die Signatur `date` verwendet, die noch gar nicht definiert ist. Die Definition von `date` liegt zum Glück recht nahe:

```
(define-record date
  make-date
  date?
  (date-year  natural)
  (date-month natural)
  (date-day   natural))
```



**Abbildung 5.2:** Zero-Bond als Schichten dargestellt

### AUFGABE 5.3

Wir haben für alle Felder von `date` die Signatur `natural` verwendet. Diese Signatur ist sehr unpräzise, zumindest für `date-month` und `date-day`.

Verbessere die Signaturen von `date-month` und `date-day` mit Hilfe von `lstinlineinteger-from-to` aus Abschnitt 2.2 auf Seite 47! □

Bevor es weitergeht, definieren wir noch zwei Beispiele für das Datum:

```
(define date1 (make-date 2019 07 26)) ; 26. Juli 2019
(define date2 (make-date 2019 12 24)) ; Weihnachten 2019
```

Mit Hilfe dieser Definitionen können wir auch endlich zwei Beispiele für `later` angeben:

```
(define later1
  (make-later date1 euro100)) ; Ich bekomme am 26.7.2019 100 €.
(define later2
  (make-later date2 euro200)) ; Ich bekomme Weihnachten 2019 200 €.
```

Außerdem können wir den Zero-Bond vom Anfang dieses Abschnitts definieren:

```
; Ich bekomme am 31. Juni 2030 1000 €.
(define zero1 (make-later (make-date 2030 06 31)
                          (make-euros 1000)))
```

Abbildung 5.2 zeigt den Aufbau des Zero-Bonds `zero1` als ineinandergeschachtelte Schichten: In den `later`-Vertrag ist ein `multiple`-Vertrag eingeschachtelt, und dort wiederum der `one-euro`-Vertrag. Aus dieser Sicht ist der `multiple`-Vertrag der «innere» Vertrag des `later`-Vertrags.

**Verträge kombinieren** | Jetzt wo die Zero-Bonds erledigt sind, können wir überlegen, ob es noch weitere Möglichkeiten geben sollte, Verträge herzustellen. Immer wenn Du Repräsentationen mit Kombinatoren baust, solltest Du nach Möglichkeiten suchen, aus zwei Dingen ein größeres Ding zu bauen.

Beispiele für dieses Prinzip haben wir in diesem Kapitel schon mehrere gesehen:

- Ein Hauptfluss und ein Nebenfluss werden zu einem Fluss kombiniert.
- Zwei Bilder werden übereinander, nebeneinander und aufeinander zu einem Bild kombiniert.

Auf die Verträge bezogen bedeutet dies, dass wir zwei Verträge zu einem zusammensetzen, der die Rechte und Pflichten von beiden kombiniert. So könnte dies gehen:

```
; Eine Kombinationsvertrag besteht aus:
; - Vertrag Nr. 1
; - Vertrag Nr. 2
```

Das sind wieder zusammengesetzte Daten mit zwei Selbstbezügen:

```
(define-record both
  make-both
  both?
  (both-contract-1 contract)
  (both-contract-2 contract))
```

Hier sind zwei Beispiele für `both`:

```
; Heute 100€, später nochmal 100€
(define both1 (make-both euro100 later1))
; Heute 200€, später nochmal 200€
(define both2 (make-both both1 later2))
```

Wir müssen noch die Definition von `contract` erweitern:

```
(define contract
  (signature (mixed one-euro multiple later both)))
```

**Der kleinste Vertrag** | Es gibt noch ein weiteres Prinzip beim Repräsentieren mit Kombinatoren: Suche immer nach dem *kleinsten* möglichen Baustein. Wir haben ja schon einen ziemlich kleinen Baustein definiert: `one-euro`. Aber oft ist der kleine Baustein eine Variation von «nichts», und so ist es auch hier: Der kleinste Vertrag ist einer ohne Rechte oder Pflichten. Ihn sollten wir unbedingt auch definieren – zunächst nur der Vollständigkeit halber, aber später werden wir ihn tatsächlich noch brauchen.

Solch ein «Nichts-Vertrag» hat, wie `one-euro` auch, keine Eigenschaften. Wir repräsentieren ihn aber, genau wie bei `one-euro`, durch einen Record, damit wir ihn mit Hilfe seiner Signatur in die Definition von `contract` einbauen können:

```
; Ein Nichts-Vertrag hat keine Eigenschaften
(define-record nothing
  make-nothing
  nothing?)
```

Entsprechend erweitern wir `contract`:

```
(define contract
  (signature
    (mixed one-euro multiple later both nothing)))
```

In Aufgabe 5.11 auf Seite 162 geht es um noch einen weiteren Kombinator.

**Verträge bewerten** | Wir haben viel Energie für die Repräsentation von Verträgen verwendet, aber noch gar nichts damit gemacht. Es fehlen noch ein paar saftige Funktionen!

Wenn wir einen Vertrag haben, wollen wir vielleicht wissen, wieviel Geld wir eigentlich insgesamt bekommen. Das nehmen wir uns als erste Aufgabe vor. Wir fangen an mit folgender Kurzbeschreibung:

```
; Für einen Vertrag berechnen, wieviel Geld wir bekommen
```

In der Kurzbeschreibung tauchen die Begriffe «Vertrag» und «Geld» auf, das übersetzen wir in folgende Signatur:

```
(: contract-payment (contract -> rational))
```

Die Funktion fasst alle Zahlungen zu einem Euro-Betrag zusammen, Zinsen ignorieren wir. Die Funktionsweise illustrieren wir wie immer mit Testfällen. Bei festen Euro-Beträgen ist recht klar, was herauskommen muss:

```
(check-expect (contract-payment euro100) 100)
(check-expect (contract-payment euro200) 200)
```

Die Beispiele `later1` und `later2` sind verzögerte Verträge über 100 € beziehungsweise 200 €. Die Verzögerung spielt bei `contract-payment` keine Rolle:

```
(check-expect (contract-payment later1) 100)
(check-expect (contract-payment later2) 200)
```

Bei den beiden `both`-Beispielen werden jeweils zwei Zahlungen von 100 € respektive 200 € kombiniert:

```
(check-expect (contract-payment both1) 200)
(check-expect (contract-payment both2) 400)
```

Die Schablone sieht so aus:

```
(define contract-payment
  (lambda (contract)
    ...))
```

Da es sich bei der Definition von `contract` um gemischte Daten handelt, sieht die erste Schablone für `contract-payment` so aus:

```
(define contract-payment
  (lambda (contract)
    (cond
      ((nothing? contract) ...)
      ((one-euro? contract) ...)
      ((multiple? contract) ...)
      ((later? contract) ...)
      ((both? contract) ...))))
```

Bei den einzelnen Fällen von `contract` handelt es sich jeweils um zusammengesetzte Daten, wir können also die Schablone um die Anwendungen der Selektoren ergänzen, soweit es welche gibt:

```
(define contract-payment
  (lambda (contract)
    (cond
      ((nothing? contract) ...)
      ((one-euro? contract) ...)
      ((multiple? contract)
       ...
       (multiple-number contract)
       (multiple-of contract)
       ...))
    ((later? contract)
     ...
     (later-date contract))
```

```

    (later-contract contract)
    ...)
  ((both? contract)
   ...
   (both-contract-1 contract)
   (both-contract-2 contract)
   ...)))

```

Es geht noch weiter:

```

(multiple-of contract),
(later-contract contract),
(both-contract-1 contract) und (both-contract-2 contract)

```

sind allesamt Selbstbezüge, wir können also dort rekursive Aufrufe hinschreiben:

```

(define contract-payment
  (lambda (contract)
    (cond
      ((nothing? contract) ...)
      ((one-euro? contract) ...)
      ((multiple? contract)
       ...
       (multiple-number contract)
       (contract-payment (multiple-of contract))
       ...)
      ((later? contract)
       ...
       (later-date contract)
       (contract-payment (later-contract contract))
       ...)
      ((both? contract)
       ...
       (contract-payment (both-contract-1 contract))
       (contract-payment (both-contract-2 contract))
       ...))))

```

Damit ist die Schablone fertig und wir können versuche, für jeden Fall etwas sinnvolles hinzuschreiben. Bei den ersten beiden ist es nicht so schwierig: Bei einem **nothing**-Vertrag wird nichts ausgezahlt, bei einem **one-euro**-Vertrag wird ein 1 € ausgezahlt:



```
(define contract-payment
  (lambda (contract)
    (cond
      ((nothing? contract) 0)
      ((one-euro? contract) 1)
      ...)))
```

Wir nehmen uns als Nächstes den `multiple`-Fall vor: Da stehen in der Schablone schon:

- `(multiple-number contract)`, also die *Anzahl*
- `(contract-payment (multiple-of contract))`, also *was bei einem ausgezahlt* wird

Diese beiden Größen müssen wir miteinander multiplizieren, um die Gesamtsumme zu berechnen, die ausgezahlt wird:

```
(define contract-payment
  (lambda (contract)
    (cond
      ...
      ((multiple? contract)
       (* (multiple-number contract)
          (contract-payment (multiple-of contract))))
      ...)))
```

Nun ist der `later`-Fall an der Reihe. Für `contract-payment` ist `(later-date contract)`, also der Zeitpunkt, zu dem der Vertrag aktiv wird, irrelevant: Es geht ja nur um den Betrag des «inneren» Vertrags, und der steht schon in der Schablone:

```
(define contract-payment
  (lambda (contract)
    (cond
      ...
      ((later? contract)
       (contract-payment (later-contract contract)))
      ...)))
```

Zu guter letzt kommt `both`. Die Zahlungen, die sich aus den beiden Teilverträgen ergeben, stehen schon in der Schablone. Wir müssen sie lediglich addieren. Die vollständige Funktion sieht dann so aus:

```
(define contract-payment
```

```
(lambda (contract)
  (cond
    ((nothing? contract) 0)
    ((one-euro? contract) 1)
    ((multiple? contract)
     (* (multiple-number contract)
        (contract-payment (multiple-of contract))))
    ((later? contract)
     (contract-payment (later-contract contract)))
    ((both? contract)
     (+ (contract-payment (both-contract-1 contract))
        (contract-payment (both-contract-2 contract))))))
```

**Verträge verwalten** | Eine weitere nützliche Funktion von Verträgen betrifft die Entwicklung eines Vertrags über die Zeit: Je länger ein Vertrag geschlossen ist, desto mehr Zahlungen werden dadurch getätigt und entsprechend bleiben weniger Zahlungen übrig. Um das abzubilden, brauchen wir eine Funktion, die ausrechnet, was von einem Vertrag nach einem bestimmten Zeitpunkt noch übrigbleibt. So könnte die Kurzbeschreibung lauten:

*; Was bleibt übrig, nachdem zu einem gegebenen Datum ausgezahlt wurde?*

Um diese Funktion zu schreiben, müssen wir uns erst einmal überlegen, wie dieser «Rest» von einem Vertrag repräsentiert werden kann. Die richtige Idee ist so einfach, dass es gar nicht so einfach ist, sie zu sehen: Den «Rest eines Vertrags» können wir wieder als Vertrag repräsentieren. Wir müssen also keine neue Datenanalyse durchführen und können mit der Signatur fortfahren. Die Funktion braucht als Eingabe einen Vertrag; als Ausgabe kommt der neue Vertrag heraus. Außerdem brauchen wir noch das Datum, das in der Kurzbeschreibung erwähnt wird:

```
(: contract-rest (contract date -> contract))
```

Kommen wir zu einigen Testfällen: Die Verträge `euro100` und `euro200` werden sofort ausgezahlt, da bleibt also nichts mehr übrig:

```
(check-expect (contract-rest euro100 date1) (make-nothing))
(check-expect (contract-rest euro200 date1) (make-nothing))
```

Das Beispiel `later1` kommt zum Datum `date1` zur Auszahlung (der 26. Juli 2019), ab diesem Datum ist also nichts mehr übrig:

```
(check-expect (contract-rest later1 date1) (make-nothing))
```

Der Beispielvertrag `later2` hingegen wird erst zu `date2` ausgezahlt, das ist Weihnachten 2019.<sup>2</sup> Zum Zeitpunkt `date1` passiert deswegen noch nichts, der Vertrag kommt also unverändert zurück:

```
(check-expect (contract-rest later2 date1) later2)
```

Interessant wird es bei den `both`-Verträgen. Bei `both1` passiert der eine Teil sofort, der andere wird bis `date1` verzögert. Zum Zeitpunkt `date1` bleibt allerdings von beiden Hälften nichts übrig:

```
(check-expect (contract-rest both1 date1) (make-nothing))
```

Bei `both2` bleibt allerdings zum Zeitpunkt `date1` noch etwas übrig, nämlich die «rechte» Hälfte, `later2`:

```
(check-expect (contract-rest both2 date1) later2)
```

Aber zum Zeitpunkt `date2` ist auch das perdu:

```
(check-expect (contract-rest both2 date2) (make-nothing))
```

Wir schreiten nun zur Funktion selbst. Hier das Gerüst:

```
(define contract-rest
  (lambda (contract date)
    ...))
```

Die Eingabe-Signatur enthält ein `contract`, wir können also wieder die Schablone für gemischte Daten verwenden:

```
(define contract-rest
  (lambda (contract date)
    (cond
      ((nothing? contract) ...)
      ((one-euro? contract) ...)
      ((multiple? contract) ...)
      ((later? contract) ...)
      ((both? contract) ...))))
```

Wir ergänzen als Nächstes die Schablonen für die einzelnen Fälle, bei denen es sich ja jeweils um zusammengesetzte Daten handelt:

```
(define contract-rest
```

<sup>2</sup> Als dieses Kapitel geschrieben wurde, lag das noch in der Zukunft.

```

(lambda (contract date)
  (cond
    ((nothing? contract) ...)
    ((one-euro? contract) ...)
    ((multiple? contract)
     ...
     (multiple-number contract)
     (multiple-of contract)
     ...))
    ((later? contract)
     ...
     (later-date contract)
     (later-contract contract)
     ...))
    ((both? contract)
     ...
     (both-contract-1 contract)
     (both-contract-2 contract)
     ...))))

```

Fast hätten wir es vergessen – in den Fällen für `multiple`, `later` und `both` gibt es ja Selbstreferenzen und damit rekursive Aufrufe:

```

(define contract-rest
  (lambda (contract date)
    (cond
      ((nothing? contract) ...)
      ((one-euro? contract) ...)
      ((multiple? contract)
       ...
       (multiple-number contract)
       (contract-rest (multiple-of contract) date)
       ...))
      ((later? contract)
       ...
       (later-date contract)
       (contract-rest (later-contract contract) date)

```

```

...
((both? contract)
...
(contract-rest (both-contract-1 contract) date)
(contract-rest (both-contract-2 contract) date)
...))
...
(date-year date)
(date-month date)
(date-day date)
...))

```

Wir können außerdem noch die Schablone für `date` bemühen, wobei es sich zum zusammengesetzte Daten handelt. Das Datum ist nur im `later`-Fall relevant – wir füllen die Schablone aus, wenn wir zu diesem Fall kommen.

Wir fangen wieder mit den einfachen Fällen an. Bei `nothing` und `one-euro` bleibt jeweils nichts übrig:

```

(define contract-rest
  (lambda (contract date)
    (cond
      ((nothing? contract) (make-nothing))
      ((one-euro? contract) (make-nothing))
      ...)))

```

#### AUFGABE 5.4

Sind beide Fälle schon durch Testfälle abgedeckt? Wenn nicht, schreibe einen!



Als Nächstes ist `multiple` an der Reihe. Der rekursive Aufruf aus der Schablone beantwortet die folgende Frage: «Wieviel ist von *einem* Vertrag nach dem Datum noch übrig?» Wenn wir mit  $n$  das Ergebnis von `(multiple-number contract)` bezeichnen, dann müssen wir die Frage beantworten, wieviel von  $n$  Verträgen übrig bleibt. Im übertragenen Sinne futtern  $n$  Esser in der Mensa an jeweils der gleichen Portion Brei, werden aber unterbrochen. Wenn wir wissen, wieviel von einer Portion jeweils übrigbleibt, wieviel bleibt dann insgesamt übrig? Wir müssen wieder mit  $n$  multiplizieren:

```

(define contract-rest

```

```
(lambda (contract date)
  (cond
    ...
    ((multiple? contract)
     (make-multiple (multiple-number contract)
                     (contract-rest (multiple-of contract) date)))
    ...)))
```

Als Nächstes ist der **later**-Fall dran. Der rekursive Aufruf beantwortet die Frage, was von dem Vertrag übrigbleibt, falls das Datum **date** schon vorbei beziehungsweise gerade heute ist. Wenn **date** noch in der Zukunft liegt, passiert gar nichts und die Funktion sollte den gesamten Vertrag **contract** zurückliefern.

Das heißt, das **date** in zwei Klassen zerfällt, nämlich (a) vor oder gleich dem Datum (**later-date contract**) oder (b) danach. Wir brauchen also eine Verzweigung danach, und dazu müssen wir die beiden Daten miteinander vergleichen. Diese Aufgabe ist hinreichend kompliziert, dass wir sie zurückstellen und uns fest vornehmen, später eine Funktion mit folgender Kurzbeschreibung und Signatur zu schreiben:

```
; Ist ein Datum früher als ein anderes?
(: date<=? (date date -> boolean))
```

Mit Hilfe dieser Funktion können wir den **later**-Fall folgendermaßen fertigstellen:

```
(define contract-rest
  (lambda (contract date)
    (cond
      ...
      ((later? contract)
       (if (date<=? (later-date contract) date)
           (contract-rest (later-contract contract) date)
           contract))
      ...)))
```

Wir haben dabei etwas geschummelt und uns darum gedrückt, die Schablone für **date** auszufüllen. Das müssen wir bei **date<=?** nachholen.

Es bleibt der Fall für **both**. Die rekursiven Aufrufe aus der Schablone liefern, was vom «linken» beziehungsweise «rechten» Vertrag übriggeblieben ist. Wir müssen die Ergebnisse nur wieder zusammensetzen:

```

(define contract-rest
  (lambda (contract date)
    (cond
      ...
      ((both? contract)
       (make-both (contract-rest (both-contract-1 contract) date)
                  (contract-rest (both-contract-2 contract) date))))))

```

Fertig!

Also fast: Wir müssen ja noch die Funktion `date<=?` schreiben. Für einen Teil davon spannen wir Dich ein:

#### AUFGABE 5.5

Schreibe ausreichende Testfälle für `date<=?`!



Hier ist das Gerüst:

```

(define date<=?
  (lambda (date1 date2)
    ...))

```

Das es sich sowohl bei `date1` als auch `date2` um zusammengesetzte Daten handelt

```

(define date<=?
  (lambda (date1 date2)
    ...
    (date-year date1) (date-year date2)
    (date-month date1) (date-month date2)
    (date-day date1) (date-day date2)
    ...))

```

Um den Rumpf richtig hinzubekommen, müssen wir sorgfältig alle möglichen Fälle abarbeiten. Zunächst müssen wir die beiden Jahreszahlen vergleichen – wenn sie bei `date1` und `date2` unterschiedlich sind, so steht das Ergebnis fest. Wenn nicht, so muss die Funktion die Monate vergleichen und schließlich die Tage. Das sieht am Ende so aus:

```

(define date<=?
  (lambda (date1 date2)
    (cond

```

```
((< (date-year date1) (date-year date2)) #t)
(> (date-year date1) (date-year date2)) #f)
((< (date-month date1) (date-month date2)) #t)
(> (date-month date1) (date-month date2)) #f)
((< (date-day date1) (date-day date2)) #t)
(> (date-day date1) (date-day date2)) #f)
(else #t))))
```

**Verträge vereinfachen** | Leider ist der Code für die `contract-rest`-Funktion zwar vollständig, funktioniert aber nicht wie erwartet: Die Tests schlagen fehl, und zwar fast alle! Wie kann das sein? Hier ist der erste:

```
(check-expect (contract-rest euro100 date1) (make-nothing))
```

Hier ist die Meldung vom Fehlschlag:

Der tatsächliche Wert `#<record:multiple 100 #<record:nothing>>`  
ist nicht der erwartete Wert `#<record:nothing>`.

Was ist passiert? Zur Erinnerung, der Vertrag `euro100` ist so definiert:

```
(define euro100 (make-multiple 100 (make-one-euro)))
```

Die `contract-rest`-Funktion hat erst einmal berechnet, was von dem einen Euro übriggeblieben ist und vom Rest dann 100 Stück zurückgegeben: 100 Stück von nichts. Das Ergebnis ist ja auch formal korrekt, aber eben nicht wie vom Testfall erwartet. Unpraktisch ist es auf jeden Fall:

#### AUFGABE 5.6

Finde noch drei weitere komplizierte Wege, «Verträge über nichts» aufzuschreiben. □

Um das Problem zu vermeiden, wäre es gut, wenn das Programm Verträge über «*n* mal nichts» gar nicht erst konstruieren würden. Dafür müssen wir beim Konstruktor `make-multiple` ansetzen. Wenn der auf einen `nothing`-Vertrag angewendet wird, muss das Ergebnis auch wieder ein `nothing`-Vertrag sein. Um das zu erreichen, müssen wir `make-multiple` umdefinieren. Da die Definition von `make-multiple` von `define-record` erledigt wird, wenden wir einen Trick an und benennen den Konstruktor erst einmal um, in dem wir die Record-Definition ändern:

```
(define-record multiple
  really-make-multiple
```



```
multiple?
(multiple-number rational)
(multiple-of contract))
```

Der Name `really-make-multiple` ist reine Konvention – solange der Name nur anders ist als `make-multiple`! Als nächstes definieren wir eine neue Funktion mit dem alten Namen und dem gleichen Vertrag. Hier Kurzbeschreibung, Signatur und ein Testfall, der gerade `euro100` entspricht:

```
; multiple-Vertrag konstruieren, dabei vereinfachen
(: make-multiple (rational contract -> contract))

(check-expect (make-multiple 100 (make-nothing)) (make-nothing))
```

Hier ist das Gerüst für die Funktion:

```
(define make-multiple
  (lambda (factor contract)
    ...))
```

Die Eingabe `contract` zerfällt nun in zwei Fälle: `nothing`-Verträge und alle anderen. Im ersten Fall geben wir wieder einen `nothing`-Vertrag zurück, im zweiten Fall rufen wir den richtigen Konstruktor auf:

```
(define make-multiple
  (lambda (factor contract)
    (if (nothing? contract)
        (make-nothing)
        (really-make-multiple factor contract))))
```

Damit funktionieren schonmal die ersten beiden Testfälle, die ursprünglich fehlgeschlagen sind. Es bleiben aber noch weitere. Der erste ist dieser hier:

Der tatsächliche Wert `#<record:both #<record:nothing> #<record:nothing>>` ist nicht der erwartete Wert `#<record:nothing>`.

Das deutet auf ein analoges Problem wie bei `multiple` hin: Nichts und nochmal nichts ist eben nichts, aber der Konstruktor von `both` weiß das nicht. Wir wenden den gleichen Trick an und benennen zunächst den Konstruktor um:

```
(define-record both
```

```
really-make-both
both?
(both-contract-1 contract)
(both-contract-2 contract))
```

Als Nächstes definieren wir einen neuen Konstruktor selber:

```
; Kombinationsvertrag konstruieren, dabei vereinfachen
(: make-both (contract contract -> contract))
```

```
(check-expect (make-both (make-nothing) (make-nothing)) (make-
  nothing))
```

Durch diesen Testfall ist «nichts und nichts» abgedeckt. Das können wir aber noch etwas verallgemeinern. Das illustrieren am besten die folgenden Testfälle, bei denen jeweils nur die linke beziehungsweise rechte Seite des `both`-Vertrags nichts ist:

```
(check-expect (make-both (make-one-euro) (make-nothing)) (make-one-
  euro))
(check-expect (make-both (make-nothing) (make-one-euro)) (make-one-
  euro))
```

Hier ist das Gerüst zur Funktion:

```
(define make-both
  (lambda (contract-1 contract-2)
    ...))
```

Die Eingaben zerfallen in drei Klassen, nämlich wenn `contract-1` nichts ist, wenn `contract-2` nichts ist und alle anderen. Entsprechend brauchen wir eine Verzweigung mit drei Zweigen:

```
(define make-both
  (lambda (contract-1 contract-2)
    (cond
      ((nothing? contract-1) contract-2)
      ((nothing? contract-2) contract-1)
      (else
       (really-make-both contract-1 contract-2))))))
```

Und nun endlich verlaufen alle Tests erfolgreich.

Die Funktionen `make-multiple` und `make-both`, welche die ursprünglichen Konstruktor-n gleichen Namens ersetzen und bei der Konstruktion die entstehenden Verträge gleich schon etwas vereinfachen heißen auch *smart constructors*.

**Kombinatorrepräsentationen finden** | Die Autoren des Papers über Finanzverträge haben Monate gebraucht, um die beschriebene elegante Repräsentation zu finden. Sie sind nach zwei Prinzipien vorgegangen:

- Sie haben Finanzverträge in ihre kleinsten Bestandteile zerlegt.
- Sie haben nach Kombinatoren für Finanzverträge gesucht und welche gefunden, entsprechend Mantra 8 auf Seite 38.

Die Prinzipien sind auf viele andere Problemstellungen übertragbar.

## AUFGABEN

### AUFGABE 5.7

Schreibe Funktionen zum Umgang mit Duschprodukten!

1. Beginne mit zwei verschiedenen Duschprodukten: Seife und Shampoo. Bei Seife ist die Farbe und der pH-Wert relevant, bei Shampoo die Farbe und der Haartyp. Schreibe eine Datendefinition für Duschprodukte und den Code dafür!
2. Füge Duschgel als weiteres Duschprodukt hinzu, das zur Hälfte aus einer beliebigen Seife und zur Hälfte aus einem beliebigen Shampoo besteht. (Wir tun einfach so, als wäre das so.)
3. Füge eine «Mixtur» als weiteres Duschprodukt hinzu, das aus zwei beliebigen Duschprodukten zusammengemixt wird, mit jeweils wählbaren Anteilen für das erste und zweite Duschprodukt.
4. Schreibe eine Funktion, die den Seifenanteil eines beliebigen Duschprodukts berechnet.

### AUFGABE 5.8

Eine geometrische Figur in der zweidimensionalen Ebene ist entweder

- ein *Quadrat* parallel zu den Koordinatenachsen,
- ein *Kreis*
- oder eine *Überlagerung* zweier geometrischer Figuren. (Die Figuren werden übereinandergelegt; die Flächen der beiden Figuren werden also vereinigt.)

Führe eine Datenanalyse für geometrische Figuren durch! Schreibe ein Programm, in dem es möglich ist, geometrische Figuren anzulegen und zu überprüfen, ob ein gegebener Punkt innerhalb der Fläche einer gegebenen geometrischen Figur liegt.

Falls Du bei dieser Aufgabe eine Funktion brauchst, die eine Quadratwurzel zieht, so hat diese den Namen `sqrt` (für «square root») und folgende Signatur:

```
(: sqrt (number -> number))
```

(Es geht allerdings auch ohne `sqrt`.)

#### AUFGABE 5.9

Schreibe einen sinnvollen *smart constructor* für `later`-Verträge.

#### AUFGABE 5.10

Betrachte folgenden Vertrag:

```
(make-multiple 100 (make-multiple 100 (make-one-euro)))
```

Kann dieser Vertrag vereinfacht werden? Wenn ja, erweitere den entsprechenden *smart constructor*.

#### AUFGABE 5.11

Füge eine weitere Klasse von Verträgen hinzu, die `give`-Verträge: Ein `give`-Vertrag dreht alle Zahlungen um, die sich aus einem Vertrag ergeben. Das heißt, ein `give`-Vertrag von 1 € führt dazu, dass der Vertragsinhaber dem Vertragspartner 1 € *geben* muss und nicht bekommt. Erweitere die beiden Funktionen, so dass sie auch mit `give`-Verträgen klarkommen. Benutze, um Zahlungen «in die andere Richtung» zu repräsentieren, ein negatives Vorzeichen.

#### AUFGABE 5.12

Die Funktionen `contract-payment` und `contract-rest` passen nicht so recht zusammen: Die eine berechnet *alle* Zahlungen, die anderen den Rest eines Vertrags ab einem bestimmten Zeitpunkt. Schreibe eine Funktion `contract-payment-until`, die alle Zahlungen bis zu einem bestimmten Zeitpunkt berechnet!

#### AUFGABE 5.13

---

Statt nullstelliger zusammengesetzter Daten wie `nothing` oder `one-euro` könnte man theoretisch auch – wie bei Aufzählungen – feste Zeichenketten wie `"nothing"` und `"one-euro"` verwenden. Versuch dies umzusetzen!



## 6 PROGRAMMIEREN MIT LISTEN

Im vorigen Kapitel haben wir bereits Daten mit Selbstbezug kennengelernt, die Informationen variabler Größe repräsentieren können: Flüsse, Bilder und Finanzverträge. Die Repräsentationen dafür waren recht speziell das jeweilige Einsatzgebiet gekoppelt. In diesem Kapitel lernen wir eine besonders praktische Datendefinition mit Selbstbezug kennen, die in nahezu allen Einsatzgebieten nützlich ist: *Listen*.

Hier sind einige Listen aus dem täglichen Leben:

Brot	Herbert	1	Axl	Dauerwelle	Pumps
Butter	Mike	2	Slash	Dreadlocks	
Käse		3	Duff	Irokese	
		4	Dizzy	Vokuhila	
		5	Izzy		
		6	Melissa		
			Brain		
			Bumblefoot		

Keine dieser Listen ist auf ihre jeweilige Länge festgelegt: Zur Liste mit «Brot» könnten beispielsweise noch «Gurken» hinzukommen, zur Liste mit Axl, Slash undsoweiter könnte zum Beispiel noch Steven dazukommen.

### 6.1 LISTEN REPRÄSENTIEREN

In der Einführung haben wir Listen mit unterschiedlichen Sorten von Elementen gesehen. Für den Anfang konzentrieren wir uns zunächst einmal auf Listen von Zahlen, um die Dinge einfach zu halten. Dann schreiben wir eine Funktion, um die Zahlen aus so einer Liste zu addieren. Es gibt verschiedene Möglichkeiten, Listen zu repräsentieren. Die Repräsentation, die wir in diesem Kapitel einführen, hat als besonders einfach und universell verwendbar herausgestellt.

Diese Repräsentation ist nach dem gleichen Prinzip entstanden die zum Finanzverträge in Abschnitt 5.3 auf Seite 140: Wir suchen zunächst nach der einfachsten Liste, die wir uns vorstellen können. Man könnte auf die Idee kommen, dass dies eine Liste mit nur einem Element ist, aber noch einfacher ist eine mit gar keinen Elementen – die *leere Liste*. Natürlich werden wir auch nichtleere Listen brauchen, darum ist der erste Entwurf einer Datendefinition wie folgt:

```
; Eine Liste aus Zahlen ist eins der folgenden:
; - eine leere Liste
; - eine nichtleere Liste
```

Die leere Liste hat keine Bestandteile, wir werden sie aber später von den nichtleeren Listen unterscheiden müssen. Wir benutzen deshalb eine leere Record-Definition:

```
; leere Liste
(define-record empty-list
  make-empty-list
  empty?)
```

Beim Namen für das Prädikat haben wir etwas gemogelt. Von Amts wegen müsste es `empty-list?` heißen. Da es aber noch öfter vorkommen wird, kürzen wir es zu `empty?` ab. Da alle leeren Listen gleich sind, definieren wir eine Abkürzung dafür, die ebenfalls noch oft vorkommen wird:

```
(define empty (make-empty-list))
```

Nun sind die nichtleeren Listen dran. Nach dem Kombinatorprinzip aus dem vorigen Kapitel sollten wir eine nichtleere Liste konstruieren, indem wir eine bestehende Liste erweitern:

```
; Eine nichtleere Liste besteht aus:
; - dem ersten Element
; - einer Liste aus Zahlen mit den restlichen Elementen
```

Diese spezielle Definition für nichtleere Listen heißt *Cons-Liste*. Das zweite Feld enthält eine «Liste aus Zahlen» – da ist der Selbstbezug.

Wir ändern die Datendefinition für Listen entsprechend ab:

```
; Eine Liste aus Zahlen ist eins der folgenden:
; - eine leere Liste
; - eine Cons-Liste

; Eine Cons-Liste besteht aus:
; - dem ersten Element
; - einer Liste aus Zahlen mit den restlichen Elementen
```

Die Datendefinition für Cons wandeln wir in eine Record-Definition um. Dafür nehmen wir an, dass die Signatur für Listen von Zahlen `list-of-numbers` heißt – die werden wir im Nachgang definieren:



```
(define-record cons-list
  cons
  cons?
  (first number)
  (rest list-of-numbers))
```

Ähnlich wie bei `empty-list` haben wir die Funktionen der Record-Definition kürzer genannt als sonst, weil sie noch oft vorkommen werden: sonst: Der Konstruktor heißt `cons` statt `make-cons-list`, das Prädikat nur `cons?` statt `cons-list?` und bei den Selektoren fehlt jeweils das `cons-` vorn.

Es fehlt noch die Definition der Signatur `list-of-numbers` für «Liste von Zahlen», die aus leeren Listen und Cons-Listen gemischt wird:

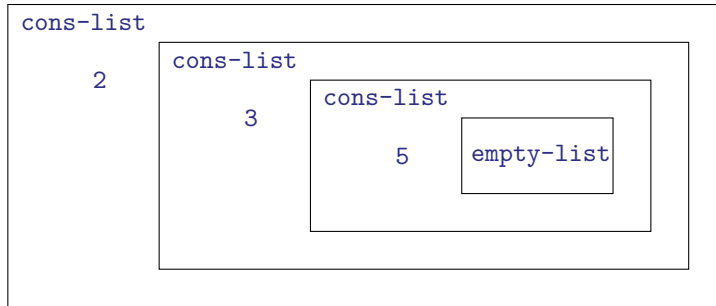
```
(define list-of-numbers
  (signature
    (mixed empty-list
            cons-list)))
```

Hier sind einige Beispiele für Listen:

```
; leere Liste
(define list0 empty)
; einelementige Liste mit der Zahl 42
(define list1 (cons 42 empty))
; Liste mit den Zahlen 1 2 3
(define list3 (cons 1 (cons 2 (cons 3 empty))))
; Liste mit den Zahlen e und pi
(define list2 (cons 2.7183 (cons 3.14159 empty)))
; Liste mit den Zahlen 2 3 5 7
(define list4 (cons 2 (cons 3 (cons 5 (cons 7 empty)))))
```

An den Beispielen kann man sehen, dass es einen Unterschied zwischen der intuitiven Struktur von Listen und ihrer Repräsentation gibt. Intuitiv besteht eine Liste aus einer Aneinanderreihung ihrer Elemente. Die Repräsentation ist aber geschachtelt, wie Abbildung 6.1 zeigt. Sie erlaubt uns insbesondere, aus kleineren Listen größere zu bauen und trotzdem die kleineren Listen wiederzuverwenden:

```
; Liste mit den Zahlen 1 2 3 5 7
(define list5 (cons 1 list4))
```



**Abbildung 6.1:** Struktur der Liste mit den Elementen 2 3 5

Dieses Beispiel zeigt außerdem, dass `cons` an eine existierende Liste *vorn* ein Element anhängt.

Nun sind wir bereit, eine Funktion über Listen zu schreiben, welche die Zahlen einer Liste aufaddiert. Kurzbeschreibung und Signatur sehen so aus:

```
; Summe der Elemente einer Liste von Zahlen berechnen
(: list-sum (list-of-numbers -> number))
```

Als Nächstes brauchen wir Testfälle, die wir mit Hilfe der Beispiel-Listen von oben (und gegebenenfalls einem Taschenrechner) konstruieren:

```
(check-expect (list-sum list2) 5.85989)
(check-expect (list-sum list3) 6)
(check-expect (list-sum list4) 17)
```

#### AUFGABE 6.1

Schreibe auch Testfälle für `list-sum` auf Basis der Beispiel-Listen `list0` und `list1`! Falls Du Dir nicht sicher bist, rate! □

Als Nächstes kommt wie immer das Gerüst:

```
(define list-sum
  (lambda (list)
    ...))
```

Die Signatur sagt, dass `list-sum` eine Liste als Eingabe akzeptiert – gemischte Daten. Die Schablone dazu sieht so aus:

```
(define list-sum
```

```
(lambda (list)
  (cond
    ((empty? list) ...)
    ((cons? list) ...))))
```

Beim `cons-list`-Fall handelt es sich um zusammengesetzte Daten, wir können also entsprechend der Schablone dafür Selektor-Aufrufe hinschreiben. Bei `rest` sitzt außerdem ein Selbstbezug, wir schreiben also auch einen rekursiven Aufruf in die Schablone:

```
(define list-sum
  (lambda (list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ...
       (first list)
       (list-sum (rest list))
       ...))))
```

Nun sind alle möglichen Konstruktionsanleitungen angewendet und wir vervollständigen die Schablone. Wir fangen an mit dem Fall für Cons-Listen. Dort steht in der Schablone `(first list)` und `(list-sum (rest list))`. Zum Verständnis hilft, wenn Du die beiden Ausdrücke ins Deutsche zu übersetzt:

- `(first list)` ist das *erste Element der Liste* und
- `(list-sum (rest list))` ist die *Summe der restlichen Elemente*.

Gefragt ist aber die Summe *aller Elemente*, dazu müssen wir die beiden noch addieren:

```
(+ (first list) (list-sum (rest list)))
```

Es bleibt der Fall der leeren Liste. Intuitiv hat die gar keine Summe, aber irgendetwas müssen wir da hinschreiben. Wir schreiben vorläufig 0 ins Programm, werden diese Wahl aber noch diskutieren:

```
(define list-sum
  (lambda (list)
    (cond
      ((empty? list) 0)
      ((cons? list)
       (+ (first list) (list-sum (rest list)))))))
```

Damit funktionieren die Testfälle und die Funktion ist fertig!

## AUFGABE 6.2

Verfolge die Auswertung der Testfälle im Stepper!



## AUFGABE 6.3

Ändere den Zweig für die leere Liste und schreibe etwas anderes als o hin. Funktionieren die Testfälle dann noch?



Tatsächlich ist o die einzig richtige Zahl, die im `empty`-Fall stehen darf. Das mag intuitiv einleuchten: die leere Liste ist schließlich «nichts» und o ist auch irgendwie «nichts». Ein weiteres Beispiel wird aber gleich zeigen, dass diese Argumentation zu einfach ist.

Wir schreiben als Nächstes eine Funktion, die alle Zahlen einer Liste *multipliziert*. Hier sind Kurzbeschreibung, Signatur, Tests und Gerüst:

```
; Produkt der Elemente einer Liste von Zahlen berechnen
(: list-product (list-of-numbers -> number))
```

```
(check-expect (list-product list1) 42)
(check-expect (list-product list3) 6)
(check-expect (list-product list4) 210)
```

```
(define list-product
  (lambda (list)
    ...))
```

Die Schablone entsteht genau wie bei `list-sum` aus den Konstruktionsanleitungen für gemischte Daten, zusammengesetzte Daten und Selbstbezüge:

```
(define list-product
  (lambda (list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ...
       (first list)
       ...
       (list-product (rest list))))))
```

Wir kümmern uns wieder zuerst um den `cons`-Fall. In der Schablone stehen `(first list)`, das erste Element und `(list-product (rest list))`, das Produkt der restlichen Elemente. Wir müssen nur noch beide miteinander multiplizieren, um das Produkt aller Elemente zu berechnen:

```
(* (first list) (list-product (rest list)))
```

Bleibt wieder der `empty`-Fall – wieder die leere Liste. Die Versuchung ist groß, wieder «nichts» hinzuschreiben, also `o`. Dann allerdings schlagen *alle* Testfälle fehl:

Der tatsächliche Wert 0 ist nicht der erwartete Wert 42.

Der tatsächliche Wert 0 ist nicht der erwartete Wert 6.

Der tatsächliche Wert 0 ist nicht der erwartete Wert 210.

Woran liegt das? Das kann man zum Beispiel anhand der Auswertung von

```
(list-product list3)
```

sehen. Die sieht so aus – wir haben einige Zwischenschritte ausgelassen:

```
(list-product list3)
↪ (list-product (cons 1 (cons 2 (cons 3 empty))))
↪ (cond ((empty? (cons 1 ...)) ...) ((cons? (cons 1 ...)) ...))
↪ (cond ((empty? (cons 1 ...)) ...) ((cons? (cons 1 ...)) ...))
↪ (* (first (cons 1 ...)) (list-product (rest (cons 1 ...))))
↪ (* 1 (list-product (rest (cons 1 ...))))
↪ (* 1 (list-product (cons 2 (cons 3 empty))))
↪ (* 1 (* 2 (list-product (cons 3 empty))))
↪ (* 1 (* 2 (* 3 (list-product empty))))
↪ (* 1 (* 2 (* 3 (cond ((empty? empty) 0) ...))))
↪ (* 1 (* 2 (* 3 (cond (#t 0) ...))))
↪ (* 1 (* 2 (* 3 0)))
```

Die Auswertung ist auf so einem guten Weg – aber um Schluss macht `list-product` alles kaputt und multipliziert mit `o`. Und so ist es mit *jeder* Eingabe. Davor stehen die Zahlen, die eigentlich miteinander multipliziert werden sollen. Wir müssten nur statt mit `o` mit einer Zahl multiplizieren, die das Ergebnis nicht kaputtmacht sondern genauso lässt, wie es ist. Im Fall der Multiplikation ist das die `1`, nicht die `o`. Die richtige Funktion sieht damit so aus:

```
(define list-product
  (lambda (list)
    (cond
```

```
((empty? list) 1)
((cons? list)
 (* (first list) (list-product (rest list))))))
```

Rückblickend können wir damit auch erklären, warum bei `list-sum` im `empty`-Fall `0` stehen muss: Die `0` macht beim Addieren das Ergebnis nicht kaputt. Darum spricht man auch davon, dass `0` das *neutrale Element* bezüglich der Addition ist, `1` das neutrale Element bezüglich der Multiplikation.

## 6.2 LISTEN MIT ANDEREN ELEMENTEN

Das Prinzip «Liste» ist nicht auf Listen von Zahlen beschränkt: Genauso gut kann man sich Listen aus Zeichenketten, den Tieren aus Kapitel 4.1 oder Bildern vorstellen. Um das zu ermöglichen, können wir die Datendefinition anpassen. Wir entfernen alle Erwähnungen von Zahlen und sprechen ausschließlich von «Elementen»:

```
; Eine Liste ist eins der folgenden:
; - eine leere Liste
; - eine Cons-Liste

; Eine Cons-Liste besteht aus:
; - dem ersten Element
; - einer Liste mit den restlichen Elementen
```

Tatsächlich können wir Listen bereits bauen, deren Elemente keine Zahlen sind:

```
(define string-list (cons "Mike" (cons "Sperber" empty)))
```

### AUFGABE 6.4

Probiere die Definition von `string-list` aus!



Diese Definition kann DrRacket zwar ausführen, aber es gibt jede Menge Signaturverletzungen.

Die verletzten Signaturen sind allesamt `number`, am prominentesten in der Record-Definition von `cons-list`. Wenn wir dort auch Zeichenketten verwenden wollen, ohne Listen aus Zahlen zu verbieten, müssen wir über `number` und `string` abstrahieren. Bei der Record-Definition geht

das, indem wir aus `cons-list` stattdessen `(cons-list-of element)` machen. Das wirkt wie ein `lambda` und wir können dann `element` innerhalb der Record-Definition benutzen:

```
(define-record (cons-list-of element)
  cons
  cons?
  (first element)
  (rest ...))
```

Diese Record-Definition definiert `cons-list` als Funktion, die eine Signatur akzeptiert (die Signatur der Elemente) und eine Signatur zurückliefert, nämlich die der Cons-Listen mit der gegebenen Signatur für die Elemente. Die Signatur können wir auch hinschreiben:

```
(: cons-list-of (signature -> signature))
```

Da `cons-list` eine Signatur konstruiert, nennen wir diese Funktion auch *Signatur-Konstruktor*. Das `element` ist ein *Signatur-Parameter*. Wir müssen also immer da, wo wir vorher `cons-list` stand, `(cons-list-of ...)` schreiben und eine Signatur für die Elemente angeben, also zum Beispiel `(cons-list-of number)` oder `(cons-list-of string)`

Die Signatur von `rest` war vorher `list-of-numbers`, die müssen wir also auch ändern. Wie genau, ergibt sich erst im Zusammenhang mit der abstrahierten Definition von `list-of-numbers`. Da es sich hier um eine ganz normale Definition mit `define` handelt, können wir ein `lambda` benutzen, um über den Signatur der Elemente zu abstrahieren:

```
(define list-of
  (lambda (element)
    (signature
     (mixed empty-list
      (cons-list-of element))))))
```

Wir können mit dieser Definition statt `list-of-numbers` also `(list-of number)` schreiben. Damit können wir auch die Record-Definition von `cons-list` vervollständigen, indem wir `(list-of element)` da hinschreiben, wo vorher `list-of-numbers` stand:

```
(define-record (cons-list-of element)
  cons
  cons?
  (first element)
  (rest (list-of element)))
```

Das erledigt die Signaturverletzungen.

Bei dieser neuen Form von `define-record` ist nicht mehr offensichtlich, wie die Signaturen des Konstruktors und der beiden Selektoren aussehen sollen, wenn wir sie ausschreiben. Wir brauchen dafür ein neues Konstrukt, um auszudrücken, dass potenziell bei jedem Aufruf von `cons`, `first` und `rest` eine andere Elementsignatur verwendet wird. Das sieht so aus:

```
(: cons (%element (list-of %element) -> (cons-list-of %element)))
(: first ((cons-list-of %element) -> %element))
(: rest ((cons-list-of %element) -> (list-of %element)))
```

Das % kennzeichnet *Signaturvariablen*. Die Signaturvariable `%element` steht für eine beliebige Signatur. Die Signatur-Deklaration von `cons` bedeutet außerdem, dass die Signatur von neuem Element (das erste `%element`) und die Signatur der Listenelemente (das `%element` in `(list %element)`) übereinstimmen sollten. (DrRacket überprüft dies aber nicht.)

Vielleicht ist Dir der Gedanke gekommen, dass statt `%element` in den obigen Signaturen auch `any` stehen könnte. Rein technisch gesehen ist das korrekt. Allerdings wären die resultierenden Signaturen weniger präzise. Nimm einmal an, dort stünde zum Beispiel das hier:

```
(: cons (any (list-of any) -> (cons-list-of any)))
```

Diese Signatur sagt «`cons` akzeptiert irgendeinen Wert und irgendeine Liste und produziert wieder irgendeine Liste». Zum Vergleich sagt die Signatur

```
(: cons (%element (list-of %element) -> (cons-list-of %element)))
```

aus, dass das neue Element, das `cons` akzeptiert, die gleiche Signatur erfüllen muss wie die schon vorhandenen Elemente der Liste und dann die Elemente der Resultatliste die gleiche Signatur erfüllen: Alle drei Vorkommen von `%element` müssen in einer konkreten Anwendung von `cons` die gleiche Signatur sein.

Mit diesen Änderungen an den Signatur-Deklarationen gehen die Signatur-Verletzungen ganz weg. Bevor das Programm aber wieder funktioniert, müssen wir noch eine weitere Fehlermeldung ausmerzen: `list-of-numbers` ist nicht mehr definiert, steht aber noch in den Signatur-Deklarationen von `list-sum` und `list-product`. Wir können entweder dort `list-of-numbers` durch `(list-of number)` ersetzen oder eine Signatur-Definition dazuschreiben:

```
(define list-of-numbers (signature (list-of number)))
```

Diese Definition muss vor den Definitionen von `list-sum` und `list-product` stehen, damit sie dort bekannt ist.

Abbildung 6.2 fasst die neue Form von `define-record` mit Signatur-Parametern zusammen.



Eine `define-record`-Form mit Signatur-Parametern hat folgende allgemeine Gestalt:

```
(define-record (t sp1 ... spk)
  c
  p
  (sel1 sig1)
  ...
  (seln sign))
```

- Die Signatur-Parameter  $sp_1 \dots sp_k$  können in den Signaturen  $sig_1 \dots sig_n$  verwendet werden.
- $t$  ist der Name des Signatur-Konstruktors mit der Signatur:

```
(: t (signature ... (k-mal) -> signature))
```

- $c, p$  und  $sel_1, \dots, sel_n$  sind wie bisher die Namen von Konstruktor, Prädikat und Selektoren.

**Abbildung 6.2:** `define-record` mit Signatur-Parametern

## 6.3 KONSTRUKTIONSANLEITUNG

Bei der Konstruktion von Funktionen, die Listen akzeptieren, haben wir drei Schablonen kombiniert: Die Schablonen für gemischte und zusammengesetzte Daten sowie die Schablone für Selbstbezüge. Wir werden noch viele Funktionen mit Listen als Eingabe schreiben, bei denen jedesmal genau die gleichen Schablonen kombiniert werden. Deshalb lohnt es sich, solchen Funktionen eine eigene Konstruktionsanleitung mit der entstandenen Schablonenkombination zu widmen:

### KONSTRUKTIONSANLEITUNG 18 (LISTEN ALS EINGABE: SCHABLONE)

Die Schablone für eine Funktion, die eine Liste akzeptiert, sieht folgendermaßen aus:

```
(: f (... (list-of elem) ... -> ...))
```

```
(define f
  (lambda (... list ...)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ...
       (first list))
```

Die eingebaute Funktion `list` erlaubt es, Listen aus ihren Elementen ohne Verwendung von `cons` zu erzeugen. Sie akzeptiert eine beliebige Anzahl von Argumenten, macht daraus eine Liste und gibt diese zurück:

```
(list 1 2 3)
↪ #<list 1 2 3>
(list "Ax1" "Slash" "Izzy")
↪ #<list "Ax1" "Slash" "Izzy">
```

Abbildung 6.3: `list`

```
...
(f ... (rest list) ...)
...
))))
```

## 6.4 EINGEBAUTE LISTEN

Da Listen in diesem Buch noch oft vorkommen und es umständlich ist, jedesmal die Definitionen für `cons` und `list-of` an den Anfang von Programmen zu setzen, macht die nächste Sprachebene das Programmieren mit Listen einfacher.

### Schalte deshalb in die nächste Sprachebene hoch:

Fange eine neue Datei an. Wähle danach wieder das Menü **Sprache** aus, darunter **Sprache auswählen** und dann **Schreibe Dein Programm!** (*ohne Anfänger*). Dann drücke noch einmal **Start**.

Ab sofort sind `cons`, `cons?`, `first` und `rest` vordefiniert genau wie der Signaturkonstruktor `list-of`. Es gibt außerdem eine praktische neue Funktion `list`, die in Abbildung 6.3 beschrieben ist.

In der neuen Sprachebene werden nichtleere Listen übersichtlicher ausgedruckt als bisher, nämlich als `#<list ...>`, wobei die Listenelemente zwischen den spitzen Klammern aufgereiht sind. Beispiele:

```
(list 1)
↪ #<list 1>
(list 1 2)
↪ #<list 1 2>
(list 1 2 3)
↪ #<list 1 2 3>
```

## 6.5 PARAMETRISCHE POLYMORPHIE

Keine Angst vor der hochtrabenden Überschrift: Wir schreiben in diesem Abschnitt eine recht einfache Funktion, welche die Länge einer Liste ermittelt. Hier ist die Kurzbeschreibung:

**; Länge einer Liste berechnen**

Der Anfang einer Signatur-Deklaration sieht so aus:

```
(: list-length ((list-of ...) -> natural))
```

Was können wir anstelle der Ellipse hinschreiben? Der Funktion sollte egal sein, auf welche Signatur die Elemente der Liste passen. Wir können dies zum Ausdruck bringen, indem wir eine Signaturvariable verwenden, genau wie bei den Signaturen von `cons`, `first` und `rest`. Das sieht so aus:

```
(: list-length ((list-of %element) -> natural))
```

Die Verwendung der Signaturvariable `%element` bedeutet, genau wie oben, dass die Signatur der Elemente der Liste bei jedem Aufruf von `list-length` anders sein kann.

Solche Funktionen, die Argumente akzeptieren, deren Signaturen Signaturvariablen enthalten, heißen *polymorph* oder auch *parametrisch polymorph* (weil die Signaturvariable eine Art Parameter abgibt), und das dazugehörige Konzept heißt *parametrische Polymorphie*: ein großes Wort, das hier für eine kleine Sache steht. Mehr Interessantere Beispiele für parametrische Polymorphie wird es in Kapitel 9 geben.

Weiter mit `list-length` – hier ist das Gerüst:

```
(define list-length
  (lambda (lis)
    ...))
```

Die Schablone ist wie gehabt:

```
(define list-length
  (lambda (lis)
    (cond
      ((empty? lis) ...)
      ((cons? lis)
       ... (first lis) ...
       ... (list-length (rest lis)) ...))))
```

Es ist `list-length` egal, was der Wert von `(first lis)` ist. Die Länge der Liste ist unabhängig davon, was für Werte sich darin befinden: entscheidend ist nur, wieviele es sind. (Dieser Umstand

ist gerade verantwortlich für die parametrische Polymorphie.) Deshalb können wir `(first lis)` aus der Schablone streichen und diese dann zum vollständigen Rumpf ergänzen:

```
(define list-length
  (lambda (lis)
    (cond
      ((empty? lis) 0)
      ((cons? lis)
       (+ 1
          (list-length (rest lis)))))))
```

Die Funktion `list-length` ist unter dem Namen `length` fest eingebaut.

## 6.6 FUNKTIONEN, DIE LISTEN PRODUZIEREN

In den vorherigen Abschnitten haben wir ausschließlich Funktionen programmiert, die Listen *akzeptieren*. In diesem Abschnitt schreiben wir Funktionen, die Listen *produzieren*. Das geht mit Techniken, die wir bereits vorgestellt haben. Wir machen die Sache interessanter, indem wir in einem ersten Beispiel Listen von zusammengesetzten Daten betrachten und in einem zweiten Beispiel zwei Listen verarbeiten.

### 6.6.1 GÜRTELTIERE ÜBERFAHREN

Auf Seite 97 haben wir die Funktion `run-over-dillo` geschrieben, die für das Überfahren von Gürteltieren zuständig ist. In diesem Abschnitt schreiben wir die Funktion, die das gleich massenweise erledigt, beispielsweise für alle Gürteltiere auf einem Highway.

Dazu übernehmen wir Daten- und Record-Definition von Gürteltieren aus Abschnitt 3.4 sowie die Funktionsdefinition von `run-over-dillo`. Du kannst auch einfach den Gürteltier-Code erweitern – dann musst Du allerdings die Sprachebene auf `Schreibe Dein Programm!` hochschalten, damit die Listen zur Verfügung stehen.

Gürteltiere können wir in Listen stecken, ebenso wie Zahlen, Zeichenketten oder boolesche Werte. Hier ist ein Beispiel:

```
; Gürteltiere auf Highway 75
(define highway75 (list dillo1 dillo2 dillo3 dillo4))
```

(`dillo1`, `dillo2`, `dillo3` und `dillo4` sind die Beispieltiere aus Abschnitt 3.4.)

Diese Liste hat die Signatur (`list-of dillo`). Wenn wir eine Funktion schreiben wollen, die alle Gürteltiere aus einer Liste überfährt, müsste diese also folgende Kurzbeschreibung und Signatur haben:

```
; Gürteltiere überfahren
(: run-over-dillos ((list-of dillo) -> (list-of dillo)))
```

Als Testfall kann obige Beispielliste erhalten:

```
(check-expect (run-over-dillos highway75)
               (list (make-dillo 55000 #f)
                     dillo2
                     (make-dillo 60000 #f)
                     dillo4))
```

Zur Erinnerung: `dillo2` und `dillo4` sind bereits tot, dementsprechend sind sie überfahren wie zuvor.

Hier ist das Gerüst:

```
(define run-over-dillos
  (lambda (dillos)
    ...))
```

Die Funktion akzeptiert eine Liste als Eingabe, wir können also, wie schon so oft, die entsprechende Schablone zum Einsatz bringen:

```
(define run-over-dillos
  (lambda (dillos)
    (cond
      ((empty? dillos) ...)
      ((cons? dillos)
       ... (first dillos) ...
       ... (run-over-dillos (rest dillos)) ...))))
```

Im ersten Zweig ist die Sache klar: Geht eine leere Liste rein, kommt auch eine leere Liste raus. Im zweiten Zweig können wir uns erst einmal um das erste Gürteltier kümmern. Wir haben ja bereits eine Funktion, die ein einzelnes Gürteltier überfährt; diese können wir auf das erste Element der Liste anwenden:

```
(define run-over-dillos
```

```
(lambda (dillos)
  (cond
    ((empty? dillos) empty)
    ((cons? dillos)
     ... (run-over-dillo (first dillos)) ...
     ... (run-over-dillos (rest dillos)) ...))))
```

Lesen wir noch einmal die beiden Ausdrücke, die im zweiten Zweig stehen:

- `(run-over-dillo (first dillos))` ist das erste Gürteltier der Liste, überfahren.
- `(run-over-dillos (rest dillos))` ist eine Liste der restlichen Gürteltiere, überfahren.

Gefragt ist eine Liste *aller* Gürteltiere, überfahren: Wir müssen also nur die Resultate der beiden Ausdrücke mit `cons` kombinieren:

```
(define run-over-dillos
  (lambda (dillos)
    (cond
      ((empty? dillos) empty)
      ((cons? dillos)
       (cons (run-over-dillo (first dillos))
              (run-over-dillos (rest dillos)))))))
```

Fertig!

Dieses Beispiel zeigt, dass wir für Funktionen, die Listen produzieren, keine neue Technik brauchen: Wenn eine Funktion eine leere Liste produzieren soll, benutzen wir an der entsprechenden Stelle `empty`, und bei nichtleeren Listen benutzen wir `cons`, bringen also die Schablone für Funktionen zum Einsatz, die zusammengesetzte Daten produzieren.

## 6.6.2 GÜRTELTIERE EXTRAHIEREN

Wir nehmen uns noch eine weitere Übung vor, bei der Listen produziert werden: Aus einer Liste von Gürteltieren wollen wir alle (noch) lebenden herausholen, vielleicht, um sie in Sicherheit zu bringen.

So sehen Kurzbeschreibung, Signatur, ein Testfall und Gerüst aus:

```
; Lebendige Gürteltiere aufsammeln
(: live-dillos ((list-of dillo) -> (list-of dillo)))
```

```
(check-expect (live-dillos highway75)
              (list dillo1 dillo3))
```

```
(define live-dillos
  (lambda (dillos)
    ...))
```

Wir fügen als Nächstes die Standard-Schablone für Funktionen ein, die Listen akzeptieren:

```
(define live-dillos
  (lambda (dillos)
    (cond
      ((empty? dillos) ...)
      ((cons? dillos)
       ... (first dillos) ...
       ... (live-dillos (rest dillos)) ...))))
```

Die Schablone ergänzen wir nun zur fertigen Funktion. Der ersten Fall betrifft die leere Liste – aus einer leeren Liste von Gürteltiere können wir auch keine lebendigen Gürteltiere herausholen, da kommt also `empty` hin.

Im zweiten Fall ist es etwas aufwendiger. Wir machen uns wieder klar, was die beiden Bestandteile bedeuten, die in der Schablone stehen:

- `(first dillo)` ist das erste Gürteltier in der Liste.
- `(live-dillos (rest dillos))` sind die lebendigen unter den restlichen Gürteltiere.

Da die Gesamtaufgabe ist, die lebendigen Gürteltiere zu extrahieren, sollten wir beim ersten Gürteltier feststellen, ob es (noch) lebt. Da steht dann folgender Zwischenstand

```
(define live-dillos
  (lambda (dillos)
    (cond
      ((empty? dillos) empty)
      ((cons? dillos)
       ... (dillo-alive? (first dillos)) ...
       ... (first dillos) ...
       ... (live-dillos (rest dillos)) ...))))
```

Der Aufruf `(dillo-alive? (first dillos))` liefert einen booleschen Wert – mit dem können wir eigentlich nur eine Sache anfangen, nämlich ihn als Bedingung in einer binären `if`-Verzweigung benutzen:

```
(define live-dillos
  (lambda (dillos)
    (cond
      ((empty? dillos) empty)
      ((cons? dillos)
       (if (dillo-alive? (first dillos))
           ...
           ...))
      ... (first dillos) ...
      ... (live-dillos (rest dillos)) ...))))
```

Wir müssen uns jetzt überlegen, was wir in die beiden Zweige des `if` schreiben. Im ersten Zweig – der Konsequente – ist das Gürteltier (`first dillos`) (noch) am Leben, im zweiten – der Alternative – ist es tot.

In der Alternative (Gürteltier tot) ist (`live-dillos (rest dillos)`) schon die richtige Antwort. (`first dillos`) fällt unter den Tisch:

```
(define live-dillos
  (lambda (dillos)
    (cond
      ((empty? dillos) empty)
      ((cons? dillos)
       (if (dillo-alive? (first dillos))
           ...
           (live-dillos (rest dillos))))
      ... (first dillos) ...))))
```

Auch, wenn wir den rekursiven Aufruf (`live-dillos (rest dillos)`) scheinbar «verbraucht» haben – im anderen Zweig (Gürteltier lebt (noch)) brauchen wir die restlichen lebendigen Gürteltiere noch einmal. Wir müssen diesmal aber darauf achten, dass wir das erste Gürteltier in der Ausgabe wieder unterbringen und vorn an die restlichen Gürteltiere dran-`cons`-en:

```
(define live-dillos
  (lambda (dillos)
    (cond
      ((empty? dillos) empty)
      ((cons? dillos)
       (if (dillo-alive? (first dillos))
```



```
(cons (first dillos)
      (live-dillos (rest dillos)))
(live-dillos (rest dillos))))))
```

Fertig!

### 6.6.3 ZWEI LISTEN ANEINANDERHÄNGEN

In unserem nächsten Beispiel ist eine Funktion `concatenate` gefragt, die zwei Listen aneinanderhängt. Kurzbeschreibung, Signatur und Gerüst sehen folgendermaßen aus:

```
; zwei Listen aneinanderhängen
(: concatenate ((list-of %element) (list-of %element)
               -> (list-of %element)))
```

Diese Funktion ist wie `list-length` polymorph – in der Signatur kommt die Signaturvariable `%element` vor. Beachte, dass `%element` *dreimal* vorkommt: Die Signatur besagt, dass die Signaturen der Elemente der beiden Eingabelisten und die Signatur der Ausgabeliste sich einander entsprechen.

Hier sind Testfall und Gerüst für `concatenate`:

```
(check-expect (concatenate (list 1 2 3) (list 4 5 6))
              (list 1 2 3 4 5 6))
```

```
(define concatenate
  (lambda (list1 list2)
    ...))
```

Konstruktionsanleitung 18 auf Seite 175 ist eigentlich nur für Funktionen gedacht, die eine einzelne Liste akzeptieren. Welche von beiden ist das *list* aus der Anleitung? Im Zweifelsfall können wir beide Alternativen ausprobieren. Wir versuchen es zunächst mit `list2`:

```
(define concatenate
  (lambda (list1 list2)
    (cond
      ((empty? list2) ...)
      ((cons? list2)
       ... (first list2) ...
       ... (concatenate list1 (rest list2)) ...))))
```

Der erste Zweig des `cond` ist noch einfach: Wenn `list2` leer ist, muss `list1` herauskommen. Jedoch wäre für das obige Beispiel der Wert von

```
(concatenate list1 (rest list2))
```

die folgende Liste:

```
#<list 1 2 3 5 6>
```

Bei dieser Liste fehlt das Element 4 in der Mitte, und es ist nicht ersichtlich, wie unsere Funktion sie passend ergänzen könnte. Diese Möglichkeit führt also in eine Sackgasse. Wir versuchen deshalb, die Schablone auf `list1` statt auf `list2` anzuwenden:

```
(define concatenate
  (lambda (list1 list2)
    (cond
      ((empty? list1) ...)
      ((cons? list1)
       ... (first list1) ...
       ... (concatenate (rest list1) list2) ...))))
```

Die erste Ellipse ist einfach zu ersetzen: Ist die erste Liste leer, ist das Ergebnis die zweite Liste `list2`. Für den zweiten Fall sollten wir uns noch einmal ins Gedächtnis rufen, was für einen Wert `(concatenate (rest list1) list2)` liefert: das Ergebnis dieses Aufrufs ist eine Liste, die aus `(rest list1)` und `list2` zusammengesetzt wurde. Auf das obige Beispiel übertragen, falls `list1` die Liste `#<list 1 2 3>` ist und `list2` die Liste `#<list 4 5 6>`, so hat `(rest list1)` den Wert `#<list 2 3>`. Der Wert von `(concatenate (rest list1) list2)` wäre also:

```
#<list 2 3 4 5 6>
```

Es fehlt das erste Element von `list1`, `(first list1)`, das vorn an das Ergebnis angehängt werden muss. Das geht mit `cons`:

```
(define concatenate
  (lambda (list1 list2)
    (cond
      ((empty? list1) list2)
      ((cons? list1)
       (cons (first list1)
             (concatenate (rest list1) list2))))))
```

Dieses Beispiel zeigt ein weiteres Schablonelement, das noch öfter vorkommen wird: Wie bei anderen zusammengesetzten Daten müssen Funktionen, die Listen konstruieren sollen, irgendwo ein `cons` enthalten.

Da viele Programme die Funktionen `list-length` und `concatenate` benötigen, sind sie in den Lehrsprachen übrigens bereits unter den Namen `length` und `append` eingebaut.

## 6.7 MINIMUM EINER LISTE UND LOKALE VARIABLEN

Wir nehmen uns noch eine weitere Übung vor, nämlich das Minimum einer Liste von Zahlen zu berechnen. Das ist scheinbar einfach, birgt aber einige Untiefen. Hier sind der erste Anlauf für Kurzbeschreibung, Tests und Gerüst:

```
; Minimum einer Liste von Zahlen berechnen
(: list-min ((list-of real) -> real))

(check-expect (list-min (list 5 3 1 4)) 1)
(check-expect (list-min (list 5 3 1 4 -4 3)) -4)
```

Vielleicht siehst Du schon das erste Problem. Die Tests sparen einen einfachen aber wichtigen Fall aus, nämlich die leere Liste – und die hat kein Minimum.

Eine Möglichkeit wäre, in dem Fall `violation` aufzurufen wie in Abschnitt 2.10 auf Seite 74. (Siehe dazu Aufgabe 6.16 auf Seite 197.) Eine andere ist, wie in Abschnitt 4.4 auf Seite 124 einen Extra-Wert für diesen Fall zu benutzen. Das hat wie dort den Vorteil, dass wir in der Signatur ablesen können, dass die Funktion auch fehlschlagen kann. Wir führen also folgenden leeren Record-Typ dafür ein:

```
; Kein Resultat
(define-record no-result
  make-no-result
  no-result?)
```

Die Signatur erweitern wir entsprechend:

```
(: list-min ((list-of real) -> (mixed real no-result)))
```

Außerdem schreiben wir noch einen Testfall dafür:

```
(check-expect (no-result? (list-min empty)) #t)
```

Nun können wir wie gewohnt die Schablone für Listen anwenden:

```
(define list-min
  (lambda (list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ... (first list) ...
       ... (list-min (rest list)) ...))))
```

Der `empty`-Fall ist gerade der, in dem es kein Resultat gibt:

```
(define list-min
  (lambda (list)
    (cond
      ((empty? list) (make-no-result))
      ((cons? list)
       ... (first list) ...
       ... (list-min (rest list)) ...))))
```

Im `cons`-Fall müssen wir darauf achten, dass der rekursive Aufruf von `list-min` gemischte Daten liefert. Wir müssen also eine Verzweigung einbauen, die zwischen `no-result` und einer Zahl unterscheidet. Wir machen das mit einer binären Verzweigung:

```
(define list-min
  (lambda (list)
    (cond
      ((empty? list) (make-no-result))
      ((cons? list)
       (if (no-result? (list-min (rest list)))
           ...
           ...))
      ... (first list) ...
      ... (list-min (rest list)) ...))))
```

Falls der rekursive Aufruf ein `no-result` liefert, haben wir nur das erste Element der Liste (`first list`) in der Hand – der Rest der Liste muss dann ja leer sein. In dem Fall muss (`first list`) auch das Minimum sein:

```
(define list-min
```

```
(lambda (list)
  (cond
    ((empty? list) (make-no-result))
    ((cons? list)
     (if (no-result? (list-min (rest list)))
         (first list)
         ...))
    ... (first list) ...
    ... (list-min (rest list)) ...))))
```

Falls der rekursive Aufruf von `list-min` ein Resultat liefert, ist dieses Resultat das Minimum der restlichen Liste. Um das Minimum der Gesamtliste zu berechnen, müssen wir es allerdings noch mit `(first list)` vergleichen, das ebenfalls das Minimum sein könnte. Abhängig vom Ergebnis des Vergleichs ist das eine oder das andere das Minimum:

```
(define list-min
  (lambda (list)
    (cond
      ((empty? list) (make-no-result))
      ((cons? list)
       (if (no-result? (list-min (rest list)))
           (first list)
           (if (< (first list)
                 (list-min (rest list)))
               (first list)
               (list-min (rest list))))))))
```

Diese Funktion ist korrekt.

Allerdings hat sie noch einen Makel: `(list-min (rest list))` taucht insgesamt dreimal auf, und es ist tatsächlich möglich, dass alle drei tatsächlich ausgewertet werden: Dreimal die gleiche Arbeit. (Du kannst Dich mit dem Stepper davon überzeugen.) Wir sollten dafür sorgen, dass der rekursive Aufruf nur einmal ausgewertet wird. Dazu führen wir ein neues Programmierkonstrukt ein, die *lokale Variable* (siehe Abbildung 6.4), für die zunächst eine Definition innerhalb des Rumpfes von `list-min` anlegen:

```
(define list-min
  (lambda (list)
    (cond
```

Im Rumpf eines **lambda** und im Zweig eines **cond** kann jeweils am Anfang eine Definition stehen. Eine solche *lokale Definition* bindet eine *lokale Variable*.

Diese Definition wird jedesmal ausgewertet, wenn die Auswertung beim Rumpf des **lambda** beziehungsweise beim Zweig des **cond** ankommt und ist auch nur innerhalb davon («lokal») sichtbar.

**Abbildung 6.4:** lokale Definition, lokale Variable

```
((empty? list) (make-no-result))
((cons? list)
 (define rest-min (list-min (rest list)))
 ...)))
```

Diese Definition wird jedesmal ausgewertet, wenn die Auswertung den **cond**-Zweig für **cons** erreicht und ist auch nur innerhalb dieses Zweiges aktiv. Wir ersetzen nun überall

```
(list-min (rest list))
```

durch die neue Variable **rest-min**:

```
(define list-min
  (lambda (list)
    (cond
      ((empty? list) (make-no-result))
      ((cons? list)
       (define rest-min (list-min (rest list)))
       (if (no-result? rest-min)
           (first list)
           (if (< (first list)
                 rest-min)
               (first list)
               rest-min))))))
```

Fertig!

Das heißt ...ein kleines bißchen verbessern können wir sie noch: Der **if**-Ausdruck am Ende berechnet das Minimum von **(first list)** und **rest-min**. Für die Berechnung des Minimums zweier Zahlen gibt es schon eine eingebaute Funktion namens **min**, die wir verwenden können:

```
(define list-min
  (lambda (list)
    (cond
      ((empty? list) (make-no-result))
```

```

((cons? list)
 (define rest-min (list-min (rest list)))
 (if (no-result? rest-min)
     (first list)
     (min (first list) rest-min)))))

```

Jetzt aber fertig!

## 6.8 FUNKTIONEN AUF NICHTLEEREN LISTEN

Das `list-min`-Beispiel war nützlich, um zu demonstrieren, wie wir «es gibt kein Ergebnis» repräsentieren können. Es gibt aber noch eine andere Möglichkeit, mit dem Problem der leeren Liste bei `list-min` fertigzuwerden: Wir verbieten sie einfach! Dann brauchen wir `no-result` gar nicht mehr.

Wir definieren also eine alternative Version von `list-min` mit Kurzbeschreibung und Signatur wie folgt:

```

; Minimum einer nichtleeren Liste von Zahlen berechnen
(: list-min-nonempty ((cons-list-of real) -> real))

```

Der Signaturkonstruktor `cons-list-of` entspricht dem gleichnamigen Signatur-Konstruktor aus Abschnitt 6.2 und beschreibt Cons-Listen – das sind ja genau die nichtleeren.

Wir benutzen die gleichen Testfälle wie vorhin:

```

(check-expect (list-min-nonempty (list 5 3 1 4)) 1)
(check-expect (list-min-nonempty (list 5 3 1 4 -4 3)) -4)

```

Die Schablone müssen wir gegenüber der für `list-of` allerdings etwas anpassen. Schließlich liefert für eine nichtleere Liste die Bedingung `(empty? list)` *immer* `#f` und `(empty? list)` *immer* `#t`, die Verzweigung damit wäre also sinnlos. Es ist stattdessen `(rest list)`, das entweder leer oder nichtleer sein kann. Das sieht dann so aus:

```

(define list-min-nonempty
  (lambda (list)
    (cond
      ((empty? (rest list)) ...)
      ((cons? (rest list))
       ... (first list) ...

```

```
... (list-min-nonempty (rest list)) ...)))))
```

Im ersten Zweig hat die Liste also genau ein Element, im zweiten Zweig hat sie mindestens zwei Elemente. Das Minimum einer Liste mit einem Element ist einfach dieses eine Element. Im zweiten Fall können wir, genau wie bei `list-min`, das Minimum von `(first list)` und dem Minimum des Rests der Liste berechnen:

```
(define list-min-nonempty
  (lambda (list)
    (cond
      ((empty? (rest list))
       (first list))
      ((cons? (rest list))
       (min (first list)
            (list-min-nonempty (rest list)))))))
```

Fertig!

#### AUFGABE 6.5

Kannst Du vorhersagen, was genau passiert, wenn Du `list-min-nonempty` trotzdem auf eine leere Liste anwendest? Probiere es aus! ☐

#### AUFGABE 6.6

Schreibe eine spezielle Konstruktionsanleitung für Funktionen, die nichtleere Listen akzeptieren! ☐

## ANMERKUNGEN

Listen sind eine Art Alleskleber: Sie taugen auch für die Repräsentation von Tabellen, Mengen und vielen anderen zusammengesetzten Daten. Für Listen gibt es eine riesige Anzahl praktischer Funktionen, von denen die Funktionen in diesem Kapitel nur die Spitze des Eisberges sind. Da ab hier Listen in den Lehrsprachen fest eingebaut sind, können sie als universelles Kommunikationsmittel zwischen Programmen dienen, weil sich Funktionen auf Listen aus einem Programm auch in einem anderen verwenden lassen. Dies unterscheidet funktionale Sprachen von vielen anderen Programmiersprachen, in denen Listen vom Programmierer selbst definiert werden müssen oder nur eine untergeordnete Rolle spielen.



## AUFGABEN

### AUFGABE 6.7

Schreibe Ausdrücke für Listen, welche die Beispiellisten vom Anfang von Kapitel 6 auf Seite 165 repräsentieren.

### AUFGABE 6.8

Schreibe folgende Funktionen auf Listen:

1. Eine Funktion `count-zeroes`, die die Anzahl von Nullen in einer Liste von Zahlen berechnet.
2. Eine Funktion `contains>10?`, die feststellt, ob eine Liste von Zahlen eine Zahl enthält, die größer als 10 ist.

### AUFGABE 6.9

Schreibe Versionen von `list-sum` und `list-product`, die nur auf nichtleeren Listen funktionieren und bei denen Du kein neutrales Element angeben musst. ☐

### AUFGABE 6.10

Auf einem Acker gibt es – vereinfacht dargestellt – Kartoffeln, Erdklumpen und Steine. Ein Ackerbestandteil ist also eine Kartoffel, ein Erdklumpen oder ein Stein. Jedes dieser Ackerbestandteile hat ein Gewicht in Gramm; Kartoffeln besitzen außerdem zusätzlich noch die Eigenschaft, ob sie essbar sind oder nicht (und daher aussortiert werden müssen). Schreibe ein Programm für eine Kartoffel-Erntemaschine, die die essbaren Kartoffeln aus den Ackerbestandteilen herausfiltern muss.

1. Führe eine Datenanalyse für Ackerbestandteile durch und schreibe passende Daten- und Record-Definitionen. Gib ein paar Beispiele für Ackerbestandteile an, die Du in den nächsten Teilaufgaben in Testfällen verwenden kannst.
2. Schreibe eine Funktion `edible-potato?`, die überprüft ob ein übergebener Ackerbestandteil eine essbare Kartoffel ist.
3. Schreibe eine Funktion `weight`, die das Gewicht eines beliebigen Ackerbestandteils zurückgibt.
4. Schreibe eine Daten- und Record-Definition für Listen von Ackerbestandteilen. Gib alle Signaturen an! Gib außerdem mindestens fünf verschiedene Beispiele für Listen mit Ackerbestandteilen an.

5. Schreibe eine Funktion `total-weight`, die eine Liste von Ackerbestandteilen akzeptiert und das Gesamtgewicht aller Ackerbestandteile in der Liste ausrechnet.
6. Schreibe eine Funktion `total-edible-potatoes-weight`, die eine Liste von Ackerbestandteilen akzeptiert und das Gesamtgewicht aller essbaren Kartoffeln in der Liste ausrechnet.
7. Schreibe eine Funktion `count`, die eine Liste von Ackerbestandteilen akzeptiert und die Anzahl der Ackerbestandteile in der Liste bestimmt.
8. Schreibe eine Funktion `count-edible-potatoes`, die eine Liste von Ackerbestandteilen akzeptiert und die Anzahl der essbaren Kartoffeln in der Liste bestimmt.
9. Eine Maschine soll nun die Kartoffeln aus dem Acker automatisch ernten. Leider ist die Maschine sehr alt und kommt deswegen mit großen Steinen und Erdklumpen nicht klar. Der Erntevorgang wird abgebrochen, sobald ein Stein oder ein Erdklumpen, der schwerer als 1kg ist, in die Maschine gerät. Schreibe die Funktion `harvester`, die diesen Vorgang simuliert. Als Eingabe bekommt diese Funktion eine Liste von Ackerbestandteilen, die sie nach und nach abarbeitet. Herauskommen soll eine Liste von Ackerbestandteilen, in der die geernteten essbaren Kartoffeln enthalten sind. Schreibe ausreichend Testfälle, beachte auch den Fall, dass die Maschine im Leerlauf betrieben wird. Verwende die Funktionen aus den vorherigen Aufgabenteilen.
10. Schreibe eine Funktion `heaviest-potato`, die aus einer Liste von Ackerbestandteilen die schwerste essbare Kartoffel heraussucht und zurückgibt. Wenn in der Liste keine essbare Kartoffel enthalten ist, soll die Funktion `(make-no-result)` zurückgeben. (Siehe Abschnitt 4.4 auf Seite 124.)
11. Schreibe eine Funktion `filter-edible-potatoes`, die eine Liste von Ackerbestandteilen akzeptiert und essbaren Kartoffeln als Liste zurückgibt.
12. Schreibe eine Funktion `drop-stones`, die eine Liste von Ackerbestandteilen akzeptiert und eine Liste zurückgibt, in der es keine Steine mehr gibt.
13. Schreibe eine Funktion `average-weight-edible-potatoes`, die eine Liste von Ackerbestandteilen akzeptiert und das durchschnittliche Gewicht der essbaren Kartoffeln berechnet.

#### AUFGABE 6.11

Wir repräsentieren einen Lithium-Ionen-Akku als eine Liste von Lithium-Ionen-Zellen. Eine Zelle besteht aus einer maximalen Ladung (in Milliamperestunden, mAh) und einer momentanen Ladung (ebenfalls in mAh).

Wenn eine Zelle eine Ladung von 1000mAh hat, bedeutet das, dass die Zelle eine Stromstärke von 1000mA (Milliampere) für eine Stunde lang liefern kann. Die momentane Ladung darf nie über die maximale Ladung steigen und nie unter 10% der maximalen Ladung fallen, in unserer Bei-

spielzelle also roomAh, sonst geht die Zelle kaputt.

1. Führe eine Datenanalyse für Li-Ionen-Akkus und Li-Ionen-Zellen durch, schreibe die Datendefinitionen auf und setze die Datendefinitionen um.
2. Schreibe folgende Funktionen:
  - `cell-full?`, die überprüft, ob eine Zelle vollständig geladen ist (das heißt ob die momentane Ladung gleich der maximalen Ladung ist)
  - `cell-empty?`, die überprüft, ob eine Zelle entladen ist (das heißt ob die momentane Ladung gleich der minimalen Ladung ist, also 10% der maximalen Ladung)
  - `cell-defect?`, die überprüft, ob eine Zelle defekt ist (das heißt ob die momentane Ladung die maximale Ladung überschreitet oder die minimale Ladung unterschreitet)
  - `cell-ok?`, die überprüft, ob eine Zelle funktioniert (das heißt ob die momentane Ladung innerhalb der minimalen und maximalen Ladung liegt)
3. Schreibe folgende Funktionen:
  - `battery-full?`, die überprüft, ob ein Akku vollständig geladen ist (das heißt ob alle Zellen voll sind); eine Batterie ohne Zelle gilt als geladen.
  - `battery-empty?`, die überprüft, ob ein Akku entladen ist (das heißt ob alle Zellen leer sind); eine Batterie ohne Zelle gilt auch als leer.
  - `battery-defect?`, die überprüft, ob ein Akku defekt ist (das heißt ob mindestens eine Zelle defekt ist); eine Batterie ohne Zelle gilt als nicht defekt.
  - `battery-ok?`, die überprüft, ob ein Akku funktioniert (das heißt ob alle Zellen funktionieren); eine Batterie ohne Zelle gilt als funktionstüchtig.

Gehe in den folgenden Teilaufgaben davon aus, dass nur *funktionierende* Zellen und Batterien übergeben werden:
4. Das Ladegerät kann eine Zelle um 500mAh pro Stunde aufladen. Schreibe eine Funktion `time-to-fully-charge-cell`, die ausrechnet, wieviele Stunden es dauert, bis das Ladegerät die Zelle aufgeladen hat.
5. Schreibe eine Funktion `charge-cell`, die eine Zelle und eine Zeit in Stunden akzeptiert und

die Zelle zurückgibt, die mit dem in der vorherigen Teilaufgabe genannten Ladegerät für die übergebene Zeit geladen wurde. Die Ladung darf aber nicht über die maximale Ladung steigen! Die restliche Zeit verfällt, wenn die Zelle bereits voll ist.

6. Schreibe eine Funktion `charge-battery`, die einen Akku auflädt. Die Funktion gibt einen Akku zurück, der mit dem Ladegerät für die übergebene Zeit geladen wurde. Ist eine Zelle vollständig geladen, so wird die nächste Zelle mit der restlichen Zeit geladen. Sind alle Zellen geladen und ist die Zeit jedoch noch nicht aufgebraucht, so verstreicht diese.
7. Das Entladen einer Zelle hängt vom Verbraucher ab: Ein Gerät hat einen Verbrauchswert, angegeben in mA. Anschaulich betrachtet heißt das, dass das Gerät pro Zeit einen gewissen Strom verbraucht. Schreibe eine Funktion `time-to-fully-discharge-cell`, die eine Zelle und einen Verbrauch pro Stunde akzeptiert und ausrechnet, wie lange es dauert, bis die Zelle entladen ist.
8. Schreibe eine Funktion `time-to-fully-discharge-battery`, die einen Akku und einen Verbrauch pro Stunde akzeptiert, und ausrechnet, wie lange es dauert, bis der Akku entladen ist.
9. Schreibe eine Funktion `discharge-cell`, die eine Zelle, einen Verbrauch pro Stunde und die Dauer (in Stunden) akzeptiert, für die der Verbraucher den Strom der Zelle verbraucht. Die Funktion soll eine Zelle zurückgeben, die für die Dauer den Verbraucher mit Strom versorgt hat. Die Ladung darf aber nicht unter die minimale Ladung fallen! Die restliche Zeit verfällt, wenn die Zelle bereits leer ist. (Tatsächlich geht dem Verbraucher einfach der Strom aus.)
10. Schreibe eine Funktion `discharge-battery`, die eine Batterie, einen Verbrauch pro Stunde und die Dauer (in Stunden) akzeptiert. Die Funktion soll eine Batterie zurückgeben, die für die Dauer den Verbraucher mit Strom versorgt hat. Ist die Ladung einer Zelle verbraucht, wird die Ladung der nächsten Zelle für die verbleibende Zeit verbraucht. Sind alle Zellen entladen und ist die Zeit jedoch noch nicht aufgebraucht, so verstreicht diese, dem Verbraucher geht

der Strom aus.

#### AUFGABE 6.12

Schreibe eine Funktion, die auf möglichst einfache Weise eine Liste von Zahlen aufsteigend sortiert! Gehe dazu wie folgt vor:

1. Schreibe eine Funktion `delete-once`, die aus einer Liste von Zahlen das erste Vorkommen einer gegebenen Zahl entfernt.  
Beispiel: `(delete-once 5 (list 1 2 5 3 4 5))` ergibt `#<list 1 2 3 4 5>`.  
(Wenn die Liste die Zahl nicht enthält, soll sie unverändert zurückgegeben werden.)
2. Schreibe eine Funktion `sort`, die eine Liste von Zahlen in aufsteigender Reihenfolge sortiert. Benutze dazu `list-min` auf Abschnitt 6.7 auf Seite 185 und `delete-once`.  
Beispiel: `(sort (list 1 5 3 2 4))` ergibt `#<list 1 2 3 4 5>`.

#### AUFGABE 6.13

In einer Firma stempeln die Mitarbeiter Stempelkarten, um ihre Arbeitszeit nachzuweisen: Auf jeder Stempelkarte ist die Nummer des Mitarbeiters vermerkt sowie eine Liste von jeweils gearbeiteten Stunden. Gehe der Einfachheit halber davon aus, dass nur ganze Stunden notiert werden. Außerdem gibt es für jeden Mitarbeiter eine Personalakte mit Name, Nummer und Stundenlohn (auch nur in ganzen Euros) des Mitarbeiters. Am Ende jedes Monats muss an jeden Mitarbeiter der Lohn überwiesen werden. Dazu werden Überweisungen erzeugt; jede Überweisung besteht aus der Mitarbeiternummer und einem Überweisungsbetrag.

Schreibe Daten- und Record-Definitionen für Stempelkarten, Personalakten und Überweisungen.

Schreibe eine Funktion, die eine Liste von Personalakten und eine Liste von Stempelkarten für einen Monat akzeptiert, für jeden Mitarbeiter die gearbeiteten Stunden aufsummiert und eine Liste der Überweisungen zurückliefert.

#### AUFGABE 6.14

Du bist Besitzer eines Parkplatzes mit Bereichen für unterschiedliche Fahrzeuge: PKWs (`car`), Lastwagen (`truck`), Wohnmobile (`caravan`), Busse (`bus`) und Fahrräder (`bike`). Wegen der Finanznot hat die Stadt Tübingen beschlossen, dass Du für jedes Fahrzeug, das auf Ihrem Parkplatz parkt, Abgaben entrichten musst, und zwar pro Tag je:

- `car`: 3 Cent

- **truck**: 5 Cent + 3 Cent pro Achse
- **caravan**: 4 Cent
- **bus**: 10 Cent + 1 Cent pro Sitzplatz
- **bike**: 1 Cent.

Der Fahrzeughalter kann die Abgaben nur für den jeweiligen Tag begleichen, so dass Du die Parkdauer nicht beachten musst.

1. Mache eine Datenanalyse und schreibe Daten- und Record-Definitionen für die Fahrzeugtypen **truck**, **bus** und **simple-vehicle**. Dabei werden unter **simple-vehicle** die Fahrzeuge zusammengefasst, bei denen keine variablen Kosten anfallen.
2. Schreibe Funktionen **tax-truck**, **tax-bus** und **tax-simple-vehicle** zur Ermittlung der Abgaben für diese Fahrzeugtypen
3. Definiere einen Datentyp **vehicle**. Schreibe eine Funktion **tax-vehicle**, die die Abgaben für ein beliebiges Fahrzeug berechnet.
4. Schreibe nun eine Funktion **calculate-tax**, die die täglichen Abgaben berechnet, die Du an das Finanzamt für alle Fahrzeuge auf ihrem Parkplatz bezahlen musst. Die Funktion akzeptiert eine Liste von Fahrzeugen und berechnet die Gesamtabgaben eines Tages.

Benutze bei jeder Funktion und jedem Record, den Du schreibst, die Konstruktionsanleitungen: Erstelle zunächst Kurzbeschreibung, Signatur, einige Testfälle und das Gerüst der Funktion. Vervollständige anschließend das Gerüst und vergewissere Dich, dass die Testfälle korrekt durchlaufen!

## AUFGABE 6.15

Da Du jetzt mit Listen programmieren kannst, bittet Dr. Knaubichler Dich nochmals um Ihre Hilfe (siehe Aufgabe 4.7 auf Seite 127): Schreibe ein Programm, das aus einer Liste von Grundkreaturen die zwei Grundkreaturen auswählt, die miteinander gekreuzt die leistungsstärkste Kreatur ergeben. Gehe dazu wie folgt vor:

1. Schreibe ein Prädikat für Kreaturen.
2. Schreibe eine Funktion **creature-power**, die eine beliebige Kreatur akzeptiert und deren Leistung berechnet. Die Leistung ist die Summe aller Eigenschaften.
3. Schreibe eine Funktion **find-optimal-breeding-partner**, die eine Grundkreatur  $c_1$  und eine Liste von Grundkreaturen akzeptiert. Die Funktion soll die Grundkreatur  $c_2$  aus der Liste auswählen und zurückgeben, die der beste Kreuzungspartner von  $c_1$  ist. Die Kreuzung von  $c_1$  und  $c_2$  soll also unter allen möglichen Kreuzungen die Kreatur mit der höchsten Leistung

ergeben. Falls die Liste der Grundkreaturen leer ist, soll die Funktion `(make-no-result)` zurückgeben. (Siehe Abschnitt 4.4 auf Seite 124.)

4. Schreibe eine Funktion `breed-optimal-partner`, die eine Grundkreatur  $c_1$  und eine Liste von Grundkreaturen akzeptiert. Die Funktion soll  $c_1$  mit der Grundkreatur aus der Liste kreuzen, die zur leistungsstärksten Kreatur führt. Falls die Liste der Grundkreaturen leer ist, soll die Funktion `(make-no-result)` zurückgeben. Benutze `find-optimal-breeding-partner`.
5. Schreibe eine Funktion `optimal-breed`, die eine Liste von Grundkreaturen akzeptiert. Die Funktion soll die beiden Grundkreaturen aus der Liste miteinander kreuzen, die unter allen möglichen Kreuzungen die Kreatur mit der höchsten Leistung ergeben. Benutze die Funktionen aus den vorherigen Teilaufgaben.

#### AUFGABE 6.16

Programmiere eine Variante der Funktion `list-min` in Abschnitt 6.7 auf Seite 185, in der Du statt `(make-no-result)` die Funktion `violation` aufrufst. Was musst Du noch alles ändern?





## 7 NATÜRLICHE ZAHLEN

In diesem Kapitel geht es um Funktionen, die etwas zählen und sich dementsprechend mit Zahlen beschäftigen. Wir haben schon eine Reihe von Programmen gesehen, die mit Zahlen rechnen, aber Zählen ist etwas besonderes und verdient deshalb ein eigenes Kapitel mit eigener Konstruktionsanleitung.

### 7.1 ZAHLEN, DIE ZÄHLEN

Du kennst bestimmt aus dem Mathematikunterricht die *Potenz* einer Zahl (der *Basis*)  $b$  zu einem *Exponenten*  $e$ :

$$b^e$$

Uns interessieren hier Exponenten, die ganze Zahlen ab 0 sind. Für solche Exponenten ist die Potenz folgendermaßen definiert:

$$b^e \stackrel{\text{def}}{=} \underbrace{b \times \dots \times b}_{e\text{-mal}}$$

Der Exponent  $e$  muss also eine natürliche Zahl sein.<sup>1</sup>

Umgangssprachlich sind natürliche Zahlen gerade die Zahlen, die Gegenstände zählen. In unserer Intuition sind die natürlichen Zahlen einfach auf einem Zahlenstrahl angeordnet und bilden eine Reihe. Für das Programmieren ist das aber wenig hilfreich. Viel mehr können wir mit der folgenden Datendefinition anfangen:

```
; Eine natürliche Zahl ist eine der folgenden:  
; - 0  
; - der Nachfolger einer natürlichen Zahl
```

In dieser Definition findet man alte Bekannte wieder: Es handelt sich um eine Fallunterscheidung, und der zweite Fall enthält einen Selbstbezug.

Dir mag der Begriff «Nachfolger» etwas befremdlich erscheinen: Der Nachfolger einer Zahl ist einfach nur die Zahl «plus 1». Weil der Nachfolger im Zusammenhang mit natürlichen Zahlen eine besondere Bedeutung hat, spendieren wir ihm zunächst eine eigene Hilfsfunktion:

---

<sup>1</sup> Spitzfindige Leser können anmerken, dass in manchen Lehrbüchern die natürlichen Zahlen mit 1 beginnen, aber wir halten uns an DIN 5473, und da ist die 0 mit dabei.

```
; Nachfolger einer Zahl
(: successor (natural -> natural))
```

```
(define successor
  (lambda (n)
    (+ n 1)))
```

Wo ein Nachfolger ist, muss es auch einen Vorgänger geben, und der zieht Eins ab:

```
; Vorgänger einer Zahl
(: predecessor (natural -> natural))
```

```
(define predecessor
  (lambda (n)
    (- n 1)))
```

Allerdings haben wir bei dieser Definition etwas geschummelt, weil der Vorgänger nur für die natürlichen Zahlen funktioniert, welche die Nachfolger anderer natürlicher Zahlen sind – also für sogenannte *positive* natürliche Zahlen. Die Signatur kann das nicht zum Ausdruck bringen, da es keine eingebaute Signatur dafür gibt. Aber wir können eine Verzweigung entsprechend der Datendefinition einbauen, um die Funktion vor Missbrauch zu schützen:

```
(define predecessor
  (lambda (n)
    (cond
      ((zero? n)
       (violation "0 does not have a predecessor"))
      ((positive? n)
       (- n 1)))))
```

Du siehst in der Funktionsdefinition die beiden eingebauten Funktionen `zero?` und `positive?`, welche die beiden Fälle der Datendefinition der natürlichen Zahlen unterscheiden.

Zurück zur eigentlichen Aufgabe, der Potenz. So könnten Kurzbeschreibung, Signatur und zwei Tests aussehen:

```
; Potenz einer Zahl berechnen
(: power (number natural -> number))
```

```
(check-expect (power 5 0) 1)
(check-expect (power 5 3) 125)
```

Als Nächstes ist das Gerüst dran:

```
(define power
  (lambda (base exponent)
    ...))
```

Für den Rumpf können wir eine Schablone aus der Datendefinition für `exponent` ableiten, das ist ja die natürliche Zahl. Die Datendefinition für natürliche Zahlen ist eine Fallunterscheidung mit zwei Fällen, dementsprechend brauchen wir eine Verzweigung mit zwei Zweigen:

```
(define power
  (lambda (base exponent)
    (cond
      ((zero? exponent) ...)
      ((positive? exponent) ...))))
```

Im zweiten Fall der Datendefinition steckt außerdem eine Selbstreferenz, wir schreiben also noch einen rekursiven Aufruf in die Schablone:

```
(define power
  (lambda (base exponent)
    (cond
      ((zero? exponent) ...)
      ((positive? exponent)
       ... (power base (predecessor exponent)) ...))))
```

Nun müssen wir noch die Lücken füllen:

Beim ersten Fall – Exponent 0 – muss als Potenz 1 herauskommen, das neutrale Element bezüglich der Multiplikation: Das ist ja nichts anderes, als die Zahlen der leeren Liste zu multiplizieren. (Siehe dazu auch Seite 172 in Abschnitt 6.1.)

Für den anderen Fall ist das Ergebnis des rekursiven Aufruf die Potenz von `base` mit dem Vorgänger von `exponent` als Exponent. Damit wir als Ergebnis `base` mit sich selbst `exponent`-mal multipliziert bekommen, fehlt noch eine Multiplikation mit `base`:

```
(define power
  (lambda (base exponent)
    (cond
      ((zero? exponent) 1)
      ((positive? base)
       (* base
```

```
(power base (predecessor exponent))))))
```

Fertig!

Der Begriff des Vorgängers ist hier eher im Weg: Er passt zwar gut zur Verwendung des Worts «Nachfolger» in der Datendefinition. (`predecessor exponent`) ist aber hier das gleiche ist wie (`- exponent 1`), und entsprechend werden wir auch in der Schablone direkt (`- ... 1`) statt `predecessor` verwenden.

#### AUFGABE 7.1

Die Funktion `predecessor` hat zwei Zweige: Warum ist der Aufruf in `power` immer gleichbedeutend zu (`- exponent 1`) – was ist mit dem anderen Zweig?

#### KONSTRUKTIONSANLEITUNG 19 (NATÜRLICHE ZAHLEN ALS EINGABE: SCHABLONE)

Eine Schablone für eine Funktion, die eine natürliche Zahl akzeptiert, sieht folgendermaßen aus:

```
(define f
  (lambda (... n ...)
    (cond
      ((zero? n) ...)
      ((positive? n)
       ...
       (f ... (- n 1) ...)
       ...
      )))
```

## 7.2 NATÜRLICHE ZAHLEN UND LISTEN

Zwischen natürlichen Zahlen und Listen besteht eine enge Beziehung: Die Datendefinition für natürliche Zahlen entspricht in ihrer Struktur der von Listen. Die Funktion `list-length` in Abschnitt 6.5 auf Seite 177 macht aus einer Liste eine natürliche Zahl. Das geht auch umgekehrt:

```
; Liste aus Kopien eines Werts erzeugen
(: copies (natural %element -> (list-of %element)))
```

```
(check-expect (copies 4 23)
  (list 23 23 23 23))
(check-expect (copies 5 "Mike")
  (list "Mike" "Mike" "Mike" "Mike" "Mike"))
```

Die Signatur der Funktion enthält die Signaturvariable `%element`, ist also polymorph und macht klar, dass die Elemente der Liste nur zur Signatur `%element` gehören können.

```
(define copies
  (lambda (count element)
    ...))
```

Die Konstruktionsanleitung schlägt folgende Schablone vor:

```
(define copies
  (lambda (count element)
    (cond
      ((zero? count) ...)
      ((positive? count)
       ... (copies (- count 1) element) ...))))
```

Für den ersten Fall brauchen wir eine Liste mit 0 Elementen. Im zweiten Fall liefert der rekursive Aufruf eine Liste mit einem Element weniger, als wir brauchen. Das ergänzen wir zur fertigen Definition so:

```
(define copies
  (lambda (count element)
    (cond
      ((zero? count)
       empty)
      ((positive? count)
       (cons element
              (copies (- count 1) element))))))
```

Eine weitere nützliche Funktion akzeptiert die Nummer eines Listenelementes – einen sogenannten *Index* – und liefert das Element:

```
; Nummeriertes Element aus einer Liste holen
(: nth ((list-of %element) natural -> %element))
```

```
(check-expect (nth (list 1 2 3 4 5) 0) 1)
(check-expect (nth (list 1 2 3 4 5) 2) 3)
```

Wir sollten uns noch überlegen, was passieren soll, wenn es den Index in der Liste gar nicht gibt. Die Funktion kann kein sinnvolles Ergebnis liefern, das zur Signatur passt – es muss ja ein Element der Liste herauskommen. Die einzig sinnvolle Möglichkeit ist, einen Fehler anzuzeigen. Daraus wird folgender Testfall:

```
(check-error (nth (list 1 2 3 4 5) 5))
```

Das Gerüst der Funktion sieht folgendermaßen aus:

```
(define nth
  (lambda (list index)
    ...))
```

Bei der Schablone haben wir die Wahl: Wir können entweder die Schablone für Listen als Eingabe benutzen oder die für natürliche Zahlen. Oder beide. Hier ist aber der Index federführend – wenn die Funktion das zweite Element einer tausendelementigen Liste extrahiert, ist die Länge irrelevant. Wir halten uns also an die Konstruktionsanleitung für natürliche Zahlen:

```
(define nth
  (lambda (list index)
    (cond
      ((zero? index) ...)
      ((positive? index)
       ... (nth (rest list) (- index 1)) ...))))
```

Im ersten Fall ist der Index 0, also das erste Element der Liste gefragt. Im zweiten Fall liefert uns der rekursive Aufruf im Rest der Liste das Element das Element an Stelle `(- index 1)`. Dies ist schon das `index`-te Element in `list`. Die Funktion sieht so aus:

```
(define nth
  (lambda (list index)
    (cond
      ((zero? index) (first list))
      ((positive? index)
       (nth (rest list) (- index 1))))))
```

Tatsächlich laufen alle Testfälle bereits durch – und das, obwohl wir für den Fehlerfall gar nichts programmiert haben.

## AUFGABE 7.2

Werte mit dieser Definition folgenden Ausdruck aus:

```
(nth (list 1 2 3 4 5) 5)
```

Die Fehlermeldung ist zwar technisch richtig, gibt aber keinen Aufschluss auf die Ursache des Fehlers. Das können wir korrigieren, indem wir doch noch die Schablone für Listen als Eingabe ergänzen und `violation` benutzen. (Siehe Abschnitt 2.10 auf Seite 74.)

```
(define nth
  (lambda (list index)
    (cond
      ((empty? list)
       (violation "nth: Die Liste ist nicht lang genug"))
      ((cons? list)
       (cond
         ((zero? index) (first list))
         ((positive? index)
          (nth (rest list) (- index 1))))))))
```

Fertig!

## 7.3 ANDERSRUM ZÄHLEN

Wir programmieren noch ein weiteres Beispiel: Zwei Zahlen bestimmen Unter- und Obergrenze eines Bereichs ganzer Zahlen und wir wollen die Liste dieser Zahlen generieren. Die Zahlen zwischen 3 und 7 sind zum Beispiel 3, 4, 5, 6, 7.

Hier sind Kurzbeschreibung, Signatur und ein Test:

```
; Liste aufeinanderfolgender ganzer Zahlen berechnen
(: between (integer integer -> (list-of integer)))

(check-expect (between 3 7) (list 3 4 5 6 7))
```

Hier ist das Gerüst für die Funktion:

```
(define between
  (lambda (from to)
    ...))
```

Es ist gar nicht so klar, was für eine Schablone wir anwenden sollten, um die Funktion zu vervollständigen: In der Signatur steht `integer`, die Eingaben können also auch negative Zahlen sein, mithin keine natürlichen Zahlen. Die Konstruktionsanleitung für natürliche Zahlen ist also nicht direkt anwendbar. Allerdings geht es bei der Funktion doch augenscheinlich ums Zählen, irgendwie sind also doch natürliche Zahlen im Spiel.

Wir können uns der Lösung nähern, indem wir zählen, wieviele Elemente die Ausgabe-Liste hat. Am Testfall sind es fünf Elemente, allgemein ist das eins mehr als die Differenz der beiden Eingabezahlen – und das ist immer eine natürliche Zahl.

Wir können die Aufgabe also leicht umformulieren: Nicht «die Zahlen zwischen 3 und 7» generieren, sondern «die 5 Zahlen ab 3». Dafür können wir eine eigene Funktion schreiben:

```
; Aufeinanderfolgende Zahlen ab einer Zahl generieren
(: count-from (integer natural -> (list-of integer)))
```

```
(check-expect (count-from 3 5) (list 3 4 5 6 7))
```

Das Gerüst ist wie folgt:

```
(define count-from
  (lambda (from count)
    ...))
```

Da `count` eine natürliche Zahl ist, können wir die entsprechende Schablone anwenden:

```
(define count-from
  (lambda (from count)
    (cond
      ((zero? count) ...)
      ((positive? count)
       ... (count-from ... (- count 1)) ...))))
```

Im ersten Fall sind null Elemente gefragt, da muss das Ergebnis `empty` sein. Da eine Liste herauskommt, werden wir wahrscheinlich im zweiten Fall `cons` benutzen, um diese zu konstruieren. Außerdem ist das gewünschte erste Element gerade `from`. Wir können die Schablone also wie folgt weiterentwickeln:



```
(define count-from
  (lambda (from count)
    (cond
      ((zero? count) empty)
      ((positive? count)
       (cons from
              ... (count-from ... (- count 1)) ...))))))
```

Für den Rest der Ausgabe-Liste brauchen wir eine Liste, die mit der nächsten Zahl anfängt, also mit `(+ from 1)`. Wenn wir das als erstes Argument des rekursiven Aufrufs benutzen, ist dieser schon genau die benötigte Rest-Liste:

```
(define count-from
  (lambda (from count)
    (cond
      ((zero? count) empty)
      ((positive? count)
       (cons from
              (count-from (+ from 1) (- count 1)))))))
```

Nun ist `count-from` fertig, aber unsere eigentliche Funktion `between` noch nicht. Wir können aber `count-from` benutzen, um `between` zu realisieren:

```
(define between
  (lambda (from to)
    (count-from from (+ (- to from) 1))))
```

Fertig!

Vielleicht erscheint Dir umständlich, die Aufgabe erst einmal umzumodeln auf eine Aufgabe über natürliche Zahlen und dann die zu bearbeiten. Geht das nicht «direkt»? Es geht.

Der Schlüssel zur direkten Lösung ist die Einsicht, dass die natürliche Zahl in der Aufgabe gerade die Differenz zwischen `from` und `to` ist. Wir können eine entsprechende Verzweigung so formulieren:

```
(define between
  (lambda (from to)
    (cond
      ((= from to) ...)
      ((< from to)
       ... (between ... ...) ...))))
```

Beim ersten Zweig brauchen wir eine einelementige Liste. Im zweiten ist es wieder sinnvoll, zunächst das `cons` hinzuschreiben, weil wir wissen, dass das erste Element `from` ist:

```
(define between
  (lambda (from to)
    (cond
      ((= from to) ...)
      ((< from to)
       (cons from (between ... ...))))))
```

Nun können wir die Argumente des rekursiven Aufrufs ergänzen. Wir benötigen eine Liste, die mit der nächsten Zahl anfängt und genau wie gehabt aufhört:

```
(define between
  (lambda (from to)
    (cond
      ((= from to) (list from))
      ((< from to)
       (cons from (between (+ from 1) to))))))
```

Fertig!

Vielleicht fällt Dir bei dieser Aufgabe auf, dass es leicht passieren kann, dass die ganzen Zahlen «um eins verrutschen». Man könnte zum Beispiel leicht auf die Idee kommen, im ersten Zweig statt `(list from)` einfach `empty` hinzuschreiben. Das passiert auch erfahrener Programmierinnen und Programmierern. Es ist deswegen sinnvoll, beim Programmieren von Funktionen über natürliche und ganze Zahlen genau hinzuschauen, ob sie korrekt sind.

## 7.4 EXKURS: POTENZ OPTIMIEREN

Wir kehren in diesem Abschnitt noch einmal zur Potenz zurück. Mit einem Trick können wir nämlich den Rechenaufwand reduzieren: Zum Beispiel können wir  $3^{10}$  folgendermaßen berechnen:

$$3^{10} = 3^5 \times 3^5$$

Wir müssen also nicht mühsam die 3 zehnmal mit sich selbst multiplizieren, es reicht,  $3^5$  zu berechnen und mit sich selbst zu multiplizieren. Bei ungeraden Exponenten geht der Trick auch, wir

müssen aber einmal extra mit der Basis multiplizieren:

$$3^9 = 3^4 \times 3^4 \times 3$$

Diesen Trick können wir auch zu einem Programm machen. Wir nennen die Funktion der Einfachheit halber `power2`, die aber ansonsten genau wie `power` funktionieren soll. Hier ist das Gerüst:

```
(define power2
  (lambda (base exponent)
    ...))
```

In `power2` wird nicht mehr direkt gezählt, obwohl der Exponent immer noch eine natürliche Zahl ist. Weil nicht gezählt wird, können wir die Schablone für natürliche Zahlen (wieder) nicht direkt verwenden. Allerdings können wir eine alternative Datendefinition für natürliche Zahlen unterstellen, die Zahlen halbiert, statt sich (wie die erste Definition) auf den Vorgänger zu beziehen.

```
; Eine natürliche Zahl ist:
; - 0
; - das Doppelte einer natürlichen Zahl
; - der Nachfolger des Doppelten einer natürlichen Zahl
```

Da die Datendefinition diesmal drei Fälle hat, muss der Rumpf aus einer Verzweigung mit drei Zweigen bestehen:

```
(define power2
  (lambda (base exponent)
    (cond
      ((zero? exponent) ...)
      ((even? exponent) ...)
      ((odd? exponent) ...))))
```

Der erste Fall entspricht der ursprünglichen `power`-Funktion – da muss eine 1 hin. Der zweite und der dritte Fall enthalten jeweils einen Selbstbezug, weswegen dort rekursive Aufrufe stehen müssen. Im zweiten Fall steht «das Doppelte», wir können also die natürliche Zahl des Selbstbezugs durch Halbierung berechnen. Im dritten Fall müssen wir vor dem Halbieren noch eins abziehen. Wir benutzen jeweils `quotient`, weil wir ganzzahlig halbieren:

```
(define power2
  (lambda (base exponent)
    (cond
```

```

((zero? exponent) 1)
((even? exponent)
 ... (power2 base (quotient exponent 2)) ...)
((odd? exponent)
 ... (power2 base (quotient (- exponent 1) 2)) ...)))

```

Wir kümmern uns nun um den zweiten Fall: Wenn dort  $b$  die Basis und  $e$  der Exponent sind, dann kommt beim rekursiven Aufruf heraus:

$$b^{\frac{e}{2}}$$

Gesucht ist

$$b^e = b^{\frac{e}{2}} \times b^{\frac{e}{2}}.$$

Dafür müssen wir  $b^{\frac{e}{2}}$  mit sich selbst multiplizieren. Damit diese Zahl nicht zweimal berechnet wird, legen wir dafür eine lokale Definition an:

```

(define power2
  (lambda (base exponent)
    (cond
      ((zero? exponent) 1)
      ((even? exponent)
       (define half (power2 base (quotient exponent 2)))
       (* half half))
      ((odd? exponent)
       ... (power2 base (quotient (- exponent 1) 2)) ...))))

```

Im dritten Fall gilt:

$$b^e = b^{\frac{e-1}{2}} \times b^{\frac{e-1}{2}} \times b.$$

Entsprechend können wir den letzten Zweig ausfüllen:

```

(define power2
  (lambda (base exponent)
    (cond
      ((zero? exponent) 1)
      ((even? exponent)
       (define half (power2 base (quotient exponent 2)))
       (* half half))
      ((odd? exponent)
       (define half (power2 base (quotient (- exponent 1) 2)))

```

```
(* half half base))))
```

Fertig!

### AUFGABE 7.3

Überzeuge Dich mit dem Stepper, dass `power2` weniger Multiplikationen benötigt als `power`!

## AUFGABEN

### AUFGABE 7.4

Die *Fakultät* einer Zahl  $n$  ist definiert als:

$$n! \stackrel{\text{def}}{=} 1 \times 2 \times \cdots \times n$$

Schreiben Sie eine Funktion, welche die Fakultät berechnet!

### AUFGABE 7.5

Programmieren Sie folgende Funktionen auf Listen:

1. Programmieren Sie eine Funktion `take`, die als Argumente eine Liste `l` und eine Zahl `n` akzeptiert, und die ersten `n` Elemente der Liste `l` zurückgibt. Beispiel:

```
(take (list 1 2 3 4 5 6 7 8) 5)
↪ #<list 1 2 3 4 5>
```

Gehen Sie davon aus, dass `l` mindestens `n` Elemente lang ist.

2. Programmieren Sie eine Funktion `drop`, die als Argumente eine Liste `l` und eine Zahl `n` akzeptiert, und die Liste `l` ohne die ersten `n` Elemente zurückgibt. Beispiel:

```
(drop (list 1 2 3 4 5 6 7 8) 3)
↪ #<list 4 5 6 7 8>
```

Gehen Sie davon aus, dass `l` mindestens `n` Elemente lang ist.

### AUFGABE 7.6

- Die eingebaute Funktion `even?` akzeptiert eine ganze Zahl und liefert `#t`, falls diese gerade ist und `#f` wenn nicht. Schreiben Sie mit Hilfe von `even?` eine Funktion namens `evens`, welche für zwei Zahlen  $a$  und  $b$  eine Liste der geraden Zahlen zwischen  $a$  und  $b$  zurückgibt:

```
(evens 1 10)
```

```
⇒ #<record:pair 2  
    #<record:pair 4  
    #<record:pair 6  
    #<record:pair 8  
    #<record:pair 10 #<empty-list>>>>>>
```

- Die eingebaute Funktion `odd?` akzeptiert eine ganze Zahl und liefert `#t`, falls diese ungerade ist und `#f` wenn nicht. Schreiben Sie mit Hilfe von `odd?` eine Funktion `odds`, welche für zwei Zahlen  $a$  und  $b$  eine Liste der ungeraden Zahlen zwischen  $a$  und  $b$  zurückgibt:

```
(odds 1 10)
```

```
⇒ #<record:pair 1  
    #<record:pair 3  
    #<record:pair 5  
    #<record:pair 7  
    #<record:pair 9 #<empty-list>>>>>>
```

## 8 EXKURS: INDUKTIVE BEWEISE UND DEFINITIONEN

In den vergangenen Kapiteln haben wir uns oft mit Selbstbezügen in Datendefinitionen beschäftigt und den rekursiven Funktionen, die sie akzeptieren. Dieses Kapitel beschäftigt sich mit der mathematischen Seite solcher Funktionen. Es ist für Dich als Lektüre sinnvoll, wenn Du Dich besonders für Mathematik interessierst oder studiumshalber dafür interessieren musst. Für das Verständnis der folgenden Kapitel kannst Du es aber auch weglassen.

Die Mathematik bildet Daten mit Selbstbezug als sogenannte *induktiv definierte* Mengen ab. Dies ermöglicht uns, dass wir uns von der Korrektheit einer Funktion mit Hilfe eines sogenannten induktiven Beweises überzeugen. Dieses Kapitel zeigt, was das ist und wie das geht.

### 8.1 AUSSAGEN ÜBER NATÜRLICHE ZAHLEN

In der Mathematik gibt es viele kuriose Aussagen über natürliche Zahlen. Legendenstatus hat z.B. die Gaußsche Summenformel. Diese besagt, dass für eine natürliche Zahl  $n$  die Summe aller Zahlen von 1 bis  $n$  mit der Formel

$$\frac{n \times (n + 1)}{2}$$

berechnet werden kann: Man muss also gar nicht mühsam die Zahlen einzeln addieren. Mathematiker lieben spezielle Schreibweisen und könnten Gauß' Behauptung folgendermaßen aufschreiben:

$$\forall n \in \mathbb{N} : \sum_{i=0}^n i = \frac{n \times (n + 1)}{2}$$

Das  $\forall$  heißt «für alle», das  $\in$  heißt «Element von»  $\mathbb{N}$  ist das Kürzel für die Menge der natürlichen Zahlen. Am Anfang steht also «für alle  $n$ , die Element von  $\mathbb{N}$  sind» oder ganz lapidar «für alle natürlichen Zahlen  $n$ ».

Das  $\sum$  ist ein griechisches Sigma und steht für «Summe». Das  $i = 0$  untendrunter und das  $n$  obendrüber bedeuten, dass eine neue Variable  $i$  eingeführt wird, und die nacheinander die Werte  $0, 1, 2, \dots, n$  annimmt, und für jeden Wert in die Formel rechts vom  $\sum$  eingesetzt wird. In der Gaußschen Summenformel steht rechts nur  $i$ , das heißt die Summe ist  $0 + 1 + 2 + \dots + n$ .

Wie können wir die Gaußsche Summenformel beweisen? Gauß' Argument war, dass, wenn er

die Summe ausschreibt (wir lassen die 0 weg):

$$1 + 2 + \dots + (n - 1) + n$$

... die beiden «äußeren» Summanden 1 und  $n$  zusammen  $n + 1$  ergeben, die beiden «nächstinne-  
ren» Summanden 2 und  $n - 1$  ebenfalls  $n + 1$  und so weiter bis zur «Mitte» – effektiv also halb so oft  
 $n + 1$  auf sich selbst addiert wird wie die Reihe selbst lang ist. Es ist also einfach nachzuvollziehen,  
wie Gauß auf die Formel kam und warum sie korrekt ist.

Was ist aber mit folgender Reihe?

$$\sum_{i=0}^n i^2$$

Der Blick auf die ausgeschriebene Form hilft nicht direkt weiter:

$$1 + 4 + 9 + 16 + \dots + (n - 1)^2 + n^2$$

Allerdings lohnt es sich, einen Blick auf die ersten paar Glieder der Reihe zu werfen, und diese  
tabellarisch über die Gaußsche Summen zu setzen:

$n$	0	1	2	3	4	5	6	...
$\sum_{i=0}^n i^2$	0	1	5	14	30	55	91	...
$\sum_{i=0}^n i$	0	1	3	6	10	15	21	...

Wenn Du die Paare der Summen der beiden Reihen lang genug anstartest, siehst Du vielleicht, dass  
sich alle als Brüche auf den Nenner 3 kürzen lassen:

$n$	0	1	2	3	4	5	6	...
$\frac{\sum_{i=0}^n i^2}{\sum_{i=0}^n i}$	?	$\frac{3}{3}$	$\frac{5}{3}$	$\frac{7}{3}$	$\frac{9}{3}$	$\frac{11}{3}$	$\frac{13}{3}$	...

Einzige Ausnahme ist der Bruch für 0: dort wird durch 0 geteilt, es ist also unklar, welcher Bruch  
an dieser Stelle in der Tabelle stehen sollte. Ansonsten suggeriert die Tabelle folgende Formel:

$$\frac{\sum_{i=0}^n i^2}{\sum_{i=0}^n i} = \frac{2n + 1}{3}$$

Die Gleichung kann mit  $\sum_{i=0}^n i = n(n + 1)/2$  multipliziert werden, um eine Antwort für die



ursprüngliche Frage zu ergeben:

$$\sum_{i=0}^n i^2 = \frac{n(n+1)(2n+1)}{6} \quad (8.1)$$

Schöne Formel – aber stimmt sie auch für alle  $n \in \mathbb{N}$ ? (Für den unklaren Fall 0 stimmt sie.) Falls Du magst, kannst Du sie noch für weitere  $n$ , zum Beispiel 7, 8, ... – ausprobieren, und tatsächlich zeigen sich zunächst keine Gegenbeispiele. Aber das ist langweilig und würde immer noch nicht reichen, um die Behauptung für *alle*  $n \in \mathbb{N}$  zu beweisen.

Wenn die Behauptung für alle  $n \in \mathbb{N}$  stimmt, also insbesondere auch für ein bestimmtes  $n \in \mathbb{N}$ , dann sollte sie auch für  $n+1$  gelten, das wäre dann die folgende Gleichung, bei der gegenüber der Gleichung oben für  $n$  jeweils  $n+1$  eingesetzt wurde:

$$\sum_{i=0}^{n+1} i^2 = \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6}$$

Das können wir etwas vereinfachen:

$$\sum_{i=0}^{n+1} i^2 = \frac{(n+1)(n+2)(2n+3)}{6} \quad (8.2)$$

Bei der Reihe  $\sum_{i=0}^{n+1} i^2$  können wir den letzten Summand ausgliedern und die Gleichung damit folgendermaßen schreiben:

$$\left(\sum_{i=0}^n i^2\right) + (n+1)^2 = \frac{(n+1)(n+2)(2n+3)}{6}$$

Damit bietet sich die Chance, die jeweiligen Seiten von Gleichung 8.1 von den Seiten von Gleichung 8.2 abzuziehen:

$$\left(\sum_{i=0}^n i^2\right) + (n+1)^2 - \sum_{i=0}^n i^2 = \frac{(n+1)(n+2)(2n+3)}{6} - \frac{n(n+1)(2n+1)}{6}$$

Es bleibt:

$$(n+1)^2 = \frac{(n+1)(n+2)(2n+3)}{6} - \frac{n(n+1)(2n+1)}{6}$$

Wenn diese Gleichung stimmt, dann stimmt auch Gleichung 8.2. Das lässt sich ausrechnen:

$$\begin{aligned}
 \frac{(n+1)(n+2)(2n+3)}{6} - \frac{n(n+1)(2n+1)}{6} &= \frac{(n+1)(n+2)(2n+3) - n(n+1)(2n+1)}{6} \\
 &= \frac{(n+1)((n+2)(2n+3) - n(2n+1))}{6} \\
 &= \frac{(n+1)(2n^2 + 3n + 4n + 6 - 2n^2 - n)}{6} \\
 &= \frac{(n+1)(6n+6)}{6} \\
 &= \frac{6(n+1)^2}{6} \\
 &= (n+1)^2
 \end{aligned}$$

Tatsache! Aber was wurde jetzt eigentlich gezeigt? Es ist leicht, bei den vielen Schritten von oben den Faden zu verlieren. Hier noch einmal die Zusammenfassung:

Es *schien* so, als ob folgende Gleichung stimmen würde:

$$\sum_{i=0}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

Es *ist* so, dass folgende Gleichung stimmt:

$$(n+1)^2 = \frac{(n+1)(n+2)(2n+3)}{6} - \frac{n(n+1)(2n+1)}{6}$$

Wenn also die Gleichung, die so scheint, stimmen würde, dann folgte daraus diese Gleichung:

$$\sum_{i=0}^{n+1} i^2 = \frac{(n+1)(n+2)(2n+3)}{6}$$

Das ist aber die gleiche Behauptung, nur für  $n+1$  statt  $n$  – mit anderen Worten folgt aus der Behauptung für  $n$  die Behauptung für  $n+1$ . Da wir oben durch Ausrechnen bereits gezeigt haben, dass die Behauptung für  $1, \dots, 6$  gilt, gilt sie auch für  $7$ . Da sie für  $7$  gilt, gilt sie auch für  $8$ . Und so weiter für alle natürlichen Zahlen. Wir müssen das nicht für jede Zahl einzeln auszuprobieren. Die Vermutung von oben ist deshalb bewiesen.

## 8.2 INDUKTIVE BEWEISE FÜHREN

Das im vorigen Abschnitt verwendete Beweisprinzip für Beweise von Sätzen der Form «für alle  $n \in \mathbb{N} \dots$ » heißt *vollständige Induktion*.<sup>1</sup> Diese funktioniert bei Aufgaben, bei denen eine Behauptung für alle  $n \in \mathbb{N}$  zu beweisen ist. Für den Beweis werden folgende Zutaten benötigt:

1. ein Beweis dafür, dass die Behauptung für  $n = 0$  stimmt und
2. ein Beweis dafür, dass, wenn die Behauptung für ein beliebiges  $n$  gilt, sie auch für  $n + 1$  gilt.

Die erste Zutat (der *Induktionsanfang*) lässt sich meist durch Einsetzen beweisen, für die zweite Zutat (den *Induktionsschluss*) ist in der Regel Algebra nötig. Da die Behauptung nach Zutat Nr. 1 für 0 gilt, muss sie nach Zutat Nr. 2 auch für 1 gelten, und deshalb auch für 2,  $\dots$  und damit für alle natürlichen Zahlen. (Es muss nicht unbedingt bei 0 losgehen, sondern kann auch bei einer beliebigen Zahl  $a$  losgehen – dann gilt aber die Behauptung nur für alle natürlichen Zahlen ab  $a$ .)

Wenn die Behauptung erst einmal formuliert ist, sind induktive Beweise oft einfach zu führen, da sie meist dem obigen Schema folgen. Das wichtigste dabei ist, das Schema auch tatsächlich einzuhalten. Darum empfiehlt es sich, dass Du folgende Anleitung befolgst – eine Art Konstruktionsanleitung für Beweise mit vollständiger Induktion:

1. Formuliere die zu beweisende Behauptung als Behauptung der Form «Für alle  $n \in \mathbb{N}$  gilt  $\dots$ », falls das noch nicht geschehen ist.
2. Schreibe die Überschrift « $n = 0$ :». Schreibe die Behauptung noch einmal ab, wobei Du das «für alle  $n \in \mathbb{N}$ » weglassen und für  $n$  die 0 einsetzt.
3. Beweise die abgeschriebene Behauptung. (Das ist oft einfach.)
4. Schreibe das Wort «Induktionsvoraussetzung:» Schreibe darunter die Behauptung noch einmal ab, wobei Du das «für alle  $n \in \mathbb{N}$ » weglässt – lass das  $n$  da, wo es ist.
5. Schreibe «Induktionsschluss (zu zeigen):». Schreibe darunter die Behauptung noch einmal ab, wobei Du für  $n$  stattdessen  $(n + 1)$  einsetzt. (Das «für alle  $n \in \mathbb{N}$ » wieder weg.)
6. Beweise den Induktionsschluss unter Verwendung der Induktionsvoraussetzung.

Wenn die Behauptung eine Gleichung der Form  $A = B$  ist, kannst Du häufig die Induktionsvoraussetzung entweder direkt oder nach einigen Umformungen in den Induktionsschluss einsetzen.

Wir gehen die Anleitung anhand des obigen Beispiels durch. Die Behauptung ist:

$$\sum_{i=0}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

<sup>1</sup> Nicht zu verwechseln mit der *philosophischen Induktion*. Die vollständige Induktion ist zwar verwandt, philosophisch gesehen aber eher eine deduktive Technik.

Die Behauptung ist durch Raten entstanden – es gibt leider keine Patentanleitung, solche Gleichungen zu finden: Hier sind Experimentierfreude und Geduld gefragt. Steht die Behauptung erst einmal fest, ist sie allerdings recht einfach zu beweisen:

1. Ausgeschrieben hat die Behauptung bereits die richtige Form:

$$\forall n \in \mathbb{N} : \sum_{i=0}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

2.  $n = 0$ :

$$\sum_{i=0}^0 i^2 = \frac{0(0+1)(2 \times 0 + 1)}{6}$$

3. Beide Seiten der Gleichung sind 0, sie ist also bewiesen.  
4. Induktionsvoraussetzung:

$$\sum_{i=0}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

5. Induktionsschluss (zu zeigen):

$$\sum_{i=0}^{n+1} i^2 = \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6}$$

6. Die Summe auf der linken Seite können wir aufteilen:

$$\left(\sum_{i=0}^n i^2\right) + (n+1)^2 = \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6}$$

Damit können wir die Induktionsvoraussetzung einsetzen, in der  $\sum_{i=0}^n i^2$  auf der linken Seite steht – dieser Kniff funktioniert fast immer bei Induktionsbeweisen für Gleichungen über Summen oder Produkte. Es entsteht folgende Gleichung:

$$\left(\frac{n(n+1)(2n+1)}{6}\right) + (n+1)^2 = \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6}$$

Die linke Seite können wir folgendermaßen vereinfachen:

$$\begin{aligned}
 \frac{n(n+1)(2n+1)}{6} + (n+1)^2 &= \frac{n(n+1)(2n+1) + 6(n+1)^2}{6} \\
 &= \frac{(n+1)(n(2n+1) + 6(n+1))}{6} \\
 &= \frac{(n+1)(2n^2 + n + 6n + 6)}{6} \\
 &= \frac{(n+1)(2n^2 + 7n + 6)}{6}
 \end{aligned}$$

Die rechte Seite können wir folgendermaßen vereinfachen:

$$\begin{aligned}
 \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6} &= \frac{(n+1)(n+2)(2(n+1)+1)}{6} \\
 &= \frac{(n+1)(n+2)(2n+2+1)}{6} \\
 &= \frac{(n+1)(n+2)(2n+3)}{6} \\
 &= \frac{(n+1)(2n^2 + 4n + 3n + 6)}{6} \\
 &= \frac{(n+1)(2n^2 + 7n + 6)}{6}
 \end{aligned}$$

Damit ist bei der linken und der rechten Seite jeweils das gleiche herausgekommen und die Behauptung ist bewiesen.

## 8.3 STRUKTUR DER NATÜRLICHEN ZAHLEN

Die vollständige Induktion aus dem vorigen Abschnitt ist nur für die Menge der natürlichen Zahlen geeignet. Sie funktioniert aber nicht für alle Mengen: Für die reellen Zahlen  $\mathbb{R}$  z.B. erreicht die Konstruktion «bei 0 anfangen und dann immer um 1 hochzählen» einfach nicht alle Elemente der Menge. Die natürlichen Zahlen sind also etwas besonderes. Das haben wir schon in Kapitel 7 gesehen: es sind Zahlen, die zählen, und auf Seite 199 steht eine Datendefinition:

; Eine natürliche Zahl ist eine der folgenden:

; - 0

; - der Nachfolger einer natürlichen Zahl

Diese Datendefinition können wir als Bastelanleitung für die Menge der natürlichen Zahlen auffassen: Diese Datendefinition erreicht *jede* natürliche Zahl – jede Zahl kann in der Form  $0 + 1 + \dots + 1$  geschrieben werden. Für diese Art Bastelanleitung für Mengen ist charakteristisch, dass es ein oder mehrere *Basiselemente* gibt (in diesem Fall die 0) und dann ein oder mehrere Regeln, die aus beliebigen «kleineren» Elementen «größere» Elemente mit Hilfe eines Selbstbezug konstruieren, in diesem Fall die Regel, die besagt, dass für jede natürliche Zahl auch ihr Nachfolger eine natürliche Zahl ist.

Das mathematische Pendant für die Datendefinition der natürlichen Zahlen sieht so aus:

#### DEFINITION 8.1 (NATÜRLICHE ZAHLEN)

Die Menge der natürlichen Zahlen  $\mathbb{N}$  ist folgendermaßen definiert:

1.  $0 \in \mathbb{N}$
2. Für  $n \in \mathbb{N}$  ist auch  $n + 1 \in \mathbb{N}$ .
3. Die obigen Regeln definieren *alle*  $n \in \mathbb{N}$ .

Eine solche Definition heißt auch *induktive Definition*, die eine *induktive Menge* definiert. Die letzte Klausel ist bei induktiven Definition immer dabei – ohne sie könnte zum Beispiel die Menge der reellen Zahlen als  $\mathbb{N}$  durchgehen, weil die Klauseln davor keinen Anspruch auf Vollständigkeit erheben könnten. Diese Klausel heißt *induktiver Abschluss*.

Wenn Du genau hinschaust, siehst Du, dass die induktive Definition genau die gleiche Struktur hat wie die Definition der Induktion von Seite 217: Es gibt eine Klausel für die 0 und eine Klausel für den Schritt von  $n$  nach  $n + 1$ . Die induktive Definition sagt, dass es zwei verschiedene Sorten von natürlichen Zahlen gibt, nämlich das Basiselement 0 aus der ersten Regel der Definition und die (positiven) Zahlen  $n + 1$  aus der zweiten Klausel. Diese Struktur entspricht außerdem der Datendefinition und der Schablone aus Konstruktionsanleitung 19 auf Seite 202.

Entsprechend muss eine Behauptung über die natürlichen Zahlen für beide Sorten bewiesen werden: Einerseits für das Basiselement 0 und andererseits für die positiven Zahlen, die keine Basiselemente sind. Die Klauseln für die Basiselemente heißen *Induktionsverankerungen*. Da die zweite Klausel einen *Selbstbezug* enthält – es muss schon eine natürliche Zahl da sein, um eine weitere zu finden – muss dort ein Induktionsschluss bewiesen werden.

Die natürlichen Zahlen haben nur ein Basiselement und es gibt nur eine Klausel mit Selbstbezug. In Abschnitt 5.3 auf Seite 5.3 ging es um Finanzverträge – da gab es mehr als eine Klausel mit Selbstbezug, und auch in diesem Kapitel werden wir noch ein solches Beispiel sehen.

## 8.4 ENDLICHE FOLGEN

Die natürlichen Zahlen sind nicht die einzige induktiv definierte Menge. Ein weiteres Beispiel sind die *endlichen Folgen*, die den Listen entsprechen:

### DEFINITION 8.2

Sei  $M$  eine beliebige Menge. Die Menge  $M^*$  der *endlichen Folgen* über  $M$  ist folgendermaßen definiert:

1. Es gibt eine *leere Folge*<sup>2</sup>  $\epsilon \in M^*$ .
2. Wenn  $f \in M^*$  eine Folge ist und  $m \in M$ , so ist  $mf \in M^*$ , also auch eine Folge.
3. Die obigen Regeln definieren *alle*  $f \in M^*$ .

Eine Folge entsteht also aus einer bestehenden Folge dadurch, dass vorn noch ein Element angehängt wird. Folgen über  $M = \{a, b, c\}$  sind deshalb etwa

$$\epsilon, a\epsilon, b\epsilon, c\epsilon, aa\epsilon, ab\epsilon, ac\epsilon, \dots, abce, \dots cbbae, \dots \quad (\text{nicht alphabetisch sortiert})$$

Da das  $\epsilon$  bei nichtleeren Folgen immer dazugehört, wird es oft nicht mitnotiert.

Die Definition der endlichen Folgen ist analog zur Definition 8.1: Die erste Klausel ist der *Basisfall*, die zweite Klausel enthält einen *Selbstbezug* und die dritte Klausel bildet den induktiven Abschluss.

### 8.4.1 FUNKTIONEN AUF ENDLICHEN FOLGEN

Dieser Abschnitt demonstriert, wie mathematische Funktionen auf endlichen Folgen formuliert werden. Als Beispiel ist eine Funktion  $s : \mathbb{R}^* \rightarrow \mathbb{R}$  gefragt, die für eine Folge deren Summe ausrechnet, also z.B.:

$$s(1\ 2\ 3\ \epsilon) = 6$$

$$s(2\ 3\ 5\ 7\ \epsilon) = 17$$

$$s(17\ 23\ 42\ 5\ \epsilon) = 87$$

Die Funktion entspricht also der Funktion `list-sum` aus Abschnitt 6.1. Da es zwei verschiedene Sorten endlicher Folgen gibt – die leere Folge und die «nichtleeren» Folgen – liegt es nahe, die

<sup>2</sup> In Kapitel 6 wurde bereits erklärt, wieso es sinnvoll ist, mit der leeren Folge anzufangen.

entsprechende Funktion mit einer Verzweigung zu schreiben, die zwischen den beiden Sorten unterscheidet:

$$s(f) \stackrel{\text{def}}{=} \begin{cases} ? & \text{falls } f = \epsilon \\ ? & \text{falls } f = mf', m \in M, f' \in M^* \end{cases}$$

Es fehlen nur noch die Teile der Definition, wo Fragezeichen stehen:

Die Summe der leeren Folge ist 0: Wenn es sich um eine andere Zahl  $m \neq 0$  handeln würde, ließe sich schließlich die Summe einer beliebigen Folge durch das Anhängen der leeren Folge um  $m$  verändern. Die erste Lücke ist also schon geschlossen:

$$s(f) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{falls } f = \epsilon \\ ? & \text{falls } f = mf', m \in M, f' \in M^* \end{cases}$$

Im zweiten Fall handelt es sich bei  $f$  um ein zusammengesetztes Objekt mit den Bestandteilen  $m$  und  $f'$ . Deshalb können  $m$  und  $f'$  für die Konstruktion des Funktionswerts herangezogen werden:

$$s(f) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{falls } f = \epsilon \\ \dots m \dots f' \dots & \text{falls } f = mf', m \in M, f' \in M^* \end{cases}$$

Soweit sind die bekannten Techniken für die Konstruktion von Funktionen auf gemischten und zusammengesetzten Daten zur Anwendung gekommen, lediglich übertragen von Programm-Funktionen auf mathematische Funktionen.

Für den nächsten Schritt passt – genau wie beim Programmieren – ein rekursiver Aufruf zum Selbstbezug. Die Funktionsdefinition verdichtet sich folgendermaßen:

$$s(f) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{falls } f = \epsilon \\ \dots m \dots s(f') \dots & \text{falls } f = mf', m \in M, f' \in M^* \end{cases}$$

Hier ist  $s(f')$  die Summe aller Folgeelemente in  $f'$ . Gefragt ist die Summe aller Folgeelemente von  $f = mf'$ . Es fehlt also zur Summe nur noch  $m$  selbst:

$$s(f) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{falls } f = \epsilon \\ m + s(f') & \text{falls } f = mf', m \in M, f' \in M^* \end{cases}$$



Mit Papier und Bleistift können wir schnell nachvollziehen, dass die Definition korrekt arbeitet:

$$\begin{aligned}
 s(17\ 23\ 42\ 5\ \epsilon) &= 17 + s(23\ 42\ 5\ \epsilon) \\
 &= 17 + 23 + s(42\ 5\ \epsilon) \\
 &= 17 + 23 + 42 + s(5\ \epsilon) \\
 &= 17 + 23 + 42 + 5 + s(\epsilon) \\
 &= 17 + 23 + 42 + 5 + 0 \\
 &= 87
 \end{aligned}$$

Die Definition von  $s$  ruft sich also an der Stelle selbst auf, an der die induktive Definition der endlichen Folgen den Selbstbezug « $f' \in M^*$  eine Folge» enthält.

Vielleicht begegnest Du der Definition von  $s$  mit Skepsis, taucht doch  $s$  sowohl auf der linken als auch auf der rechten Seite auf. Es sieht so aus, als sei  $s$  «durch sich selbst definiert». Tatsächlich ist dies jedoch kein Problem, da:

- sich  $s(f)$  stets selbst auf einer *kürzeren* Folge  $f'$  aufruft, und
- schließlich bei der leeren Folge landet, bei der die Verzweigung greift und keinen weiteren rekursiven Aufruf mehr vornimmt.

Solange eine rekursive Funktion dem Schema von  $s$  folgt und damit der Struktur der Folgen selbst, sind diese beiden Bedingungen automatisch erfüllt.

#### 8.4.2 FOLGENINDUKTION

Das Gegenstück zur vollständigen Induktion heißt bei den Folgen *Folgeninduktion*. Der «Schluss von  $n$  auf  $n + 1$ » wird bei der Folgeninduktion zu je einem Schluss von  $f$  auf  $mf$  für alle Folgen  $f$  und alle  $m \in M$ . Oft kommt es dabei auf das Folgeelement  $m$  gar nicht an.

Zum Beweis der Behauptung, dass eine bestimmte Behauptung für alle  $f \in M^*$  gilt, genügt es, die folgenden Beweise zu führen:

1. Die Behauptung gilt für  $f = \epsilon$  (*Induktionsanfang*)
2. Wenn die Behauptung für eine Folge  $f$  gilt, so gilt sie auch für alle Folgen  $mf$  wobei  $m \in M$ . (*Induktionsschluss*).

Die Folgeninduktion funktioniert, weil sie der Struktur der Definition der endlichen Folgen 8.2 genauso folgt wie die vollständige Induktion der Struktur der natürlichen Zahlen. Das Lemma lässt sich auch mit Hilfe der vollständigen Induktion beweisen – siehe Aufgabe 8.7.

Entsprechend der vollständigen Induktion gibt es auch für die Folgeninduktion eine Anleitung:

1. Formuliere die zu beweisende Behauptung als Behauptung der Form «Für alle  $f \in M^*$  gilt ...», falls das noch nicht geschehen ist.
2. Schreibe die Überschrift « $f = \epsilon$ ». Schreibe die Behauptung noch einmal ab, wobei Du das «für alle  $f \in M^*$ » weglässt und für  $f$  das  $\epsilon$  einsetzt.
3. Beweise die abgeschriebene Behauptung. (Das ist oft einfach.)
4. Schreibe das Wort «Induktionsvoraussetzung:» Schreibe darunter die Behauptung noch einmal ab, wobei Du das «für alle  $f \in M^*$ » weglässt – lass das  $f$  da, wo es ist.
5. Schreibe «Induktionsschluss (zu zeigen):». Schreibe darunter die Behauptung noch einmal ab, wobei Du für  $f$  stattdessen  $mf$  einsetzt. (Lass das «für alle  $f \in M^*$ » wieder weg.)
6. Beweise den Induktionsschluss unter Verwendung der Induktionsvoraussetzung. Denke daran, die Behauptung für *alle*  $m \in M$  zu beweisen – das ist meist aber nicht immer einfach. Wenn die Behauptung eine Gleichung der Form  $A = B$  ist, kannst Du häufig die Induktionsvoraussetzung entweder direkt oder nach einigen Umformungen in den Induktionsschluss einsetzen.

Für ein sinnvolles Beispiel dient die Funktion  $\text{cat}$ , die Folgen aneinanderhängt:

$$\text{cat}(f_1, f_2) \stackrel{\text{def}}{=} \begin{cases} f_2 & \text{falls } f_1 = \epsilon \\ m \text{ cat}(f'_1, f_2) & \text{falls } f_1 = mf'_1, m \in M, f'_1 \in M^* \end{cases}$$

Es soll bewiesen werden, dass  $\text{cat}$  assoziativ ist, das heißt für alle  $u, v, w \in M^*$  gilt

$$\text{cat}(u, \text{cat}(v, w)) = \text{cat}(\text{cat}(u, v), w).$$

1. Die Form der Behauptung ist noch problematisch, da in ihr drei Folgen auftauchen – und keine davon heißt  $f$ . Im schlimmsten Fall musst Du raten, über welche Folge die Induktion geht und gegebenenfalls alle Möglichkeiten durchprobieren. Wir entscheiden uns für die erste Folge  $u$  und benennen sie in  $f$  um, damit der Beweis auf die Anleitung passt:

$$\forall f \in M^* : \forall v, w \in M^* : \text{cat}(f, \text{cat}(v, w)) = \text{cat}(\text{cat}(f, v), w)$$

2.  $f = \epsilon$ : Hier gilt

$$\text{cat}(\epsilon, \text{cat}(v, w)) = \text{cat}(v, w)$$

$$\text{cat}(\text{cat}(\epsilon, v), w) = \text{cat}(v, w)$$

nach der definierenden Gleichung.

3. Induktionsvoraussetzung:

$$\text{cat}(f, \text{cat}(v, w)) = \text{cat}(\text{cat}(f, v), w)$$

4. Induktionsschluss (zu zeigen):

$$\text{cat}(mf, \text{cat}(v, w)) = \text{cat}(\text{cat}(mf, v), w)$$

5. Wir benutzen die Definition von  $\text{cat}$ , um die linke Seite weiter auszurechnen:

$$\text{cat}(mf, \text{cat}(v, w)) = \text{mcat}(f, \text{cat}(v, w))$$

Hier steht aber  $\text{cat}(f, \text{cat}(v, w))$  – und das steht auch auf der linken Seite der Induktionsvoraussetzung. Wir können also einsetzen:

$$= \text{mcat}(\text{cat}(f, v), w)$$

Ebenso können wir die rechte Seite ausrechnen:

$$\begin{aligned} \text{cat}(\text{cat}(mf, v), w) &= \text{cat}(\text{mcat}(f, v), w) \\ &= \text{mcat}(\text{cat}(f, v), w) \end{aligned}$$

Das ist aber das gleiche, das auch bei der linken Seite herausgekommen ist – der Beweis ist fertig.

## 8.5 NOTATION FÜR INDUKTIVE DEFINITIONEN

Induktive Mengen kommen in der Informatik enorm häufig vor – so häufig, dass sich eine Kurzschreibweise für ihre induktiven Definitionen eingebürgert hat, die sogenannte *kontextfreie Grammatik*. In den meisten Informatik-Büchern geht mit dem Begriff eine langwierige mathematische Definition einher, die wir für dieses Buch nicht in aller Ausführlichkeit benötigen. Wir benutzen die kontextfreie Grammatik (ab hier einfach nur «Grammatik») informell als Abkürzung für eine länger ausgeschriebene induktive Definition. Hier die Grammatik für die natürlichen Zahlen:

$$\langle \mathbb{N} \rangle \rightarrow 0 \mid \langle \mathbb{N} \rangle + 1$$

In der Grammatik steht  $\langle \mathbb{N} \rangle$  für «eine natürliche Zahl». Lesen kannst Du die Grammatik so: Eine natürliche Zahl ist entweder die 0 oder hat die Form  $n + 1$  wobei  $n$  wiederum eine natürliche Zahl

ist. Der obligatorische induktive Abschluss («Die obigen Regeln definieren *alle*  $n \in \mathbb{N}$ .») wird stillschweigend vorausgesetzt und darum weggelassen.

Das Zeichen  $\rightarrow$  kann also als «ist» oder «hat die Form» gelesen werden, das Zeichen  $|$  als «oder». Die Notation  $\langle X \rangle$  definiert eine entsprechende Menge  $X$ . Die Grammatik ist also eine Art mathematische Schreibweise für zusammengesetzte Daten, gemischte Daten ( $|$ ) und Selbstbezüge, nur eben für mathematische Objekte.

Für die Menge der Folgen über einer Menge  $M$  genügt also folgende Notation:

$$\langle M^* \rangle \rightarrow \epsilon \mid \langle M \rangle \langle M^* \rangle$$

In beiden Definitionen ist jeweils der Selbstbezug klar zu sehen: Bei den natürlichen Zahlen taucht  $\langle \mathbb{N} \rangle$  in einer Klausel auf, bei den Folgen  $\langle M^* \rangle$ . Alle anderen Klauseln – also die ohne Selbstbezug – beschreiben Basisfälle.

## 8.6 STRUKTURELLE REKURSION

In Abschnitt 8.4 ist zu sehen, dass die Definition der beiden Beispielfunktionen auf Folgen ( $s$  und  $cat$ ) jeweils der induktiven Definition der Folgen «folgt» – die Definition beider Funktionen hat die Form:

$$F(f, \dots) \stackrel{\text{def}}{=} \begin{cases} \dots & \text{falls } f = \epsilon \\ \dots F(f') \dots & \text{falls } f = mf', m \in M, f' \in M^* \end{cases}$$

Das ist kein Zufall: Die rekursive Funktionsdefinition gehört zur induktiven Mengendefinition wie Pech zu Schwefel. Zwei Grundregeln legen diese Form fest:

1. Für jede Klausel gibt es eine Verzweigung mit einem Zweig der induktiven Definition.
2. Bei Selbstbezügen steht im entsprechenden Zweig ein Selbstaufruf.

Auch bei den natürlichen Zahlen lassen sich viele Operationen rekursiv aufschreiben. Zum Beispiel die Potenz:

$$b^n \stackrel{\text{def}}{=} \begin{cases} 1 & \text{falls } n = 0 \\ bb^{n'} & \text{falls } n = n' + 1, n' \in \mathbb{N} \end{cases}$$

Die Notation  $n = n' + 1$  folgt zwar der induktiven Definition, ist hier aber gleichbedeutend mit  $n > 0$  und  $n' = n - 1$ . Deshalb werden induktive Definitionen auf natürlichen Zahlen meist so

geschrieben:

$$b^n \stackrel{\text{def}}{=} \begin{cases} 1 & \text{falls } n = 0 \\ bb^{n-1} & \text{falls } n > 0 \end{cases}$$

Die Korrespondenz zwischen induktiven Definitionen und den rekursiven Funktionen lässt sich auch allgemein formulieren. Angenommen, die Menge  $X$  ist durch einer Grammatik definiert, die  $n$  Klauseln hat:

$$\langle X \rangle \rightarrow C_1 \mid \dots \mid C_n$$

Eine Funktion auf dieser Menge braucht dann – genau wie bei gemischten Daten – eine Verzweigung mit  $n$  Zweigen, eine für jede Klausel:

$$F(x) \stackrel{\text{def}}{=} \begin{cases} R_1 & \text{falls } x = F_1 \\ \dots & \\ R_n & \text{falls } x = F_n \end{cases}$$

Die Bedingung  $x = F_i$  ergibt sich jeweils aus der entsprechenden Klausel. Wenn dort Bezüge zu anderen Mengen oder Selbstbezüge stehen, so werden diese durch Variablen ersetzt und entsprechende Mengenzugehörigkeiten. Angenommen, die Klausel  $C_i$  hat beispielsweise diese Form:

$$[ \langle A \rangle \ \& \ \langle B \rangle ]$$

Dann müsste  $F_i$  entsprechend so aussehen:

$$x = [ a \ \& \ b ], a \in A, b \in B$$

Eine Klausel  $C_i$  mit Selbstbezug sähe beispielsweise so aus:

$$\{ \langle X \rangle \}$$

Das dazugehörige  $F_i$  sähe so aus:

$$x = \{ x' \}, x' \in X$$

Dementsprechend steht höchstwahrscheinlich in der rechten Seite  $R_i$  ein rekursiver Aufruf  $F(x')$ .

Funktionen dieser Form, die der Struktur einer induktiven Menge direkt folgen, heißen *strukturell rekursiv*.

Hier ein Beispiel für eine induktiv definierte Menge, deren Struktur etwas reichhaltiger ist, als die der natürlichen Zahlen oder der Folgen. Die Grammatik beschreibt einfache aussagenlogische Ausdrücke:

$$\begin{aligned}
 \langle E \rangle &\rightarrow \top \mid \perp \\
 &\mid \langle E \rangle \wedge \langle E \rangle \\
 &\mid \langle E \rangle \vee \langle E \rangle \\
 &\mid \neg \langle E \rangle
 \end{aligned}$$

Gedacht ist das folgendermaßen:  $\top$  steht für «wahr»,  $\perp$  für «falsch»,  $\wedge$  für «und»,  $\vee$  für «oder» und  $\neg$  für «nicht». Beispiele für Ausdrücke sind:

$$\begin{aligned}
 &\top \\
 &\top \vee \perp \\
 &\perp \vee \top \\
 &\neg \perp
 \end{aligned}$$

Beim Ausdruck  $\top \vee \perp \wedge \top$  ist nicht klar, wie er gemeint ist, da die Klammerung fehlt. In solchen Fällen schreiben wir einen Ausdruck genau wie in der Arithmetik mit Klammern:

$$\begin{aligned}
 &(\top \vee \perp) \wedge \top \\
 &\top \vee (\perp \wedge \top)
 \end{aligned}$$

Jeder solche Ausdruck hat einen Wahrheitswert, also «wahr» oder «falsch», und eine strukturell rekursive Funktion kann diesen Wahrheitswert errechnen. Diese Funktion «heißt»  $\llbracket e \rrbracket$  – für einen Ausdruck  $e$  liefert  $\llbracket e \rrbracket$  den Wert 1, falls der Ausdruck «wahr» ist und 0, falls «falsch». Die doppelten eckigen Klammern heißen in der Fachsprache «Semantikklammern» und werden oft benutzt, um Funktionen auf Mengen zu schreiben, die durch eine Grammatik definiert sind.

Die Funktionsdefinition besteht, wie oben beschrieben, auf jeden Fall aus einer Verzweigung, die für jede Klausel der Grammatik einen Zweig aufweist:

$$\llbracket e \rrbracket \stackrel{\text{def}}{=} \begin{cases} ? & \text{falls } e = \top \\ ? & \text{falls } e = \perp \\ ? & \text{falls } e = e_1 \wedge e_2 \\ ? & \text{falls } e = e_1 \vee e_2 \\ ? & \text{falls } e = \neg e' \end{cases}$$

In den Zweigen für  $\wedge$ ,  $\vee$  und  $\neg$  sind rekursive Aufrufe angezeigt, weil in der Grammatik dort Selbst-

bezüge stehen:

$$\llbracket e \rrbracket \stackrel{\text{def}}{=} \begin{cases} ? & \text{falls } e = \top \\ ? & \text{falls } e = \perp \\ \dots \llbracket e_1 \rrbracket \dots \llbracket e_2 \rrbracket & \text{falls } e = e_1 \wedge e_2 \\ \dots \llbracket e_1 \rrbracket \dots \llbracket e_2 \rrbracket & \text{falls } e = e_1 \vee e_2 \\ \dots \llbracket e' \rrbracket \dots & \text{falls } e = \neg e' \end{cases}$$

Die ersten beiden Fälle liefern 1 und 0 respektive, für die weiteren Fälle werden folgende Hilfsfunktionen definiert:

$$a(t_1, t_2) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{falls } t_1 = 1 \text{ und } t_2 = 1 \\ 0 & \text{sonst} \end{cases} \quad o(t_1, t_2) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{falls } t_1 = 0 \text{ und } t_2 = 0 \\ 1 & \text{sonst} \end{cases}$$

$$n(t) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{falls } t = 0 \\ 0 & \text{falls } t = 1 \end{cases}$$

Die Funktion  $a$  liefert nur dann 1, wenn  $t_1$  und  $t_2$  1 sind (sonst 0). Die Funktion  $o$  liefert dann 1, wenn  $t_1$  oder  $t_2$  1 sind (sonst 0). Die Funktion  $n$  liefert dann 1, wenn  $t$  nicht 1 ist (sonst 0). Diese Funktionen vervollständigen nun die Definition von  $\llbracket \cdot \rrbracket$ :

$$\llbracket e \rrbracket \stackrel{\text{def}}{=} \begin{cases} 1 & \text{falls } e = \top \\ 0 & \text{falls } e = \perp \\ a(\llbracket e_1 \rrbracket, \llbracket e_2 \rrbracket) & \text{falls } e = e_1 \wedge e_2 \\ o(\llbracket e_1 \rrbracket, \llbracket e_2 \rrbracket) & \text{falls } e = e_1 \vee e_2 \\ n(\llbracket e' \rrbracket) & \text{falls } e = \neg e' \end{cases}$$

## 8.7 STRUKTURELLE INDUKTION

Beweise von Eigenschaften strukturell rekursiver Funktionen funktionieren – wieder entsprechend den natürlichen Zahlen und den Folgen – mit *struktureller Induktion*. Strukturelle Induktion folgt ebenfalls der Struktur der Grammatik – vollständige Induktion und Folgeninduktion sind also Spezialfälle. Folgendes Beispiel – aufbauend auf den aussagenlogischen Ausdrücken aus dem vorigen

Abschnitt – illustriert die Technik:

SATZ 8.3 Aus einem aussagenlogischen Ausdruck  $e$  entsteht der Ausdruck  $\bar{e}$ , indem jedes  $\top$  durch  $\perp$ , jedes  $\perp$  durch  $\top$ , jedes  $\wedge$  durch  $\vee$  und jedes  $\vee$  durch  $\wedge$  ersetzt wird. Es gilt für jeden Ausdruck  $e$ :

$$\llbracket \bar{e} \rrbracket = n(\llbracket e \rrbracket)$$

Die Behauptung hat bereits die für einen Induktionsbeweis geeignete Form «für alle  $e$ », wobei  $e$  aus einer induktiv definierten Menge kommt. Die Behauptung muss jetzt für alle möglichen Fälle von  $e$  bewiesen werden – da die Grammatik für Ausdrücke fünf Fälle hat, sind auch fünf Fälle zu beweisen:  $e = \top$ ,  $e = \perp$ ,  $e = e_1 \wedge e_2$ ,  $e_1 \vee e_2$  und  $e = \neg e'$ , wobei  $e_1$ ,  $e_2$  und  $e'$  ihrerseits Ausdrücke sind.

Hier sind die Beweise für die Fälle  $e = \top$ ,  $e = \perp$ :

$$\begin{array}{ll} \llbracket \top \rrbracket = 1 & \llbracket \perp \rrbracket = 0 \\ \llbracket \bar{\top} \rrbracket = \llbracket \perp \rrbracket & \llbracket \bar{\perp} \rrbracket = \llbracket \top \rrbracket \\ = 0 & = 1 \\ = n(1) & = n(0) \end{array}$$

Als Nächstes ist der Fall  $e = e_1 \wedge e_2$  an der Reihe. Da die Klausel

$$\langle E \rangle \rightarrow \dots \mid \langle E \rangle \wedge \langle E \rangle$$

zwei Selbstbezüge hat, gibt es auch eine zweiteilige Induktionsvoraussetzung – die Behauptung kann für  $e_1$  und  $e_2$  angenommen werden:  $\llbracket \bar{e}_1 \rrbracket = n(\llbracket e_1 \rrbracket)$ ,  $\llbracket \bar{e}_2 \rrbracket = n(\llbracket e_2 \rrbracket)$ . Der Induktionschluss dazu sieht so aus:

$$\begin{aligned} \llbracket e_1 \wedge e_2 \rrbracket &= a(\llbracket e_1 \rrbracket, \llbracket e_2 \rrbracket) \\ \llbracket \overline{e_1 \wedge e_2} \rrbracket &= \llbracket \bar{e}_1 \vee \bar{e}_2 \rrbracket && \text{Definition von } \bar{\phantom{x}} \\ &= o(\llbracket \bar{e}_1 \rrbracket, \llbracket \bar{e}_2 \rrbracket) \\ &= o(n(\llbracket e_1 \rrbracket), n(\llbracket e_2 \rrbracket)) && \text{Induktionsvoraussetzung} \\ &= n(a(\llbracket e_1 \rrbracket, \llbracket e_2 \rrbracket)) \end{aligned}$$

Der letzte Schritt ergibt sich aus genauer Betrachtung der Definitionen von  $o$  und  $a$  beziehungsweise durch Einsetzen aller möglichen Werte für  $t_1$  und  $t_2$ .



Der Fall  $e = e_1 \vee e_2$  folgt analog. Wieder gibt es zwei Selbstbezüge, also gilt wieder die zweiteilige Induktionsvoraussetzung  $\llbracket \overline{e_1} \rrbracket = n(\llbracket e_1 \rrbracket)$ ,  $\llbracket \overline{e_2} \rrbracket = n(\llbracket e_2 \rrbracket)$ . Induktionsschluss (zu zeigen):

$$\begin{aligned}
 \llbracket e_1 \vee e_2 \rrbracket &= o(\llbracket e_1 \rrbracket, \llbracket e_2 \rrbracket) \\
 \llbracket \overline{e_1 \vee e_2} \rrbracket &= \llbracket \overline{e_1} \wedge \overline{e_2} \rrbracket && \text{Definition von } \overline{\phantom{x}} \\
 &= a(\llbracket \overline{e_1} \rrbracket, \llbracket \overline{e_2} \rrbracket) \\
 &= a(n(\llbracket e_1 \rrbracket), n(\llbracket e_2 \rrbracket)) && \text{Induktionsvoraussetzung} \\
 &= n(o(\llbracket e_1 \rrbracket, \llbracket e_2 \rrbracket))
 \end{aligned}$$

Schließlich fehlt noch der Fall  $e = \neg e'$ . Hier gibt es nur einen Selbstbezug, entsprechend lautet die Induktionsvoraussetzung:  $\llbracket \overline{e'} \rrbracket = n(\llbracket e' \rrbracket)$ . Induktionsschluss (zu zeigen):

$$\begin{aligned}
 \llbracket \neg e' \rrbracket &= n(\llbracket e' \rrbracket) \\
 \llbracket \overline{\neg e'} \rrbracket &= \llbracket \neg e' \rrbracket && \text{Definition von } \overline{\phantom{x}} \\
 &= n(\llbracket e' \rrbracket) \\
 &= n(n(\llbracket e' \rrbracket)) && \text{Induktionsvoraussetzung}
 \end{aligned}$$

Das Beispiel zeigt, dass strukturell induktive Beweise durchaus mehr als einen «Induktionsanfang» haben können – hier die Fälle  $\top$  und  $\perp$ . Ebenso gibt es mehr als einen Induktionsschluss – einen für jede Klausel mit Selbstbezug, hier sind das drei Fälle.

Für Beweise mit struktureller Induktion gilt also folgende Anleitung:

1. Formuliere Du die zu beweisende Behauptung als Behauptung der Form «Für alle  $x \in X$  gilt ...» (wobei  $X$  eine induktiv definierte Menge ist), falls das noch nicht geschehen ist.
2. Führe jetzt einen Beweis für jede einzelne Klausel  $C_i$  der induktiven Definition:
3. Schreibe die Überschrift « $x = F_i$ », wobei  $F_i$  eine Bedingung ist, die der Klausel entspricht. Schreibe die Behauptung noch einmal ab, wobei Du das «für alle  $x \in X$ » weglässt und für  $x$  stattdessen  $F_i$  einsetzt.
4. Wenn die Klausel keinen Selbstbezug enthält, so beweise die Behauptung direkt.
5. Wenn die Klausel einen oder mehrere Selbstbezüge enthält, so stehen in der Überschrift Variablen  $x_j$  oder  $x'$  o.ä., die ihrerseits Element von  $X$  sind. Schreibe dann die Überschrift «Induktionsvoraussetzung:». Schreibe darunter für jeden Selbstbezug die Behauptung noch einmal ab, wobei Du das «für alle  $x \in X$ » weglässt und für  $x$  stattdessen  $x_j$  beziehungsweise  $x'$  einsetzt.

Schreibe darunter das Wort «Induktionsschluss» und beweise die Behauptung. Denke daran,

die Induktionsvoraussetzung zu benutzen.

## ANMERKUNGEN

Die erste konstruktive Definitionen der natürlichen Zahlen wurde bereits im Jahr 1889 von PEANO vorgeschlagen:

### DEFINITION 8.4 (PEANO-AXIOME)

Die Menge  $\mathbb{N}$  der natürlichen Zahlen ist durch folgende Eigenschaften, die *Peano-Axiome*, gegeben:

1. Es gibt eine natürliche Zahl  $0 \in \mathbb{N}$ .
2. Zu jeder Zahl  $n \in \mathbb{N}$  gibt es eine Zahl  $n' \in \mathbb{N}$ , die *Nachfolger von  $n$*  heißt.
3. Für alle  $n \in \mathbb{N}$  ist  $n' \neq 0$ .
4. Aus  $n' = m'$  folgt  $n = m$ .
5. Eine Menge  $M$  von natürlichen Zahlen, welche die 0 enthält und mit jeder Zahl  $m \in M$  auch deren Nachfolger  $m'$ , ist mit  $\mathbb{N}$  identisch.

Diese Definition ist äquivalent zu Definition 8.1 von Seite 220. Wie in Definition 8.1, lässt sich  $\mathbb{N}$  sich aus den Peano-Axiomen schrittweise konstruieren. Aus 1. und 2. folgt, dass es natürliche Zahlen

$$0, 0', 0'', 0''', 0''', \dots$$

gibt. Ohne die 0 am Anfang entsteht die Darstellung der natürlichen Zahlen durch Striche. Die Axiome 3 und 4 besagen, dass jede Zahl nur eine Strichdarstellung besitzt: jedes Anfügen eines Striches «'» erzeugt eine völlig neue Zahl.

Die Peano-Formulierung ist zwar ähnlich zu Definition 8.1, weist aber auch interessante Unterschiede auf:

- Für « $n'$ » ist im täglichen Umgang natürlich die Bezeichnung « $n + 1$ » gebräuchlich.
- Die dritte und vierte Bedingung zusammen besagen, dass die Nachfolgerfunktion injektiv ist, das heißt durch fortgesetzte Anwendung der Nachfolgerfunktion entstehen immer neue Elemente.
- Schließlich beschreibt die fünfte Bedingung den *induktiven Abschluss*, der festlegt, dass außer den solchermaßen erzeugten Elementen keine weiteren existieren. Axiom 5 wird auch *Induktionsaxiom* genannt. In Definition 8.1 hat das Induktionsaxiom die üblichere Form «Die obigen Regeln definieren *alle*  $n \in \mathbb{N}$ .» Ebenfalls üblich ist eine Formulierung wie wie «...die *kleinste* Menge mit den Eigenschaften ...».

## AUFGABEN

### AUFGABE 8.1

Beweise die Gaußsche Summenformel mit vollständiger Induktion:

$$\forall n \in \mathbb{N} : \sum_{i=0}^n i = \frac{n \times (n+1)}{2}$$

### AUFGABE 8.2

Betrachte folgende Tabelle:

$$\begin{aligned} 1 &= 1 \\ 1 - 4 &= -(1 + 2) \\ 1 - 4 + 9 &= 1 + 2 + 3 \\ 1 - 4 + 9 - 16 &= -(1 + 2 + 3 + 4) \end{aligned}$$

Rate die Gleichung, die dieser Tabelle zugrundeliegt und schreibe in mathematischer Notation auf. Beweise die Gleichung!

### AUFGABE 8.3

Beweise, dass folgende Gleichung für alle  $n \in \mathbb{N}$  gilt:

$$\sum_{i=0}^n i^3 = \left( \sum_{i=0}^n i \right)^2$$

### AUFGABE 8.4

Was ist falsch an folgendem Induktionsbeweis?

**Behauptung** Alle Pferde haben die gleiche Farbe.

**Beweis**

**Induktionsanfang** Für eine leere Menge von Pferden gilt die Behauptung trivialerweise.

**Induktionsschluss** Gegeben sei eine Menge von  $n+1$  Pferden. Nimm ein Pferd aus der Menge heraus – die restlichen Pferde haben per Induktionsannahme die gleiche Farbe. Nimm ein anderes

Pferd aus der Menge heraus – wieder haben die restlichen Pferde per Induktionsannahme die gleiche Farbe. Da die übrigen Pferde die Farbe in der Zwischenzeit nicht plötzlich gewechselt haben können, war es in beiden Fällen die gleiche Farbe, und alle  $n + 1$  Pferde haben diese Farbe.  $\square$

#### AUFGABE 8.5

Beweise mittels Induktion, dass über Folgen für  $u, v \in M^*$  gilt:

$$\text{len}(\text{cat}(u, v)) = \text{len}(u) + \text{len}(v)$$

Dabei sei  $\text{len}$  die Länge von Folgen, definiert als:

$$\text{len}(f) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{falls } f = \epsilon \\ \text{len}(f') + 1 & \text{falls } f = mf' \end{cases}$$

#### AUFGABE 8.6

Beweise durch Folgeninduktion:

$$\text{cat}(v, w) = \text{cat}(z, w) \Rightarrow v = z$$

$$\text{cat}(v, w) = \text{cat}(v, z) \Rightarrow w = z$$

#### AUFGABE 8.7

Beweise die Korrektheit der Folgeninduktion aus Abschnitt 8.4.2 (Seite 223) mit Hilfe der vollständigen Induktion. Beschreibe dazu, wie sich eine Aussage über alle Folgen in eine Aussage über die Längen der Folgen umwandeln lässt. Setze dann die Klauseln der vollständigen Induktion mit den entsprechenden Klauseln der Folgeninduktion in Beziehung.

#### AUFGABE 8.8

Gib eine induktive Definition der umgangssprachlich definierten Funktion  $\bar{\phantom{x}}$  aus Abschnitt 8.7 (Seite 230) an.

#### AUFGABE 8.9

Gib eine Datendefinition für die aussagenlogischen Ausdrücke aus Abschnitt 8.7 an. Programmiere die Funktion  $\llbracket \_ \rrbracket$  sowie die Funktion  $\neg$ .

#### AUFGABE 8.10

Die *Fibonacci*-Funktion auf den natürlichen Zahlen ist folgendermaßen definiert:

$$fib(n) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{falls } x = 0 \\ 1 & \text{falls } x = 1 \\ fib(n-2) + fib(n-1) & \text{sonst} \end{cases}$$

Beweise, dass  $fib(n)$  die ganze Zahl ist, die am nächsten zu  $\Phi^n / \sqrt{5}$  liegt, wobei  $\Phi = (1 + \sqrt{5})/2$ .

Anleitung: Zeige, dass  $fib(n) = (\Phi^n - \Psi^n) / \sqrt{5}$ , wobei  $\Psi = (1 - \sqrt{5})/2$ .

#### AUFGABE 8.11

Eine partielle Ordnung  $\preceq$  auf einer Menge  $M$  heißt *wohlfundiert* oder *noethersch*, wenn es keine unendlichen Folgen  $(x_i)_{i \in \mathbb{N}}$  gibt, so dass für alle  $i \in \mathbb{N}$  gilt

$$x_{i+1} \preceq x_i \text{ und } x_{i+1} \neq x_i$$

auch geschrieben als:

$$x_{i+1} \prec x_i .$$

Das Beweisprinzip der *wohlfundierten* oder *noetherschen Induktion* ist dafür zuständig, die Gültigkeit einer Eigenschaft  $P$  auf  $M$  zu beweisen, wenn es eine wohlfundierte Ordnung  $\preceq$  auf  $M$  gibt. Es besagt, dass es ausreicht,  $P(z)$  unter der Voraussetzung nachzuweisen, dass  $P(y)$  für alle Vorgänger  $y$  von  $z$  gilt. Anders gesagt:

$$\forall z \in M : (\forall y \in M : y \prec z \Rightarrow P(y)) \Rightarrow P(z) \Rightarrow \forall x \in M : P(x)$$

Leite die vollständige und die strukturelle Induktion als Spezialfälle der wohlfundierten Induktion her.



## 9 HIGHER-ORDER-PROGRAMMIERUNG

Der renommierte Informatiker Paul Hudak wurde einst gefragt, was die drei wichtigsten Dinge beim Programmieren seien. Er antwortete: «Abstraktion, Abstraktion, Abstraktion». Entsprechend ist das Mantra 4 auf Seite 37 auch eins der wichtigsten:

MANTRA 4

Wenn Du zwei Programmteile siehst, die sich nur an wenigen Stellen unterscheiden und die inhaltlich verwandt sind, abstrahiere!

In diesem Kapitel werden wir dieses Mantra konsequent anwenden. Dabei kommt oft eine besondere Art Funktion heraus: Funktionen, die andere Funktionen als Eingabe akzeptieren oder als Ausgabe liefern. Solche Funktionen heißen auch *Funktionen höherer Ordnung* oder *Higher-Order-Funktion*. Um die resultierende Art der Programmierung – die *Higher-Order-Programmierung* – geht es in diesem Kapitel.

### 9.1 FUSSBALL-FAKTEN ERMITTELN

Eine nahezu unerschöpfliche Quelle für Diskussionen stellen die Fußballergebnisse dar. Hier stellen sich so bedeutsame Fragen wie:

- Wie hat diese Saison Bayern München gegen 1. FC Kaiserslautern auswärts gespielt?<sup>1</sup>
- Wie viele Tore sind in dieser Saison gefallen?
- Welche Mannschaft ist abstiegsgefährdet?

All diese Informationen speisen sich aus Ergebnissen der Spiele einer Saison. Ein Spiel können wir folgendermaßen charakterisieren:

```
; Ein Spiel hat folgende Eigenschaften:  
; - Spieltag  
; - Heimmannschaft  
; - Heimmannschaft-Tore  
; - Gastmannschaft
```

---

<sup>1</sup> Wir wissen, das ist schon eine Weile her.

; - Gastmannschaft-Tore

Es handelt sich sichtlich um zusammengesetzte Daten. Wir müssen uns überlegen, welche Signaturen zu den einzelnen Bestandteilen passen. Spieltag und Tore sind allesamt natürliche Zahlen. Eine Mannschaft können wir durch ihren Namen als Zeichenkette repräsentieren. Das ergibt folgende Record-Definition:

```
(define-record game
  make-game game?
  (game-matchday natural)
  (game-home-team string)
  (game-home-goals natural)
  (game-guest-team string)
  (game-guest-goals natural))
```

Es folgen hier beispielhaft die Ergebnisse des ersten Spieltags der Bundesliga-Saison 2009/2010:

```
(define game1 (make-game 1 "Wolfsburg" 2 "Stuttgart" 0))
(define game2 (make-game 1 "Mainz" 2 "Bayer 04" 2))
(define game3 (make-game 1 "Hertha" 1 "Hannover" 0))
(define game4 (make-game 1 "Bremen" 2 "Frankfurt" 3))
(define game5 (make-game 1 "Nürnberg" 1 "Schalke" 2))
(define game6 (make-game 1 "Dortmund" 1 "1. FC Köln" 0))
(define game7 (make-game 1 "Hoffenheim" 1 "Bayern" 1))
(define game8 (make-game 1 "Bochum" 3 "Gladbach" 3))
(define game9 (make-game 1 "Freiburg" 1 "Hamburg" 1))
```

```
(define day1
  (list game1 game2 game3 game4 game5 game6 game7 game8 game9))
```

Die kompletten Ergebnisse dieser Saison lassen sich unter dem Namen `soccer.rkt` von der Webseite des Buchs – <https://www.deinprogramm.de/> – herunterladen.

Eine recht einfache Frage ist die Bestimmung der Punktzahl, welche die Gastgebermannschaft in einem bestimmten Spiel erzielt hat. Die Punktzahl ist zwar eine natürliche Zahl, es gibt aber nur drei Möglichkeiten: 0, 1 und 3. Wir können also eine präzisere Signatur als `points` definieren:

```
; Punktzahl in Spiel
(define points
  (signature (enum 0 1 3)))
```



Unsere Funktion für die Bestimmung der Punktzahl hat Kurzbeschreibung, Signatur, Tests und Gerüst wie folgt:

```
; Punktzahl für Heimmannschaft berechnen
(: home-points (game -> points))

(check-expect (home-points game1) 3)
(check-expect (home-points game2) 1)
(check-expect (home-points game3) 3)
(check-expect (home-points game4) 0)
```

```
(define home-points
  (lambda (game)
    ...))
```

Entsprechend der Signatur der Ausgabe fallen die Spiele in drei unterschiedliche Klassen, das heißt wir brauchen eine Verzweigung mit drei Zweigen entsprechend den drei möglichen Punktzahlen:

```
(define home-points
  (lambda (game)
    (cond
      (... 3)
      (... 0)
      (... 1))))
```

Die Bedingungen der drei Zweige kommen aus den Fußballregeln – wir müssen die Tore von Heim- und Gastmannschaft vergleichen:

```
(define home-points
  (lambda (game)
    (cond
      ((> (game-home-goals game) (game-guest-goals game)) 3)
      ((< (game-home-goals game) (game-guest-goals game)) 0)
      ((= (game-home-goals game) (game-guest-goals game)) 1))))
```

Das funktioniert schon korrekt, ist aber noch unelegant. Das `(game-home-goals game)` ist dreimal wiederholt – wir sollten eine lokale Definition einführen, damit es nur einmal dasteht. Ebenso für den Aufruf von `game-guest-goals`. Das sieht dann so aus:

```
(define home-points
```

```
(lambda (game)
  (define goals1 (game-home-goals game))
  (define goals2 (game-guest-goals game))
  (cond
    ((> goals1 goals2) 3)
    ((< goals1 goals2) 0)
    ((= goals1 goals2) 1))))
```

Du könntest berechtigterweise fragen, warum die lokalen Variablen `goals1` und `goals2` heißen, und nicht `home-goals` und `guest-goals` heißen. Das wird bei der nächsten Funktion (hoffentlich) klar, welche die Punkte ausrechnet, die dem Gästeteam zustehen:

```
; Punktzahl für Gastmannschaft berechnen
(: guest-points (game -> points))
```

```
(define guest-points
  (lambda (game)
    (define goals1 (game-guest-goals game))
    (define goals2 (game-home-goals game))
    (cond
      ((> goals1 goals2) 3)
      ((< goals1 goals2) 0)
      ((= goals1 goals2) 1))))
```

Diese beiden Funktionen sind weitgehend identisch, der einzige Unterschied ist die Definition der lokalen Variablen `goals1` und `goals2`. Eine solche Duplizierung von Code ist immer schlecht, vor allem, wenn sich etwas ändert. Theoretisch könnte der Fußballbund die Regeln für die Vergabe von Punkten ändern, und dann müssten wir beide Funktionen in gleicher Weise anpassen. Die Lösung für das Problem zeigt sich dadurch, dass wir aus dem gemeinsamen Code der beiden Funktionen eine neue Funktion machen, entsprechend dem Abstraktions-Mantra:

#### MANTRA 4

Wenn Du zwei Programmteile siehst, die sich nur an wenigen Stellen unterscheiden und die inhaltlich verwandt sind, abstrahiere!

Um zu abstrahieren, kopieren wir eine der beiden Funktionsdefinitionen und geben ihr einen neuen Namen, in diesem Fall `compute-points`:

```
(define compute-points
  (lambda (game)
    (define goals1 (game-guest-goals game))
    (define goals2 (game-home-goals game))
    (cond
      ((> goals1 goals2) 3)
      ((< goals1 goals2) 0)
      ((= goals1 goals2) 1))))
```

Nun identifizieren wir in der kopierten Funktion die Stellen, die sich bei den beiden ursprünglichen Funktionen unterscheiden und ersetzen diese Stellen durch neue Namen. Hier sind das gerade `game-guest-goals` und `game-home-goals`, wir ersetzen sie durch die Namen `get-goals-1` und `get-goals-2`:

```
(define compute-points
  (lambda (game)
    (define goals1 (get-goals-1 game))
    (define goals2 (get-goals-2 game))
    (cond
      ((> goals1 goals2) 3)
      ((< goals1 goals2) 0)
      ((= goals1 goals2) 1))))
```

Diese neuen Variablen sind noch ungebunden, wir müssen sie deshalb noch im `lambda` unterbringen:

```
(define compute-points
  (lambda (get-goals-1 get-goals-2 game)
    (define goals1 (get-goals-1 game))
    (define goals2 (get-goals-2 game))
    (cond
      ((> goals1 goals2) 3)
      ((< goals1 goals2) 0)
      ((= goals1 goals2) 1))))
```

Damit können wir die bisherigen Definitionen von `home-points` und `guest-points` ersetzen durch die folgenden:

```
(define home-points
```

```
(lambda (game)
  (compute-points game-home-goals game-guest-goals game)))
```

```
(define guest-points
  (lambda (game)
    (compute-points game-guest-goals game-home-goals game)))
```

Diese neuen Definitionen lassen die Gemeinsamkeiten und Unterschiede beider Funktionen klar erkennen und vermeiden die doppelte Definition der Fußball-Punkteregeln.

**Wichtig:** Diese beiden Definitionen müssen *nach* der Definition von `compute-points` stehen, da davor `compute-points` noch nicht definiert ist.

Der neuen Funktion fehlt noch eine Signatur. Die Funktion akzeptiert drei Argumente. Das letzte hat die Signatur `game`. Die ersten beiden werden aus den Funktionen `game-home-goals` und `game-guest-goals` bestückt, die jeweils die Signatur `(game -> natural)` haben. Daraus können wir die Signatur für `compute-points` zusammensetzen:

```
(: compute-points ((game -> natural) (game -> natural) game ->
  points))
```

In der Signatur tauchen mehrere Pfeile auf, weil die Funktion `compute-points` ihrerseits zwei Funktionen als Argumente akzeptiert. Solche Funktionen mit mehreren Pfeilen in der Signatur heißen *Funktionen höherer Ordnung* oder *Higher-Order-Funktionen*.

Die Abstraktion bei der Entwicklung von `compute-points` hätten wir auch etwas anders anstellen können: Wir haben bei `compute-points` die beiden Parameter `get-goals-1` und `get-goals-2` zu dem schon bestehenden `lambda` hinzugefügt. Wir können stattdessen auch ein neues `lambda` einfügen. (Bei Definitionen, in denen noch kein `lambda` steht, müssten wir das sowieso.) Das sieht dann so aus:

```
(define make-compute-points
  (lambda (get-goals-1 get-goals-2)
    (lambda (game)
      (define goals1 (get-goals-1 game))
      (define goals2 (get-goals-2 game))
      (cond
        ((> goals1 goals2) 3)
        ((< goals1 goals2) 0)
        ((= goals1 goals2) 1))))))
```

Die Signatur dieser Funktion unterscheidet sich leicht von der Signatur von `compute-points`. Sie akzeptiert nicht mehr drei Argumente sondern nur zwei, aber liefert dafür eine Funktion, die von einem Spiel die Punkte liefert:

```
(: make-compute-points ((game -> natural) (game -> natural)
                        -> (game -> points)))
```

Auch mit dieser Funktion können wir alternative Definitionen von `make-compute-points` schreiben. Auf den ersten Blick ist das umständlich, wenn wir das genauso machen wie mit `compute-points`, da wir noch mehr Klammern brauchen:

```
(define home-points
  (lambda (game)
    ((make-compute-points game-home-goals game-guest-goals) game)))

(define guest-points
  (lambda (game)
    ((make-compute-points game-guest-goals game-home-goals) game)))
```

Vielleicht fällt Dir aber das Muster auf, das diese beiden Definitionen gemeinsam haben:

```
(define f
  (lambda (x)
    (g x)))
```

Das heißt, wenn das Programm `f` aufruft, ruft diese Funktion direkt `g` auf, mit der gleichen Eingabe: `f` und `g` sind also äquivalent, und wir könnten genauso gut schreiben:

```
(define f g)
```

Das gleiche Prinzip können wir auch auf `home-points` und `guest-points` anwenden:

```
(define home-points
  (make-compute-points game-home-goals game-guest-goals))

(define guest-points
  (make-compute-points game-guest-goals game-home-goals))
```

Man kann auch schön an der Signatur von `make-compute-points` sehen, dass diese als Resultat eine Funktion mit Signatur `(game -> points)` liefert, das ist gerade die gewünschte Signatur von `home-points` und `guest-points`.

Vielleicht hast Du Dich über den Namen `make-compute-points` gewundert: Während die Funktion `compute-points` direkt eine Punktzahl berechnet, macht `make-compute-points` eine Funktion, welche die Punktzahl berechnet: Darum der Präfix `make-`. Es handelt sich also um eine Art «Funktionsfabrik».

#### AUFGABE 9.1

Wir hätten das neue `lambda` auch unterhalb des alten `lambda` einfügen können statt oberhalb. Welche Signatur hätte die Funktion `make-compute-points` dann? Warum ist das keine so gute Idee? ☐

Die Abstraktionstechnik, die wir für die Konstruktion von `compute-points` und `make-compute-points` verwendet haben, wenden wir in diesem Buch noch öfter an. Sie verdient deshalb eine eigene Konstruktionsanleitung:

#### KONSTRUKTIONSANLEITUNG 20 (ABSTRAKTION)

Wenn Du zwei Definitionen geschrieben hast, die inhaltlich verwandt sind und viele Ähnlichkeiten aufweisen, abstrahiere wie folgt:

1. Kopiere eine der beiden Definitionen und gib ihr einen neuen Namen.
2. Ersetze die Stellen, bei denen sich die beiden Definitionen unterscheiden, jeweils durch eine neue Variable.
3. Füge die neuen Variablen als Parameter zum `lambda` der Definition hinzu oder füge ein neues `lambda` mit diesen Parametern ein. Du musst gegebenenfalls rekursive Aufrufe der Funktion anpassen.
4. Schreibe eine Signatur für die neue Funktion.
5. Ersetze die beiden alten Definitionen durch Aufrufe der neuen Definition.

## 9.2 HIGHER-ORDER-FUNKTIONEN AUF LISTEN

Als Nächstes wollen wir herausbekommen, welche Spiele einer Saison unentschieden ausgingen. Dazu ist zunächst eine Funktion notwendig, die feststellt, ob *ein* bestimmtes Spiel unentschieden war:

```
; Ist Spiel unentschieden?
(: game-draw? (game -> boolean))
```

```
(define game-draw?
  (lambda (g)
    (= 1 (home-points g))))
```

Die Spiele der Saison liegen in einer Liste, wir schreiben also eine Funktion, die aus einer Liste von Spielen die unentschiedenen herausholt. Kurzbeschreibung, Signatur und Gerüst folgen der Konstruktionsanleitung für Funktionen auf Listen. Hier ist die Schablone:

```
; Unentschiedene Spiele rausfiltern
(: drawn-games ((list-of game) -> (list-of game)))

(check-expect (drawn-games day1) (list game2 game7 game8 game9))

(define drawn-games
  (lambda (list)
    ...))
```

Die Funktion hat eine Liste als Eingabe, es kommt also die entsprechende Schablone zur Anwendung:

```
(define drawn-games
  (lambda (list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ... (first list) ...
       ... (drawn-games (rest list)) ...))))
```

Der rekursive Aufruf liefert die unentschiedenen Spiele des Rests, wir müssen also nur noch beim ersten Spiel entscheiden, ob es unentschieden ausging:

```
(define drawn-games
  (lambda (list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ... (game-draw? (first list)) ...
       ... (drawn-games (rest list)) ...))))
```

Die Funktion `game-draw?` liefert einen booleschen Wert, den können wir eigentlich nur in eine binäre Verzweigung stecken:

```
(define drawn-games
  (lambda (list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       (if (game-draw? (first list))
           ...
           ...))
      ... (drawn-games (rest list)) ...))))
```

Abhängig davon, ob das Spiel unentschieden ist oder nicht, bringen wir es im Ergebnis unter:

```
(define drawn-games
  (lambda (list)
    (cond
      ((empty? list) empty)
      ((cons? list)
       (if (game-draw? (first list))
           (cons (first list) (drawn-games (rest list)))
           (drawn-games (rest list))))))
```

Fertig!

Analog zu `drawn-games` schreiben wir als Nächstes eine Funktion, die alle Spiele aus einer Liste extrahiert, die die Heimmannschaft gewonnen hat. Dazu brauchen wir erst einmal eine Funktion analog zu `game-draw?`:

```
; Hat die Heimmannschaft gewonnen?
(: home-won? (game -> boolean))
```

```
(check-expect (home-won? game1) #t)
(check-expect (home-won? game2) #f)
(check-expect (home-won? game4) #f)
```

```
(define home-won?
  (lambda (game)
    (= 3 (home-points game))))
```



Die Funktion `home-won-games` entsteht genau analog zu `drawn-games`. Am Ende kommt folgendes heraus:

```
; Spiele herausfiltern, bei denen die Heimmannschaft gewann
(: home-won-games ((list-of game) -> (list-of game)))

(check-expect (home-won-games day1) (list game1 game3 game6))

(define home-won-games
  (lambda (list)
    (cond
      ((empty? list) empty)
      ((cons? list)
       (if (home-won? (first list))
           (cons (first list)
                 (home-won-games (rest list)))
           (home-won-games (rest list)))))))
```

Die beiden Funktionen `drawn-games` und `home-won-games` sind bis auf den Namen fast identisch. Der einzige Unterschied ist, dass an der Stelle von `game-draw?` in der zweiten Funktion `home-won?` steht. Darüber können wir nach Konstruktionsanleitung 20 auf Seite 244 abstrahieren. Dazu kopieren wir die letzte Funktion und denken uns einen neuen Namen aus – da es in beiden Fällen ums «extrahieren» geht, nehmen wir `extract-games`:

```
(define extract-games
  (lambda (list)
    (cond
      ((empty? list) empty)
      ((cons? list)
       (if (home-won? (first list))
           (cons (first list)
                 (extract-games (rest list)))
           (extract-games (rest list)))))))
```

Als Nächstes müssen wir die Stelle, die sich bei beiden Definitionen unterscheidet (`game-draw?` beziehungsweise `home-won?`), durch eine neue Variable bezeichnen. Die nennen wir `f?`:

```
(define extract-games
  (lambda (list)
```

```
(cond
  ((empty? list) empty)
  ((cons? list)
   (if (f? (first list))
       (cons (first list)
             (extract-games (rest list)))
       (extract-games (rest list))))))
```

Als Nächstes müssen wir die neue Variable als Parameter zum **lambda** hinzufügen. Aufpassen: Die beiden rekursiven Aufrufe müssen ebenfalls um den Parameter erweitert werden:

```
(define extract-games
  (lambda (f? list)
    (cond
      ((empty? list) empty)
      ((cons? list)
       (if (f? (first list))
           (cons (first list)
                 (extract-games f? (rest list)))
           (extract-games f? (rest list)))))))
```

(Es ist Geschmackssache, ob der neue Parameter vor oder hinter die alten kommt.)

Die Funktionsdefinition ist damit fertig, aber wir müssen noch eine Signatur schreiben. Wir fangen mit der Signatur von **home-won-games** an und erweitern um den zusätzlichen Parameter:

```
(: extract-games (... (list-of game) -> (list-of game)))
```

Der zusätzliche Parameter ist **f?**, den wir für **game-draw?** und **home-won?** eingesetzt haben. Diese beiden Funktionen haben die Signatur **(game -> boolean)**, das können wir also für die Ellipse einsetzen:

```
(: extract-games ((game -> boolean) (list-of game) -> (list-of game)))
```

Nun können wir Aufrufe der alten Funktionen durch Aufrufe der neuen ersetzen. Wir tun das, indem wir die Testfälle für **drawn-games** und **home-won-games** anpassen:

```
(check-expect (extract-games game-draw? day1) (list game2 game7 game8
  game9))
(check-expect (extract-games home-won? day1) (list game1 game3 game6))
```

Fertig!

Vielleicht hast Du das Gefühl, das Muster der Funktion `extract-games` schon gesehen zu haben. In der Tat sieht die Funktion `live-dillos` in Abschnitt 6.6.2 auf Seite 6.6.2 ebenfalls den Funktionen `drawn-games` und `home-won-games` sehr ähnlich. Hier ist sie noch einmal zur Erinnerung:

```
(: live-dillos ((list-of dillo) -> (list-of dillo)))
```

```
(define live-dillos
  (lambda (dillos)
    (cond
      ((empty? dillos) empty)
      ((cons? dillos)
       (if (dillo-alive? (first dillos))
           (cons (first dillos)
                 (live-dillos (rest dillos)))
           (live-dillos (rest dillos)))))))
```

Und tatsächlich – wenn wir `dillos` in `list` umbenennen und den gleichen Abstraktionsprozess wie bei `drawn-games` und `home-won-games` durchlaufen, dann entsteht die identische Funktionsdefinition zu `extract-games`. Allerdings gibt es dann immer noch einen Unterschied – die Signaturen sind unterschiedlich. Die Funktion `extract-games` kann nur Listen von Spielen verarbeiten, während `live-dillos` eine Liste von Gürteltieren akzeptiert. Können auch darüber – einmal `game`, das andere Mal `dillo` – abstrahieren?

Können wir! Dazu sollten wir die Funktion noch umbenennen, wir wählen `extract-list` statt `extract-games`. Vor allem aber müssen wir die Signatur ändern, die bisher auf `game` abonniert ist. Auf Signaturebene abstrahieren wir durch Signaturvariablen und ersetzen einfach mal probierhalber jedes `game` durch `%a`:

```
(: extract-list ((%a -> boolean) (list-of %a) -> (list-of %a)))
```

Diese Funktion können wir nun versuchen zu verstehen: Sie akzeptiert eine Liste von `%as` und eine Funktion, die `%as` akzeptiert. Man kann daran fast schon sehen, was die Funktion macht: Das einzige, was sie mit der Funktion anfangen kann, ist sie auf die Elemente der Liste anzuwenden – sonst gibt es ja weit und breit keine anderen `%as`. Dass die Funktion einen booleschen Wert produziert, suggeriert schon, dass anhand dieses Werts unterschieden wird. Die Signatur liefert also wertvolle Informationen darüber, was die Funktion macht – wenn auch nicht alle.

Schreibe eine andere sinnvolle Funktion mit der gleichen Signatur wie `extract-list`! □

Wir betrachten noch eine weitere Anwendung von `extract-list`: Wir wollen aus einer Liste von Spielen die Spiele extrahieren, an denen eine bestimmte Mannschaft teilgenommen hat. Auch dazu definieren wir eine Hilfsfunktion, die feststellt, ob eine Mannschaft bei einem Spiel die Heim- oder die Gastmannschaft ist:

```
; Spielt Mannschaft bei Spiel?
(: plays-game? (team game -> boolean))

(check-expect (plays-game? "Wolfsburg" game1) #t)
(check-expect (plays-game? "Stuttgart" game1) #t)
(check-expect (plays-game? "Hannover" game1) #f)

(define plays-game?
  (lambda (team game)
    (or (string=? team (game-home-team game))
        (string=? team (game-guest-team game)))))
```

Nehmen wir an, wir wollen alle Spiele mit Beteiligung von Nürnberg extrahieren, indem wir `extract-list` verwenden. Um `extract-list` zu verwenden, brauchen wir eine Funktion mit Signatur `(game -> boolean)`. Die können wir folgendermaßen definieren:

```
; Spielt Nürnberg mit?
(: plays-nürnberg? (game -> boolean))

(check-expect (plays-nürnberg? game1) #f)
(check-expect (plays-nürnberg? game5) #t)

(define plays-nürnberg?
  (lambda (game)
    (plays-game? "Nürnberg" game)))
```

Damit können wir zum Beispiel alle Nürnberg-Spiele aus der Saison extrahieren:

```
(extract-list plays-nürnberg? season-2009/2010)
```

Soweit so gut. Aber was, wenn wir auch die Spiele des HSV extrahieren wollen? Wir könnten eine Funktion `plays-hamburg?` definieren, aber wenn wir viele solcher Anfragen stellen, wird es auf Dauer umständlich, für jeden Verein extra eine Definition zu schreiben.

Einfacher ist es, stattdessen die Funktion, die wir an `extract-list` übergeben, «ad hoc» mit einem `lambda` zu konstruieren:

```
(extract-list (lambda (game) (plays-game? "Nürnberg" game))
              season-2009/2010)
(extract-list (lambda (game) (plays-game? "Hamburg" game))
              season-2009/2010)
```

Wir sind von der Signatur her vorgegangen: `extract-list` erwartet eine Funktion mit einem Parameter der Signatur `game`, also schreiben wir ein `lambda` mit einem entsprechenden Parameter. Wir wollen alle Spiele extrahieren, bei denen Nürnberg beziehungsweise Hamburg dabei ist, also schreiben wir den entsprechenden Ausdruck in den Rumpf.

### AUFGABE 9.3

Schreibe eine Funktion `games-playing`, die alle Spiele aus einer Liste extrahiert, bei denen eine bestimmte Mannschaft dabei war. □

Die Funktion `extract-list` ist so praktisch, dass sie unter dem Namen `filter` fest eingebaut ist.

## 9.3 LISTEN UMWANDELN

Bestimmte Fußballstatistiken beschäftigen sich damit, wieviele Tore in einem Spiel, an einem Spieltag oder in einer ganzen Saison geschossen wurden. Grundlage dafür ist die Anzahl Tore eines einzelnen Spiels. Hier sind Kurzbeschreibung, Signatur, Tests und Gerüst dafür:

```
; Gesamttore eines Spiels berechnen
(: total-goals (game -> natural))

(check-expect (total-goals game1) 2)
(check-expect (total-goals game2) 4)
```

```
(define total-goals
  (lambda (game)
    ...))
```

Als Schablone machen wir davon Gebrauch, dass die Eingabe zusammengesetzte Daten sind:

```
(define total-goals
  (lambda (game)
    ... (game-matchday game) ...
    ... (game-home-team game) ...
    ... (game-home-goals game) ...
    ... (game-guest-team game) ...
    ... (game-guest-goals game) ...))
```

Für die Gesamt-Torzahl benötigen wir nur `(game-home-goals game)` und `(game-guest-goals game)`, die addiert werden:

```
(define total-goals
  (lambda (game)
    (+ (game-home-goals game)
       (game-guest-goals game))))
```

Als Nächstes wollen wir die Gesamttore jeweils aller Spiele eines Spieltags herausbekommen. Dazu wollen wir aus einer Liste der Spiele eine Liste der zugehörigen Gesamttore machen:

```
; Gesamttore in einer Liste von Spielen berechnen
(: list-total-goals ((list-of game) -> (list-of natural)))
```

Das erste Element des Ergebnisses soll die Toranzahl des ersten Spiels der Eingabeliste sein, das zweite Element die Toranzahl des zweiten Spiels undsoweiter. Dieser Test illustriert die Funktionsweise:

```
(check-expect (list-total-goals day1)
  (list 2 4 1 5 3 1 2 6 2))
```

Es folgt das Gerüst mit der obligatorischen Schablone für Funktionen auf Listen:

```
(define list-total-goals
  (lambda (list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ... (first list))
```

```
... (list-total-goals (rest list)) ...)))))
```

Falls die Eingabe eine leere Liste ist, so kann auch die Ausgabe nur leer sein. Im anderen Zweig steht in der Schablone das erste Spiel: von dem muss die Funktion die Toranzahl ermitteln, für den Rest ist das schon passiert. Die beiden Ergebnisse kombinieren wir mit `cons`:

```
(define list-total-goals
  (lambda (list)
    (cond
      ((empty? list) empty)
      ((cons? list)
       (cons (total-goals (first list))
             (list-total-goals (rest list)))))))
```

#### AUFGABE 9.4

Benutze die Funktion `list-sum` aus Abschnitt 6.1 auf Seite 168, um eine Funktion zu schreiben, welche die durchschnittliche Toranzahl einer Liste von Spielen berechnet. □

Fans sind vielleicht weniger an Gesamttoren interessiert als an den Toren ihrer Mannschaft. Damit wir uns da einen Überblick verschaffen können, benötigen wir wieder erst einmal eine Funktion, welche für ein einzelnes Spiel die Anzahl der Tore einer bestimmten Mannschaft berechnet:

```
; Tore einer Mannschaft aus einem Spiel berechnen
(: team-goals (team game -> natural))
```

```
(check-expect (team-goals "Wolfsburg" game1) 2)
(check-expect (team-goals "Stuttgart" game1) 0)
```

Was soll passieren, wenn die gewünschte Mannschaft bei dem Spiel gar nicht dabei war? Das sollte zu einem Fehler führen:

```
(check-error (team-goals "Hannover" game1))
```

Gerüst und Schablone machen zunächst wieder davon Gebrauch, dass es sich bei `game` zum zusammengesetzte Daten handelt:

```
(define team-goals
  (lambda (team game)
    ... (game-matchday game) ...
```

```
... (game-home-team game) ...
... (game-home-goals game) ...
... (game-guest-team game) ...
... (game-guest-goals game) ...))
```

Außerdem gibt es eine Fallunterscheidung in den Daten, je nachdem, ob die gesuchte Mannschaft die Heim- oder die Gastmannschaft ist, was zu einer Verzweigung mit zwei Zweigen führt:

```
(define team-goals
  (lambda (team game)
    (cond
      ((string=? team (game-home-team game))
       (game-home-goals game))
      ((string=? team (game-guest-team game))
       (game-guest-goals game))))))
```

Jetzt wollen wir auch diese Funktion benutzen, um eine ganze Liste von Spielen, bei denen eine bestimmte Mannschaft mitspielt, in die Liste der Tore dieser Mannschaft verwandelt. Signatur und Kurzbeschreibung sehen so aus:

```
; Tore einer Mannschaft aus einer Liste von Spielen auflisten
(: list-team-goals (team (list-of game) -> (list-of natural)))
```

#### AUFGABE 9.5

Schreibe einen Testfall, der eine Liste der Tore von Hamburg in einer Saison berechnet. Benutze dazu `filter`! □

Gerüst und Schablone sehen so aus:

```
(define list-team-goals
  (lambda (team list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ... (first list) ...
       ... (list-team-goals team (rest list)) ...))))
```



Und wieder kommt bei einer leeren Liste als Eingabe eine leere Liste heraus, und wir wenden im `cons`-Fall die zuvor geschriebene Hilfsfunktion auf das erste Element an und bauen mit `cons` die Ergebnis-Liste:

```
(define list-team-goals
  (lambda (team list)
    (cond
      ((empty? list) empty)
      ((cons? list)
       (cons (team-goals team (first list))
              (list-team-goals team (rest list)))))))
```

Fertig!

Allerdings ruft mal wieder das Abstraktions-Mantra: Die Funktionen `list-total-goals` und `list-team-goals` ähneln sich stark. Sie unterscheiden sich nur darin, dass bei der ersten `total-goals` auf dem ersten Element der Liste aufgerufen wird und beim zweiten `team-goals`.

Um zu abstrahieren, gehen wir wieder nach Konstruktionsanleitung 20 auf Seite 244 vor. Wir kopieren eine der beiden Funktionen (wir nehmen die erste) und geben ihr einen neuen Namen:

```
(define list-apply
  (lambda (list)
    (cond
      ((empty? list) empty)
      ((cons? list)
       (cons (total-goals (first list))
              (list-apply (rest list)))))))
```

Die Funktion heißt `list-apply`, weil sie eine Funktion auf alle Elemente einer Liste anwendet.

Als Nächstes müssen wir eine Variable dort einführen, wo die beiden Funktionen sich unterscheiden, nämlich bei `total-goals`. Wir nennen die Funktion einfach `f`:

```
(define list-apply
  (lambda (list)
    (cond
      ((empty? list) empty)
      ((cons? list)
       (cons (f (first list))
              (list-apply (rest list)))))))
```

Nun müssen wir die neue Variable noch im `lambda` unterbringen, ohne zu vergessen, sie auch im rekursiven Aufruf aufzuführen:

```
(define list-apply
  (lambda (f list)
    (cond
      ((empty? list) empty)
      ((cons? list)
       (cons (f (first list))
              (list-apply f (rest list)))))))
```

#### AUFGABE 9.6

Ändere die Definition von `list-total-goals` so, dass sie `list-apply` benutzt. □

Wir brauchen noch eine Signatur für die neue Funktion. Wir versuchen sie mal, «von null» zu konstruieren, indem wir aufschreiben, was wir wissen. Wir wissen, dass `list-apply` zwei Argumente hat, von denen das erste eine einstellige Funktion und das zweite eine Liste ist. Außerdem produziert die Funktion auch eine Liste als Ergebnis. Das sieht schriftlich so aus:

```
(: list-apply ((... -> ...) (list-of ...) -> (list-of ...)))
```

Wir haben zwar über zwei Funktionen abstrahiert, die sich mit Spielen beschäftigt. Aber `list-apply` hat nichts Fußball-spezifisches mehr an sich. Wir wissen also nichts über die Signatur der Elemente der Eingabeliste. Wir schreiben dort also eine Signaturvariable hin:

```
(: list-apply ((... -> ...) (list-of %element) -> (list-of ...)))
```

Nun wissen wir, dass die Elemente der Eingabeliste als Eingabe für die Funktion `f` verwendet werden, wir können also auch dort `%element` hinschreiben:

```
(: list-apply ((%element -> ...) (list-of %element) -> (list-of ...)))
```

Die Definition von `list-apply` legt auch für das Ergebnis von `f` keine Signatur fest. Wir schreiben also auch da eine Signaturvariable hin:

```
(: list-apply ((%element -> %result) (list-of %element) -> (list-of
...)))
```

Es fehlt nur noch die Signatur für die Elemente der Ergebnisliste. Dafür müssen wir den Aufruf von `cons` betrachten, dessen erstes Argument gerade das Ergebnis des Aufrufs von `f` ist. Dort muss deshalb auch `%result` hin:

```
(: list-apply ((%element -> %result) (list-of %element) -> (list-of %
  result)))
```

Bei solchen Signaturen, in denen überhaupt nichts über die Signaturvariablen bekannt ist, ist es üblich, kürzere Namen für die Signaturvariablen zu wählen, zum Beispiel so:

```
(: list-apply ((%a -> %b) (list-of %a) -> (list-of %b)))
```

Fertig!

#### AUFGABE 9.7

Schreibe auch `list-team-goals` so um, dass es `list-apply` verwendet. Du musst einen ähnlichen Trick verwenden wie bei der Kombination von `extract-list` und `plays-game?` im vorigen Abschnitt auf Seite 251. □

Die Funktion `list-apply` ist ähnlich praktisch wie `filter` aus dem vorigen Abschnitt und ist darum unter dem Namen `map` fest eingebaut.

## 9.4 LISTEN ZUSAMMENFALTEN

In Abschnitt 6.1 haben wir auf Seite 168 die Funktionen `list-sum` und `list-product` geschrieben, welche die Summe einer Liste von Zahlen bildet:

```
; Summe der Elemente einer Liste von Zahlen berechnen
```

```
(: list-sum ((list-of number) -> number))
```

```
(define list-sum
```

```
  (lambda (list)
```

```
    (cond
```

```
      ((empty? list) 0)
```

```
      ((cons? list)
```

```
        (+ (first list) (list-sum (rest list))))))
```

```
; Produkt der Elemente einer Liste von Zahlen berechnen
```

```
(: list-product (list-of-numbers -> number))
```

```
(define list-product
  (lambda (list)
    (cond
      ((empty? list) 1)
      ((cons? list)
       (* (first list) (list-product (rest list)))))))
```

Die Definitionen von `list-sum` und `list-product` unterscheiden sich, bis auf den Namen, nur an zwei Stellen: im Zweig für `empty`, wo das jeweilige neutrale Element steht, und im Zweig für `cons`, wo die Funktion steht, die benutzt wird, um das erste Element mit dem Ergebnis des rekursiven Aufrufs zu kombinieren. Es ist also eine gute Idee zu abstrahieren!

Wir gehen wieder nach Konstruktionsanleitung 20 auf Seite 244 vor. Uns ist beim ersten Mal, als wir das gemacht haben, kein guter Name eingefallen. Das ist keine Schande: Wir fangen deshalb mit `xxx` an in der Hoffnung, dass uns später ein besserer Name einfällt:

```
(define xxx
  (lambda (list)
    (cond
      ((empty? list) 1)
      ((cons? list)
       (* (first list) (xxx (rest list)))))))
```

Als Nächstes müssen wir die Unterschiede zwischen den beiden Funktionen durch neue Variablen ersetzen, das sind hier gerade `1` und `*`. Die neuen Variablen nennen wir `for-empty` und `for-cons`, weil sie in den Zweigen für `empty` respektive `cons` stehen:

```
(define xxx
  (lambda (list)
    (cond
      ((empty? list) for-empty)
      ((cons? list)
       (for-cons (first list) (xxx (rest list)))))))
```

Für den nächsten Schritt fügen wir die neuen Variablen zum bestehenden `lambda` hinzu. Wir müssen wieder darauf achten, die Variablen auch zum rekursiven Aufruf hinzuzufügen:

```
(define xxx
  (lambda (for-empty for-cons list)
    (cond
```

```

((empty? list) for-empty)
((cons? list)
 (for-cons (first list)
           (xxx for-empty for-cons (rest list))))))

```

Diese Funktion kann sowohl die Aufgabe von `list-sum` als auch die von `list-product` erledigen:

```

(xxx 0 + (list 1 2 3 4))
↪ 10

```

und so aufmultiplizieren:

```

(xxx 1 * (list 1 2 3 4))
↪ 24

```

Als Nächstes steht nach Konstruktionsanleitung die Signatur für die neue Funktion an. Es liegt nahe, die Signatur von `list-sum` beziehungsweise `list-product` zu verallgemeinern. Für den neuen Parameter `for-empty` verwenden wir die Signatur von 0 und 1, also `number`. Für `for-cons` verwenden wir die Signatur von + und \*, also `(number number -> number)`:

```

(: xxx (number (number number -> number) (list-of number) -> number))

```

Bei näherer Betrachtung steht aber in der Definition von `xxx` allerdings nichts «zahlenspezifisches», wir könnten also versuchen, `number` durch eine Signaturvariable zu ersetzen, wie wir es auch bei `list-apply` beziehungsweise `map` im vorigen Abschnitt gemacht haben:

```

(: xxx (%a (%a %a -> %a) (list-of %a) -> %a))

```

Das heißt, wir könnten versuchen, `xxx` mit etwas aufzurufen, das keine Liste von Zahlen ist. Wir brauchen dafür ein passendes zweites Argument, also eine Funktion, deren Signatur die Form `(%a %a -> %a)` hat. Ein Beispiel ist die Funktion `append`, die zwei Listen aneinanderhängt und folgende Signatur hat:

```

((list-of %element) (list-of %element) -> (list-of %element))

```

Weil wir `(list-of %element)` für `%a` einsetzen, brauchen wir als weitere Argumente für `xxx` noch eine Liste als erstes Argument und eine Liste von Listen als drittes Argument. Das könnte so aussehen:

```

(xxx empty append (list (list 1 2 3) (list 4 5 6) (list 7 8 9)))
↪ #<list 1 2 3 4 5 6 7 8 9>

```

Wir können also `xxx` benutzen, um eine Liste von Listen «flachzuklopfen».

Aber ist das schon die beste Signatur für `xxx`? Wir können wie bei `list-map` untersuchen, ob die `%a`s tatsächlich alle gleich sein müssen. Um das herauszubekommen, könnten wir mal annehmen, sie seien alle verschieden:

```
(: xxx (%a (%b %c -> %d) (list-of %e) -> %f))
```

Wenn wir aber die Definition von `xxx` genau anschauen, dann sehen wir, dass jeweils einige Signaturvariablen gleich sein müssen:

- Die Signatur `%a` von `for-empty` muss die gleiche sein wie `%f`, weil `for-empty` als Ergebnis von `xxx` verwendet wird.
- Die Signatur `%b` des ersten Arguments von `for-cons` muss die gleiche sein wie die der Listenelemente `%a`, weil als erstes Argument (`first list`) verwendet wird.
- Die Signatur `%c` des Arguments von `for-cons` muss die gleiche sein wie `%f`, weil das Ergebnis von `xxx` dort verwendet wird.
- Die Signatur `%d` des Ergebnisses von `for-cons` muss ebenfalls die gleiche sein wie `%f`, weil der Aufruf von `for-cons` als Ergebnis von `xxx` verwendet wird.

Aber immerhin müssen `%a` und `%b` nicht gleich sein. Es bleibt:

```
(: xxx (%a (%b %a -> %a) (list-of %b) -> %a))
```

## AUFGABE 9.8

Verfolge die Auswertung von

```
(xxx 0 + (list 1 2 3))
```

im Stepper!



Im Stepper kannst Du sehen, dass `xxx` folgendermaßen arbeitet:

```
(xxx e c #<list a1 ... an >)
↔...↔
(c a1 (c a2 (... (c an u) ...)))
```

Im Fall von `0` und `+` wird die Liste dann zu einem einzelnen Wert reduziert oder «zusammengefasst», weswegen wir die Funktion `list-fold` nennen.

Die Funktionsweise von `list-fold` können wir daran veranschaulichen, dass sich die ursprüngliche Liste auch als

```
(cons a1 (cons a2 (... (cons an empty ...)))
```

schreiben lässt. Das heißt, an die Stelle von `cons` tritt `c` und an die Stelle von `empty` tritt `e`.

#### AUFGABE 9.9

Kannst Du daraus ableiten, was hierbei herauskommt, ohne DrRacket zu benutzen?

```
(list-fold empty cons (list 1 2 3))
```

□

Eine andere, praktische Darstellung von `list-fold` ist, die Gleichung mit dem Operator *zwischen* den Operanden zu schreiben (und nicht davor). Um das deutlich zu machen, schreiben wir für `for-cons` nicht `c` sondern  $\odot$ :

```
(list-fold c  $\odot$  #<list a1 ... an>)
```

$\hookrightarrow \dots \hookrightarrow$

```
a1  $\odot$  (a2  $\odot$  (... (an  $\odot$  e) ...))
```

Die Funktion `list-fold` ist erstaunlich vielseitig und kann die Aufgaben vieler anderer Funktionen auf Listen erledigen. Zum Beispiel können wir uns `list-apply` noch einmal vorknöpfen und versuchen, die Definition mit `list-fold` zu schreiben. Hier zur Erinnerung die bisherige Fassung:

```
(define list-apply
  (lambda (f list)
    (cond
      ((empty? list) empty)
      ((cons? list)
       (cons (f (first list))
              (list-apply f (rest list)))))))
```

Wir wollen `list-apply` in folgende Form umschreiben:

```
(define list-apply
  (lambda (f list)
    (list-fold ...
      ...
      list)))
```

Wir müssen passende Werte für die beiden Parameter von `list-fold` finden, `for-empty` und `for-cons`. In `list-apply` steht im `empty`-Zweig `empty`.

Bei `for-cons` ist es nicht so offensichtlich. Wir können uns aber an der Signatur orientieren. Gefragt ist eine Funktion mit zwei Argumenten: das erste ist das erste Listenelement, das zweite das Resultat des rekursiven Aufrufs. Wir können also so anfangen:

```
(define list-apply
  (lambda (f list)
    (list-fold empty
      (lambda (first-list result)
        ...)
      list)))
```

Im Rumpf des neuen `lambda` müssen wir nun Code hinschreiben, der dem Code aus `list-apply` im `cons`-Zweig entspricht, also:

```
(cons (f (first list))
      (list-apply f (rest list)))
```

Allerdings müssen wir die Parameter aus dem `lambda` benutzen, also `first-list` und `result`: `first-list` ist gerade der Wert von `(first-lists)`, `result` das Resultat des rekursiven Aufrufs. Wir ersetzen also «rückwärts»:

```
(define list-apply
  (lambda (f list)
    (list-fold empty
      (lambda (first-list result)
        (cons (f first-list)
              result))
      list)))
```

Fertig!

#### AUFGABE 9.10

Schreibe auf die gleiche Weise `extract-list` mit Hilfe von `list-fold` um!



Die Funktion `list-fold` ist so praktisch, dass auch sie fest eingebaut ist, und zwar unter dem Namen `fold`.



## 9.5 SCHÖNFINKELN

Vielleicht erinnerst Du Dich an die Funktion `plays-game?` auf Seite 250. Wir haben sie geschrieben, um mit Hilfe von `extract-list` Spiele zu extrahieren, bei denen eine bestimmte Mannschaft mitspielt. Allerdings konnten wir `plays-game?` nicht direkt mit `extract-list` verwenden, weil `extract-list` eine einstellige Funktion akzeptiert, `plays-game?` aber zweistellig ist. Wir mussten `lambda` verwenden, um eine passende Funktion zu konstruieren. Zur Erinnerung, das sah so aus:

```
(extract-list (lambda (game) (plays-game? "Nürnberg" game))
              season-2009/2010)
(extract-list (lambda (game) (plays-game? "Hamburg" game))
              season-2009/2010)
```

Die beiden `lambda`s sind bis auf den Namen der Mannschaft fast identisch, das schreit mal wieder nach Abstraktion! Dafür brauchen wir eine Funktion, die solche `lambda`s wie oben erzeugt. Wenn wir eins davon kopieren und gerade die unterschiedliche Stelle durch eine Variable ersetzen, sieht das so aus:

```
(lambda (game) (plays-game? team game))
```

Jetzt müssen wir noch `team` irgendwo binden. Es ist nicht sinnvoll, es zum existierenden `lambda` hinzuzufügen – dann hätten wir wieder eine zweistellige Funktion. Stattdessen machen wir ein `lambda` drumherum und geben dem ganzen einen Namen:

```
(define plays-game?/team
  (lambda (team)
    (lambda (game) (plays-game? team game))))
```

(Der Schrägstrich gehört ganz normal zum Namen.)

Diese Funktion akzeptiert eine Mannschaft und liefert wieder eine Funktion, die für ein Spiel einen booleschen Wert liefert. Sie hat also folgende Signatur:

```
(: plays-game?/team (team -> (game -> boolean)))
```

Jetzt können wir die obigen Aufrufe von `extract-list` mit Hilfe von `plays-game?/team` umschreiben:

```
(extract-list (plays-game?/team "Nürnberg")
              season-2009/2010)
(extract-list (plays-game?/team "Hamburg")
              season-2009/2010)
```

Eine ähnliche Konstruktion können wir für `team-goals` durchführen, um es zum Beispiel einfacher mit `list-apply` zu verwenden. Dabei kommt folgende Definition heraus:

```
(: team-goals/team (team -> (game -> natural)))
(define team-goals/team
  (lambda (team)
    (lambda (game)
      (team-goals team game))))
```

Wenn Du genau hinschaust, kannst Du sehen, dass wir schon wieder abstrahieren können, nämlich über `plays-game?/team` und `team-goals/team` – die sind nämlich identisch bis auf die Funktion, die am Ende aufgerufen wird. Wir kopieren also wieder und ersetzen diese Funktion durch die neue Variable `f`. Weil wir noch nicht absehen können, wo die Reise hingeht, nennen wir die Funktion erstmal wieder `xxx`:

```
(define xxx
  (lambda (team)
    (lambda (game)
      (f team game))))
```

Für den neuen Namen machen wir wieder ein neues `lambda`:

```
(define xxx
  (lambda (f)
    (lambda (team)
      (lambda (game)
        (f team game))))))
```

Wir brauchen eine Signatur. Da drei `lambdas` involviert sind, brauchen wir auch drei Pfeile:

```
(: xxx (... -> (... -> (... -> ...)))
```

Außerdem ist das erste Argument für `f` eine Funktion, die zwei Argumente akzeptiert:

```
(: xxx ((... ... -> ...) -> (... -> (... -> ...)))
```

Viele Pfeile – lass Dich nicht aus der Ruhe bringen!

Wir abstrahieren zwar über zwei Funktionen, die sich mit Fußball beschäftigen, aber *schon wieder* ist es so, dass die entstandene Funktion nichts fußball-spezifisches hat. Wir versuchen es deshalb erst einmal wieder mit Typvariablen und fangen mit der Signatur von `f` an. Wir nehmen einfach mal drei verschiedene Variablen – gegebenenfalls finden wir später heraus, dass einige davon gleich sein müssen:

```
(: xxx ((%a %b -> %c) -> (... -> (... -> ...))))
```

Innerhalb von `xxx` wird `f` aufgerufen mit `team` und `game`, die zu den Signaturen `%a` respektive `%b` gehören. Die weitere Analyse wird darum vereinfacht, wenn wir `team` in `a` und `game` in `b` umbenennen:

```
(define xxx
  (lambda (f)
    (lambda (a)
      (lambda (b)
        (f a b)))))
```

Mit diesen Namen wir den Rest der Signatur ablesen:

```
(: xxx ((%a %b -> %c) -> (%a -> (%b -> %c))))
```

Wie sollte die Funktion heißen? Tatsächlich ist sie berühmt und geht auf Arbeiten der Mathematiker MOSES SCHÖNFINKEL und HASKELL CURRY zurück. Im englischsprachigen Raum heißt das Verb dazu darum *curry*, im deutschsprachigen Raum *schönfinkeln* oder *curryfizieren*. Hier also die finale Version:

```
; Funktion schönfinkeln
(: curry ((%a %b -> %c) -> (%a -> (%b -> %c))))
```

```
(define curry
  (lambda (f)
    (lambda (a)
      (lambda (b)
        (f a b)))))
```

Hier ist ein Testfall, der `curry` zusammen mit `plays-game?` verwendet:

```
(check-expect (list-apply ((curry team-goals) "Hamburg")
                          (filter ((curry plays-game?) "Hamburg")
                                season-2009/2010))
              (list 1 4 4 3 3 1 1 3 0 3 2 2 0 1 0 4 2 2 0 1 3 3 0 0 1
                    2 2 0 0 2 0 1 4 1))
```

#### AUFGABE 9.11

Schreibe ein Beispiel, das `curry` zusammen mit `team-goals` verwendet!



Die `curry`-Funktion macht aus einer zweistelligen eine einstellige Funktion, die aber im Prinzip das gleiche macht. Um die zweistellige Funktion aufzurufen, muss das Programm beide Argumente auf einmal liefern, im anderen Fall kommen sie nacheinander – aber das Ergebnis ist am Ende das gleiche.

Wenn wir aus einer zweistelligen Funktion eine einstellige machen können (also eine einstellige Funktion, die eine einstellige Funktion liefert), geht das auch umgekehrt? Na klar! Dazu drehen wir die Signatur gegenüber `curry` um, indem wir die Teile vor und hinter dem Pfeil in der Mitte vertauschen:

```
; Funktion entschönfinkeln
(: uncurry ((%a -> (%b -> %c)) -> (%a %b -> %c)))
```

Als Test lassen wir `uncurry` auf das Resultat von `curry` los:

```
(check-expect ((uncurry (curry +)) 3 4) 7)
```

Für die Konstruktion orientieren wir uns wie bei `curry` an der Signatur: `curry` akzeptiert eine Funktion und liefert eine zweistellige Funktion. Wir fangen deshalb folgendermaßen an:

```
(define uncurry
  (lambda (f)
    (lambda (a b)
      ...)))
```

Die innere Funktion soll einen Wert mit Signatur `%c` produzieren. Um den zu bekommen, müssen wir zunächst `f` mit einem Wert der Signatur `%a` aufrufen – den haben wir mit `a`:

```
(define uncurry
  (lambda (f)
    (lambda (a b)
      ... (f a) ...)))
```

Der Aufruf `(f a)` liefert wiederum eine Funktion mit Signatur `(%b -> %c)`, wir müssen also noch einmal Klammern drummachen und ein `%b` liefern – das ist in diesem Fall `b`:

```
(define uncurry
  (lambda (f)
    (lambda (a b)
      ((f a) b))))
```

Fertig!

Damit ist die Transformation ein *Isomorphismus*; es gilt folgende Gleichung für Funktionen  $f$  mit zwei Parametern:

$$(\text{uncurry } (\text{curry } f)) = f$$

#### AUFGABE 9.12

Versuche, Funktionen `curry3` und `uncurry3` zu schreiben. Sie sollen wie `curry` und `uncurry` arbeiten, aber für Funktionen mit drei statt nur mit zwei Eingaben. □

## 9.6 LEXIKALISCHE BINDUNG

In diesem Abschnitt schreiben wir ausnahmsweise keine neuen Funktionen, sondern untersuchen einen subtilen Aspekt des Verhaltens von Funktionen: den Umgang mit Variablen. Dir ist sicher schon aufgefallen, dass wir in unterschiedlichen Funktionen oft die gleichen Namen verwenden, zum Beispiel hier:

```
(define feed-dillo
  (lambda (dillo)
    ...))
```

```
(define run-over-dillo
  (lambda (dillo)
    ...))
```

Obwohl der Parameter in beiden Funktionen `dillo` heißt, haben die beiden Parameter in Wirklichkeit nichts miteinander zu tun. Das kannst Du im Stepper nachvollziehen, wenn Du bei der Auswertung zusiehst. Du kannst es aber auch einfach am Programm sehen. Um der Sache auf den Grund zu gehen, betrachten wir folgendes Programm:

```
(define pi 3.14159265)

; Fläche eines Kreises ausrechnen
(: circle-area (real -> real))

(check-expect (circle-area 5) 78.53981625)

(define circle-area
```

```
(lambda (radius)
  (* pi radius radius)))
```

In dem Programm interessieren uns die Variablen `pi` und `radius`, die mehrmals auftauchen – man spricht darum auch von unterschiedlichen *Vorkommen* dieser Variablen.

Die Auswertung des Tests führt dazu, dass im Rumpf des `lambda` von `circle-area` die Variable `radius` ersetzt wird. Dieser Vorgang heißt auch *Bindung* und das ist so, weil `radius` als Parameter im `lambda` steht. Dieses Vorkommen als Parameter heißt auch *Bindung* der Variable. Die beiden Vorkommen von `radius` im Rumpf `(* pi radius radius)`, die ersetzt werden, heißen *gebundene Vorkommen*. Auch das `pi` im Rumpf von `circle-area` wird ersetzt und ist darum ein gebundenes Vorkommen. Das dazugehörige bindende Vorkommen steht im `define` obendrüber.

Je nachdem, ob die Bindung mit `define` oder `lambda` passiert, verhält sie sich unterschiedlich. Das kannst Du sehen, wenn Du das obige Programm laufen lässt und in der REPL `radius` und `pi` auswertest:

```
> radius
radius: Variable ist nicht definiert
> pi
3.14159265
```

Das heißt, `pi` können wir überall benutzen: Man sagt auch, dass `pi` überall im Programm *sichtbar* ist. Es handelt sich bei `pi` um eine *globale Variable*. Globale Variablen werden auch von `define-record` gebunden.

Im Gegensatz dazu ist `radius` nur im Rumpf des `lambda` sichtbar, dessen Parameter es ist – es ist eine sogenannte *lokale Variable*.

Lokale Variablen entstehen außerdem durch lokale Definitionen wie zum Beispiel bei der Funktion `home-points` in Abschnitt 9.1 auf Seite 239. Wenn eine lokale Definition am Anfang eines `lambda`-Rumpfes steht, ist sie in dem Rumpf sichtbar. Wenn sie im Zweig eines `cond` steht, ist sie nur in diesem Zweig sichtbar.

Die Beziehung zwischen den Bindung und gebundenen Vorkommen kannst Du in DrRacket visualisieren, indem Du auf den Knopf *Syntaxprüfung* (beziehungsweise *Check Syntax*) drückst und dann mit dem Mauszeiger über Variablenvorkommen streichst. Abbildung 9.1 zeigt exemplarisch, wie das aussieht.

#### AUFGABE 9.13

Probiere die *Syntaxprüfung* auf dem Fußball-Code aus und untersuche zum Beispiel die Variablen in der Funktion `home-points`. □

```

(define pi 3.14159265)

; Fläche eines Kreises ausrechnen
(: circle-area (real -> real))

(check-expect (circle-area 5) 78.53981625)

(define circle-area
  (lambda (radius)
    (* pi radius radius)))

```

2 gebundene Vorkommen

**Abbildung 9.1:** Anzeige von Bindung nach Syntaxprüfung

Eine dritte Kategorie von Bindung gibt es auch noch, nämlich die «eingebauten» Variablen wie `string-append` und `+`, die ähnlich global sichtbar sind wie die globalen Variablen. Wenn Du über sie nach Syntaxprüfung streichst, wird `importiert` aus angezeigt.

Damit könnte man meinen, es sei alles zur Bindung gesagt: Globale und eingebaute Variablen sind überall sichtbar; eine lokale Variable ist innerhalb ihres `lambda`-Rumpfes beziehungsweise `cond`-Zweiges sichtbar. Aber was ist mit der folgenden Funktion?

```

(define f
  (lambda (x)
    (define y (+ x 1))
    (lambda (x)
      (+ x y))))

```

Die ist zugegeben künstlich konstruiert. Aber wie verhält sie sich?

#### AUFGABE 9.14

Was liefert die Auswertung von `((f 1) 2)`?



In dieser Funktion gibt es *zwei* Variablen namens `x`: Zwei Bindungen als Parameter und auch zwei gebundene Vorkommen. Welche gebundenen Vorkommen gehören zu welchen Bindungen? Hier gilt die Regel der sogenannten *lexikalischen Bindung*:

#### DEFINITION 9.1 (LEXIKALISCHE BINDUNG)

Du findest die zu einem gebundenen Vorkommen zugehörige Bindung, indem Du von innen nach außen in der Klammerstruktur der Funktion suchst: Die erste Bindung als Parameter oder lokale Definition, die Du so findest, gehört zu dem gebundenen Vorkommen.

Wenn Du auf diese Art und Weise keine Bindung findest, halte Ausschau nach einer globalen Definition. Wenn Du keine globale Definition findest, muss es sich um eine eingebaute beziehungsweise importierte Definition handeln.

Die lexikalische Bindung ordnet einer Bindung ihre gebundenden Vorkommen zu. Das bedeutet, dass Du, wenn Du eine Bindung und all ihre gebundenen Vorkommen unbenennst, das Verhalten des Programms nicht änderst. Entsprechend sind also Variablen, die zu unterschiedlichen Bindungen gehören, unterschiedliche Variablen. Wenn Du nach einer Syntaxprüfung über einer Variable das Kontextmenü («rechte Maustaste») aufrufst, gibt es dort einen Punkt **umbenennen**, der das Umbenennen automatisiert.

Der Begriff «lexikalische Bindung» suggeriert, dass es noch andere Formen der Bindung gibt: Tatsächlich funktioniert Bindung in einigen anderen, vornehmlich älteren Sprachen, nach anderen Prinzipien, oft unter dem Begriff *dynamische Bindung* zusammengefasst. Lexikalische Bindung wird oft auch *statische Bindung* genannt. «Statisch» bedeutet im Kontext der Programmierung «bevor das Programm läuft» und statische Bindung bedeutet, dass Du die Beziehung zwischen Bindungen und gebundenen Vorkommen am Programmtext ablesen kannst. Bei dynamischer Bindung ergibt sich diese Beziehung erst, wenn das Programm läuft.

Die Definition von **f** kommt Dir vielleicht esoterisch vor. Tatsächlich haben wir aber schon oft Programme geschrieben, bei denen mehrere Variablen gleichen Namens auftauchen, wie zum Beispiel **home-points**:

```
(: home-points (game -> points))
...
(define home-points
  (lambda (game)
    ...))
```

Hier gibt es zwei Variablen namens **game**: In der Signaturdeklaration ist das die Signatur **game** aus der gleichnamigen Record-Definition. In der Funktion ist es der Parameter **game**. Dieser erfüllt die Signatur **game** – darum haben die Signatur und der Parameter ja gerade den gleichen Namen, um diese Beziehung deutlich zu machen.

Man könnte meinen, *jetzt* wäre alles gesagt zur Bindung. Aber es gibt noch eine weitere Besonderheit, nämlich wenn in einer Definition (gleich ob lokal oder global) die Variable, die definiert wird, auch gleich wieder auftaucht. Auch davon haben wir schon ständig Gebrauch gemacht, nämlich in rekursiven Funktionen, wo für den rekursiven Aufruf genau die Variable benutzt wird, in deren Definition sie sich befindet.



## AUFGABE 9.15

Zu welchen Bindungen gehören die gebundenen Vorkommen von `x` in folgendem Programm?

```
(define g
  (lambda (x)
    (define x (+ x 1))
    (define y (+ x 1))
    (lambda (x)
      (+ x y))))
```

Kannst Du vorhersagen – ohne es in der REPL auszuprobieren – was `((g 1) 2)` liefert?



## AUFGABEN

## AUFGABE 9.16

Schreibe eine Funktion `any?`, die eine Liste akzeptiert sowie eine Funktion, die für jedes Listenelement entweder `#t` oder `#f` liefert. `Any?` soll `#t` zurückgeben, wenn mindestens die Funktion für mindestens ein Element der Liste `#t` zurückgibt, sonst `#f`.

Schreibe eine analoge Funktion `every?`, die dann `#t` zurückgibt, wenn die Funktion für alle Elemente der Liste `#t` zurückgibt, sonst `#f`. Schreibe zunächst eine Fassung nach dem Muster von `any?`. Schreibe eine zweite Fassung, die einfach `any?` aufruft und selbst keine Rekursion benutzt.

## AUFGABE 9.17

Schreibe Funktionen `feed-animals` und `run-over-animals` die mit Hilfe der Funktionen in Abschnitt 4.2 auf Seite 115 alle Tiere auf dem Highway (die sich in einer Liste befinden) füttert respektive überfährt.

## AUFGABE 9.18

Schreibe folgende Funktionen unter der Verwendung von `filter`!

- Schreibe eine Funktion `evens`, welche die ungeraden Zahlen aus einer Liste entfernt,
- eine Funktion `count-zeroes`, die in einer Liste von Zahlen die Nullen zählt,
- und eine Funktion `multiples`, die eine Zahl `n` und eine Liste von Zahlen akzeptiert, und eine Liste alle Vielfachen der Zahl `n` liefert.

## AUFGABE 9.19

Programmiere eine Funktion `filter-map`, die als Argumente eine Funktion `p` mit einem Argument sowie eine Liste `l` akzeptiert. `Filter-map` soll als Ergebnis die Liste der Rückgabewerte von `p` für diejenigen Elemente von `l` zurückgeben, für die `p` nicht `#f` zurückgibt. Beispiel:

```
(filter-map (lambda (x)
              (if (even? x)
                  (+ x 1)
                  #f))
            (list 1 2 5 17 24 13))
⇒ #<list 3 25>
```

## AUFGABE 9.20

Verwende die Funktion `list-fold` um folgende Funktionen zu schreiben:

1. Schreibe eine Funktion `list-map`, die eine Funktion auf jedes Element einer Liste anwendet.
2. Schreibe eine Funktion `list-or`, die eine Funktion auf alle Elemente einer Liste anwendet, die jeweils ein Boolean liefert, und die Resultate mit `or` verknüpft.
3. Schreibe eine Funktion `count-trues`, die eine Funktion mit boolescher Rückgabe auf alle Elemente einer Liste anwendet und zählt, wie häufig `#t` zurückgegeben wird.
4. Schreibe eine Funktion `contains?`, die feststellt, ob ein Element in einer Liste enthalten ist.
5. Schreibe eine Funktion `remove-duplicates`, die alle doppelten Elemente aus einer Liste filtert.

## AUFGABE 9.21

Fußballreporter zitieren gern obskure Fakten über vergangene Spiele, wenn sie zum Spielverlauf wenig zu sagen haben.

Hier ein Beispiel:

«In der Saison 2009/2010 gab es ja immerhin vier Spieltage, an denen die größte Tordifferenz kleiner als 3 war.»

1. Stelle dem Fußballreporter Hilfsfunktionen zur Verfügung, die ihm erlauben, eine Antwort auf solch eine Frage zu finden:

*An wie vielen Spieltagen der Saison 2009/2010 war die größte Tordifferenz kleiner als 3?*

Schreibe mit Hilfe dieser Funktionen einen Ausdruck, der testet, ob die am Anfang dieser Aufgabe zitierte Aussage über die Tordifferenzen in der Saison 2009/2010 stimmt!

- (a) Schreibe eine Funktion, die aus einer Liste die Dubletten entfernt. Diese Funktion akzeptiert eine Liste sowie eine Funktion, die zwei Elemente der Liste akzeptiert und zurückliefert, ob diese gleich sind. (Also zum Beispiel = bei Listen von Zahlen.) Sie liefert dann eine Liste, in der jedes Element «nur einmal vorkommt», also kein anderes, das dazu gleich ist.
  - (b) Schreibe eine Funktion, die aus der Liste aller Spiele eine Liste aller Spielplannummern extrahiert.
  - (c) Schreibe eine Funktion, die aus der Liste aller Spiele eine Liste von Listen aller Spiele macht – so dass in jeder Teilliste alle Spiele eines Spieltags zusammengefasst sind.
  - (d) Schreibe eine Funktion, welche die Tordifferenz eines Spiels zurückliefert.
  - (e) Schreibe eine Funktion, welche das maximale Element einer Liste berechnet. Die Funktion sollte neben der Liste eine Funktion akzeptieren, die berechnet, ob ein Element «kleiner oder gleich» einem anderen ist.
  - (f) Schreibe schließlich einen Ausdruck, der die Anzahl der Spieltage der Saison 2009/2010 liefert, bei denen die größte Tordifferenz kleiner als 3 war.
2. Schreibe Ausdrücke, die folgende Fragen beantworten – entwickle dazu, falls nötig, weitere Hilfsfunktionen:
- An welchem Spieltag gab es die meisten Heimsiege?
  - An welchem Spieltag fielen die meisten Tore?
  - Gab es mehr Siege für die Heimmansschaften an ungeraden Spieltagen als an geraden?

## AUFGABE 9.22

Betrachte das folgende Programm:

```
(define x 2)           ; --> zwei
(define y -1)          ; --> minuseins
(define z -3)          ; --> minusdrei

(define f
  (lambda (x z)
    (+ (* x x) z y)))

(f 4 -2)
```

Benenne die Variablen `x`, `y` und `z`, die in den ersten drei Zeilen des Programms definiert werden, im kompletten Programm um, und zwar `x` in `zwei`, `y` in `minuseins` und `z` in `minusdrei`. Achte bei der Umbenennung auf die lexikalische Bindung. Benenne keine Parameter der Funktion `f` um.

Nachdem Du die Umbenennung durchgeführt hast, welches Ergebnis liefert der Ausdruck `(f 4 -2)`? Berechne das Ergebnis von Hand mit Hilfe des Substitutionsmodells und halte die Zwischenschritte fest.

#### AUFGABE 9.23

Betrachte das folgende Programm:

```
(define x 2)                ; --> zwei
(define y 4)                ; --> vier

(define z                    ; --> f
  (lambda (x y z)
    (+ x (z y))))

(z y x (lambda (z) (+ x z)))
```

Benenne die Variablen `x`, `y` und `z`, die in den ersten drei Zeilen des Programms definiert werden, im kompletten Programm um. Der neue Name der Variable steht als Kommentar im Programm hinter dem Pfeil (`-->`). Achte bei der Umbenennung auf die lexikalische Bindung. Benenne keine Parameter der Funktion `z` um.

Berechne, nachdem Du die Umbenennung durchgeführt haben, von Hand `(z y x (lambda (z) (+ x z)))` und halte die Zwischenschritte fest.

#### AUFGABE 9.24

Betrachte folgendes Programm:

```
(define x 1)
(define y 3)
(define z 5)

(define f
  (lambda (x)
    ((lambda (y)
```

```

((lambda (z)
  (+ z (* x y)))
 (+ x z)))
(+ x y)))

```

```
(f y)
```

Benenne hier alle lokalen Variablen, die innerhalb der Funktion `f` gebunden werden, um. Verändere nicht den Namen der Variablen `x`, `y` und `z` aus den ersten drei Zeilen des Programms.

Nachdem Du die Umbenennung durchgeführt haben, welches Ergebnis liefert der Ausdruck `(f y)`? Berechne das Ergebnis von Hand mit Hilfe des Substitutionsmodells und halte die Zwischenschritte fest.

*Hinweis:* In diesen Aufgaben findest Du keine Kommentare und Signaturen zu den Funktionen. Hier kannst Du an einem Beispiel sehen, dass es sehr wichtig ist, diese Informationen anderen Programmierern immer zur Verfügung zu stellen. Denn es kann auch bei kleinen Programmen schwer sein, die Funktionsweise der einzelnen Funktionen ohne Kommentare zu verstehen.

#### AUFGABE 9.25

Funktionen können nicht nur als Argumente an andere Funktionen übergeben, sondern auch als Ergebnisse zurückgeliefert werden. Dies soll in dieser Aufgabe genutzt werden, um ein einfaches Telefonbuch zu implementieren: Ein Telefonbuch ist als Funktion repräsentiert, die den Namen einer Person akzeptiert und die Telefonnummer der Person zurückliefert.

1. Definiere zunächst eine Signatur `phonebook-result` für das Ergebnis des Nachschlagens in einem Telefonbuch. Ein solches Ergebnis ist entweder eine Telefonnummer oder ein Wert, der «nicht gefunden» darstellt. Die Signatur des Telefonbuchs ist `(string -> phonebook-result)`.
2. Definiere einen Wert `empty-phonebook`, der das leere Telefonbuch repräsentiert.
3. Definiere eine Funktion namens `add-to-phonebook`, welche ein Telefonbuch, einen Namen und eine Telefonnummer erwartet und das um den neuen Eintrag erweiterte Telefonbuch zurückliefert.
4. Schreibe eine Funktion namens `lookup-in-phonebook`, welche ein Telefonbuch und einen Namen einer Person erwartet und die Nummer der Person im Telefonbuch nachschlägt und zurückliefert. Beispiele:
  - `(lookup-in-phonebook empty-phonebook "Hans")` liefert «nicht gefunden».
  - `(lookup-in-phonebook`

```
(add-to-phonebook empty-phonebook "Hans" "754829")
"Hans")
```

liefert die Nummer 754829.

- (lookup-in-phonebook  

```
(add-to-phonebook empty-phonebook "Hans" "754829")
"Lea")
```

liefert «nicht gefunden».

#### AUFGABE 9.26

Diese Aufgabe ist für Dich geeignet, falls Du das entsprechende Wissen aus der Mathematik besitzt:

Ein Polynom ist eine Funktion von  $\mathbb{R}$  nach  $\mathbb{R}$  und hat folgende Form:

$$p(x) = a_0 + a_1 \times x + a_2 \times x^2 + \dots + a_n \times x^n$$

Ein Polynom wird also eindeutig durch die Koeffizienten  $a_0$  bis  $a_n$  bestimmt.

1. Schreibe die Datendefinition für Polynome.
2. Programmiere eine Funktion `polynomial+`, die zwei Polynome addiert. Schreibe ggf. zutreffende Eigenschaften auf und überprüfe diese!
3. Programmiere eine Funktion `polynomial*`, die zwei Polynome multipliziert. Beispielsweise werden die Polynome  $a_0 + a_1x$  und  $b_0 + b_1x$  nach folgendem Schema multipliziert:

$$\begin{aligned} & a_0 \times (b_0 + b_1 \times x) \\ + & a_1x \times (b_0 + b_1 \times x) \\ = & a_0 \times b_0 + a_0 \times b_1x + a_1 \times b_0x + a_1 \times b_1x^2 \end{aligned}$$

Schreibe ggf. zutreffende Eigenschaften auf und überprüfe diese!

4. Schreibe eine Funktion `polynomial-function`, die ein Polynom akzeptiert und eine Funktion liefert, die ein Polynom an einer bestimmten Stelle auswertet, also gerade der Funktion  $p$  aus der Definition entspricht.
5. Die Ableitung eines Polynoms  $p$  wie oben ist bekanntlich durch

$$p'(x) = a_1 + 2 \times a_2 \times x + 3 \times a_3 \times x^2 + \dots + n \times a_n \times x^{n-1}$$

gegeben. Schreibe die Funktion `polynomial-derivative`, die von einem gegebenen Polynom das abgeleitete Polynom berechnet.

## AUFGABE 9.27

Listen lassen sich nach verschiedenen Kriterien sortieren: zum Beispiel aufsteigend, absteigend, oder abhängig von einem Feld der Elemente. So kann eine Liste von Fußballspielen nach der Gesamtanzahl der Tore, der Anzahl der Tore des Heimteams oder der Anzahl der Tore des Gästeteams oder der Anzahl der Tore der gewinnenden Mannschaft sortiert werden. Die genaue Anordnung wird durch eine Operation festgelegt, die bestimmt, ob ein Element vor einem anderen stehen soll.

1. Schreibe eine Funktion, die eine Liste nach einem beliebigen Kriterium sortiert. Die Funktion sollte folgende Signatur haben:

```
(: list-sort ((%a %a -> boolean) (list-of %a) -> (list-of %a)))
```

Das erste Argument ist eine Funktion, die eine *Ordnung* realisiert, also zwei Elemente vergleicht, und `#t` zurückliefert, falls sie schon in der richtigen Reihenfolge sind und `#f`, falls nicht. Zum Beispiel können `<=` oder `>=` für das Sortieren von Listen von Zahlen verwendet werden.

2. Benutze diese Funktion, um eine Liste von Fußballspielen nach den folgenden Kriterien zu sortieren:
  - Gesamtanzahl der Tore,
  - Anzahl der Tore der Heimmannschaft,
  - Anzahl der Tore der Gastmannschaft oder
  - Anzahl der Tore der gewinnenden Mannschaft

## AUFGABE 9.28

Folgen mit unendlicher Länge lassen sich als zusammengesetzte Daten mit zwei Komponenten repräsentieren: Dabei ist die erste Komponente das erste Element der Folge und die zweite eine Funktion ohne Parameter, die, wenn sie angewendet wird, eine Folge mit den restlichen Elementen ohne das erste liefert. Solche unendlichen Folgen heißen *Streams*. Schreibe Daten- und Record-Definition für Streams!

Hier ist eine Funktion, die einen Stream aus natürlichen Zahlen, angefangen bei einer Zahl  $n$ , liefert:

```
; Stream mit Zahlen ab n erzeugen
(: from (natural -> stream))
(define from
  (lambda (n)
    (make-stream n
```

```
(lambda () (from (+ n 1))))))
```

(Dabei ist angenommen, dass der Konstruktor der Record-Definition `make-stream` heißt.) Zur Betrachtung von Streams ist folgende Funktion nützlich, welche die ersten  $n$  Elemente eines Streams als Liste extrahiert:

```
; erste Elemente eines Streams in eine Liste extrahieren
(: stream-take (natural stream -> (list-of %a)))
(define stream-take
  (lambda (n stream)
    (if (= n 0)
        empty
        (cons (stream-first stream)
                (stream-take (- n 1)
                             ((stream-rest-proc stream)))))))
```

(Dabei ist angenommen, dass die Selektoren für Streams `stream-first` und `stream-rest-proc` heißen.) `Stream-take` lässt sich z.B. auf das Ergebnis von `from` anwenden:

```
(stream-take 17 (from 4))
↪ #<list 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20>
```

Programmiere einige intellektuelle Herausforderungen mit Streams!

1. Programmiere eine Funktion `stream-drop`, die eine natürliche Zahl  $n$  und einen Stream akzeptiert, und einen neuen Stream liefert, der aus dem alten durch Weglassen der ersten  $n$  Elemente entsteht:

```
(stream-take 17 (stream-drop 3 (from 4)))
↪ #<list 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23>
```

2. Programmiere eine Funktion `stream-filter` analog zu `filter`:

```
(stream-take 10 (stream-filter odd? (from 1)))
↪ #<list 1 3 5 7 9 11 13 15 17 19>
```



3. Programmiere eine Funktion `drop-multiples`, die eine Zahl  $n$  und einen Stream von Zahlen  $s$  akzeptiert. `Drop-multiples` soll einen Stream liefern, in dem gegenüber  $s$  alle Vielfachen von  $n$  entfernt wurden:

```
(stream-take 10 (drop-multiples 3 (from 1)))  
↪ #<list 1 2 4 5 7 8 10 11 13 14>
```

4. Schreibe eine Funktion `sieve`, die aus einem Stream von Zahlen all diejenigen Zahlen entfernt, die Vielfache von Vorgängern im Stream sind:

```
(stream-take 10 (sieve (from 2)))  
↪ #<list 2 3 5 7 11 13 17 19 23 29>
```

Um was für Zahlen handelt es sich in dem Beispielaufruf und warum?

5. Schreibe eine Funktion `powers`, die für eine Zahl  $n$  einen Stream ihrer Potenzen liefert:

```
(stream-take 10 (powers 2))  
↪ #<list 2 4 8 16 32 64 128 256 512 1024>
```

6. Schreibe eine Funktion `stream-map` analog zu `list-map`:

```
(stream-take 10 (stream-map (lambda (x) (+ x 1)) (from 1)))  
↪ #<list 2 3 4 5 6 7 8 9 10 11>
```

7. Schreibe eine Funktion `merge`, die zwei aufsteigende Streams von Zahlen zu einem aufsteigenden Stream der Elemente beider Streams vereinigt:

```
(stream-take 10 (merge (powers 2) (powers 3)))  
↪ #<list 2 3 4 8 9 16 27 32 64 81>
```

8. Schreibe eine Definition für einen Stream aufsteigend sortierter Potenzen von Primzahlen:

```
(stream-take 10 prime-powers)  
↪ #<list 2 3 4 5 7 8 9 11 13 16>
```

Definiere dazu zunächst einen Stream aus Streams von Potenzen

```
(define prime-powers-stream (stream-map powers (sieve (from 2))))
```

Definiere eine Funktion `merge-streams`, welche diesen Stream akzeptiert und die Elemente der Streams aus `prime-powers-stream` mit Hilfe von `merge` aufsteigend sortiert.

#### AUFGABE 9.29

Betrachte folgende mysteriöse Funktion:

```
(: // ((%a -> (%b -> %b)) (list-of %a) -> (%b -> %b)))
```

```
(define //
  (lambda (proc lis)
    (cond
      ((empty? lis)
       (lambda (y)
         y))
      ((cons? lis)
       (lambda (y)
         ((proc (first lis))
          ((// proc (rest lis))
           y)))))))
```

**Hinweise:** Beachte die Signaturen! In mehreren Teilaufgaben gibt es Gelegenheiten, `curry` beziehungsweise `uncurry` zu benutzen.

- Vergleiche die Funktion mit `list-fold` und beschreibe, wie `//` und `list-fold` zueinander in Beziehung stehen. Schreibe, falls möglich, eine Definition von `//`, die `list-fold` benutzt und umgekehrt.
- Schreibe mit Hilfe von `//` eine Funktion `list-sum`, welche die Elemente einer Liste addiert.
- Schreibe eine Funktion `insert`, die eine reelle Zahl  $n$  akzeptiert und eine Funktion zurückliefert, die eine aufsteigend sortierte Liste von reellen Zahlen konsumiert und eine Liste zurückliefert, in der  $n$  an die entsprechende Stelle der Liste einsortiert wurde.
- Schreibe mit Hilfe von `//` eine Funktion, die eine Liste von reellen Zahlen aufsteigend sortiert.

#### AUFGABE 9.30

Betrachte folgendes mysteriöse Programm:

```

(define y
  (lambda (f)
    ((lambda (x)
      (f (lambda (z) ((x x) z))))
     (lambda (x)
      (f (lambda (z) ((x x) z)))))))

(define m
  (y
   (lambda (f)
     (lambda (x)
      (cond
        ((= x 1)
         1)
        (> x 1)
        (* x (f (- x 1))))))))

```

Probiere `m` aus – was macht die Funktion? `Y` wird auch *Fixpunktkombinator* genannt. Das Thema kommt in Abschnitt 16.3.3 auf Seite 500 noch einmal ausführlich:

#### AUFGABE 9.31

Beweise, dass für Funktionen  $p_1$  mit einem Parameter, die einparametrische Funktionen zurückgeben, und Funktionen  $p_2$  mit zwei Parametern gilt:

$$\begin{aligned}
 (\text{curry } (\text{uncurry } p_1)) &= p_1 \\
 (\text{uncurry } (\text{curry } p_2)) &= p_2
 \end{aligned}$$

#### AUFGABE 9.32

Eine Funktion  $f$  ist *idempotent*, wenn gilt:

$$(\text{compose } f \ f) = f$$

Zeige, dass folgende Funktion idempotent ist:

```
(define abs
  (lambda (x)
    (if (negative? x)
        (- x)
        x)))
```

Welche anderen idempotente Funktionen kennst Du?

## 10 PROGRAMMIEREN MIT AKKUMULATOREN

Manche Berechnungen funktionieren am einfachsten, wenn sie ein Zwischenergebnis mitführen und aktualisieren. Die bisherigen Konstruktionsanleitungen für Funktionen, die Listen oder natürliche Zahlen verarbeiten, können das aber nicht. Wir brauchen dafür eine neue Programmier-technik, das Programmieren mit *Akkumulatoren*, und entsprechend angepasste Konstruktionsanleitungen. Beides behandeln wir in diesem Kapitel.

### 10.1 ZWISCHENERGEBNISSE MITFÜHREN

Wir fangen mit einer scheinbar einfachen Funktion an: Gefragt ist eine Funktion, die eine Liste invertiert, also die Reihenfolge ihrer Elemente umdreht:

```
; Liste umdrehen
(: invert ((list-of %a) -> (list-of %a)))

(check-expect (invert empty) empty)
(check-expect (invert (list 1 2 3 4)) (list 4 3 2 1))
```

Gerüst und Schablone sehen wie folgt aus:

```
(define invert
  (lambda (list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ... (invert (rest list)) ...
       ... (first list) ...))))
```

Um den Rumpf zu vervollständigen, können wir uns an dem zweiten Testfall orientieren – da ist `list` die Liste mit den Elementen 1 2 3 4. Entsprechend ist `(first list)` die Zahl 1, `(rest list)` die Liste mit den Elementen 2 3 4, das heißt der rekursive Aufruf liefert die Liste mit den Elementen 4 3 2. Um das gewünschte Ergebnis mit den Elementen 4 3 2 1 zu bekommen, müssen wir deshalb `(first list)` hinten an das Ergebnis des rekursiven Aufrufs anhängen.

Um ein Element *hinten* an eine Liste zu hängen, haben wir bisher noch keine fertige Funktion – die müssen wir erst noch schreiben. Wir könnten die Arbeit jetzt unterbrechen, um das zu tun. Wir

machen erstmal nur eine Notiz, dass wir das später machen – in Form einer Kurzbeschreibung und einer Signatur:

```
; Element an Liste anhängen
(: append-element ((list-of %a) %a -> (list-of %a)))
```

Wenn wir diese Funktion voraussetzen, können wir `invert` recht einfach fertig schreiben:

```
(define invert
  (lambda (list)
    (cond
      ((empty? list) empty)
      ((cons? list)
       (append-element (invert (rest list))
                        (first list))))))
```

Die Funktion `append-element` ist ganz ähnlich der Funktion `concatenate` aus Abschnitt 6.5. Zunächst Testfälle:

```
(check-expect (append-element (list 1 2 3) 4) (list 1 2 3 4))
(check-expect (append-element empty 4) (list 4))
```

Gerüst und Schablone:

```
(define append-element
  (lambda (list element)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ... (first list) ...
       ... (append-element (rest list) element) ...))))
```

Im `cons`-Fall können wir auch hier die Lösung anhand eines Beispiels finden: Im Testfall hat `list` die Elemente 1 2 3, der rekursive Aufruf liefert also 2 3 4. Wir müssen `(first list)`, also die 1, nur noch vorne dranhängen, mit `cons`. Im `empty`-Fall hängen wir `element` hinten an eine leere Liste – wir brauchen deshalb eine einelementige Liste mit `element` drin, das könnten wir mit der eingebauten `list` machen:

```
(define append-element
  (lambda (list element)
    (cond
```

```

(empty? list) (list element))
(cons? list)
(cons (first list)
      (append-element (rest list) element))))))

```

Hier wollten wir gern «Fertig!» wie sonst auch schreiben, aber die Funktion funktioniert nicht: DrRacket beschwert sich, dass nach der öffnenden Klammer von `(list element)` keine Funktion steht sondern `#<empty-list>`. Wups! Das liegt daran, dass wir zwar in `(list element)` die eingebaute Funktion `list` gemeint haben, wir aber auch einen Parameter mit dem Namen `list` benutzen, und der überdeckt nach den Regeln der lexikalischen Bindung aus Abschnitt 9.6 auf Seite 267 die eingebaute Funktion.

Wir können das Problem auf zwei Arten lösen – wir benennen den Parameter um oder wir konstruieren die einelementige Liste «von Hand». Wir haben uns für letzteres entschieden:

```

(define append-element
  (lambda (list element)
    (cond
      ((empty? list) (cons element empty))
      ((cons? list)
       (cons (first list)
             (append-element (rest list) element))))))

```

Doch zurück zu `invert`. Obwohl die zu erledigende Aufgabe einfach erscheint, dauert schon das Invertieren von Listen der Länge 1000 eine ganze Weile. Du kannst das zum Beispiel ausprobieren, indem Du die Funktion `copies` aus Abschnitt 7.2 auf Seite 202 verwendest und das hier in der REPL auswertest:

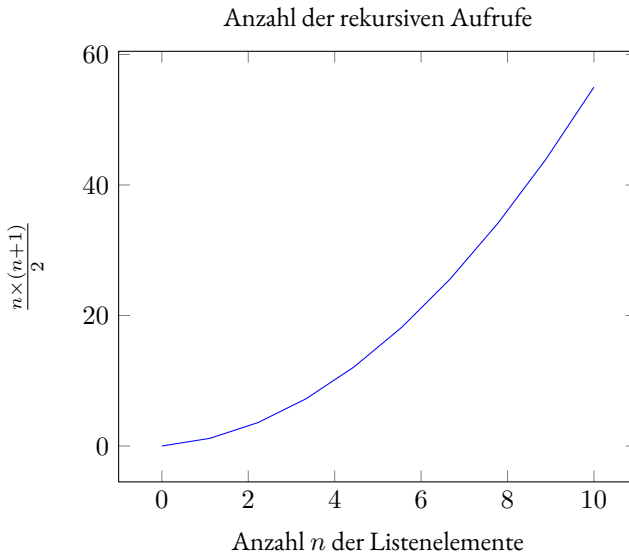
```

(invert (copies 1000 42))

```

Das dauert 2020 auf dem kaum ein Jahr alten Computer von Michael Sperber immerhin ein paar Sekunden: Das wäre vielleicht in den 70er Jahren noch akzeptabel gewesen. Tatsächlich ist es so, dass zum Beispiel das Invertieren einer Liste der Länge 400 *mehr* als doppelt so lang wie das Invertieren einer Liste der Länge 200 benötigt. Das liegt daran, dass `invert` bei jedem rekursiven Aufruf `append-element` aufruft, und `append-element` selbst macht so viele rekursive Aufrufe wie die Liste lang ist. Das sind für eine Liste der Länge 20 für den ersten Aufruf von `append-element` 19 Aufrufe, für den zweiten 18 Aufrufe und so weiter.

Das sind also  $19 + 18 + 17 + \dots + 1$  rekursive Aufrufe. Vielleicht erinnerst Du Dich – das ist



**Abbildung 10.1:** Funktionsaufrufe bei `invert`

ein Beispiel für die Gaußsche Summenformel, die wir in Abschnitt 8.1 auf Seite 213 bewiesen haben:

$$\forall n \in \mathbb{N} : \sum_{i=0}^n i = \frac{n \times (n + 1)}{2}$$

Wenn Du diese Formel für immer größer werdende  $n$  ausrechnest, wird die Zahl schwindelerregend schnell größer, wie Abbildung 10.1 zeigt. Woran liegt das?

Wenn Du die rechte Seite ausmultiplizierst, steht da:

$$\frac{n \times (n + 1)}{2} = \frac{n^2 + n}{2}$$

In der Formel bestimmt das  $n^2$  das schnelle Wachstum, auch *quadratisch* genannt.

Tatsächlich gibt es eine bessere Methode, eine Liste umzudrehen: Die obige `invert`-Funktion konstruiert die Ergebnisliste, indem stets Elemente *hinten* angehängt werden. Das entspricht nicht der «natürlichen» Konstruktion von Listen mit `cons`, das ein Element *vorn* anhängt. Das Ergebnis ließe sich aber durch Anhängen vorn ganz einfach konstruieren, und zwar, indem in folgender Reihenfolge Zwischenergebnisse berechnet werden, wie in folgendem Beispiel für den Testfall (`invert (list 1 2 3 4)`):



```
#<empty-list>
#<list 1>
#<list 2 1>
#<list 3 2 1>
#<list 4 3 2 1>
```

Jedes Zwischenergebnis entsteht aus dem vorhergehenden, indem ein Element vorn an die Liste darüber angehängt wird. Dies geschieht in der Reihenfolge, in der die Elemente in der ursprünglichen Liste auftreten: scheinbar einfach. Allerdings erlaubt die normale Konstruktionsanleitung für Listen nicht, dieses Zwischenergebnis mitzuführen: Das Ergebnis des rekursiven Aufrufs (`invert (rest lis)`) ist unabhängig vom Wert von (`first lis`). Damit aber ist es der Funktion aus der normalen Konstruktionsanleitung unmöglich, die obige Folge von Zwischenergebnissen nachzuvollziehen, da von einem Zwischenergebnis zum nächsten gerade (`first lis`) vorn angehängt wird. Wir müssen also etwas anders an das Problem herangehen.

Um das Zwischenergebnis mitzuführen, benutzen wir einen separaten Parameter, einen sogenannten *Akkumulator*. Dieser sammelt die invertierte Liste der bisher schon «gesehenen» Elemente auf. Hier ist die Signatur der neuen Funktion `invert-helper`:

```
; Hilfsfunktion zum Umdrehen einer Liste
(: invert-helper ((list-of %a) (list-of %a) -> (list-of %a)))
```

Der folgende Testfall soll illustrieren, wie die Funktion arbeitet:

```
(check-expect (invert-helper (list 4 5 6) (list 3 2 1))
               (list 6 5 4 3 2 1))
```

Die zweite Liste – der Akkumulator – enthält die «bereits invertierten» Elemente. Die erste Liste ist noch nicht verarbeitet; die Elemente werden nacheinander an den Akkumulator vorn drangehängt. Für die Definition der Funktion setzen wir erst einmal die bereits bekannte Schablone ein für Funktionen, die eine Liste akzeptieren:

```
(define invert-helper
  (lambda (list inverted)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ... (first list) ...
       ... (invert-helper (rest list) ...) ...))))
```

Ziemlich viele Lücken noch! Füllen wir erstmal die einfachste, im `empty`-Fall: Dann nämlich hat die Funktion schon «alle Elemente gesehen» und diese in den Akkumulator `inverted` «hineinakkumuliert» – der enthält dann die invertierte Eingabeliste:

```
(define invert-helper
  (lambda (list inverted)
    (cond
      ((empty? list) inverted)
      ((cons? list)
       ... (first list) ...
       ... (invert-helper (rest list) ...) ...))))
```

Im `cons`-Fall müssen wir insbesondere das Argument zum zweiten Parameter von `invert-helper` ergänzen: Wir sind natürlich versucht, da einfach `inverted` hinzuschreiben wie bei vielen anderen rekursiven Funktionen vorher. Aber wir müssen doch die Elemente von `list` da noch «hineinakkumulieren» und aus dem vorigen Zwischenergebnis das nächste machen. Für das «Hineinakkumulieren» nehmen wir das erste Element und hängen es vorn an, so wie wir es in der Beispielrechnung auch gemacht haben:

```
(define invert-helper
  (lambda (list inverted)
    (cond
      ((empty? list) inverted)
      ((cons? list)
       ...
       (invert-helper (rest list)
                      (cons (first list) inverted))
       ...))))
```

Was müssen wir noch dazuschreiben? Die Funktion arbeitet schon die Eingabeliste ab und akkumuliert ihre Elemente in `inverted` hinein. Sie ist bereits fertig, wir müssen nur die Ellipsen wegmachen:

```
(define invert-helper
  (lambda (list inverted)
    (cond
      ((empty? list) inverted)
      ((cons? list)
       (invert-helper (rest list)
```

```
(cons (first list) inverted))))))
```

Die Funktion kann schon was, ist aber natürlich nicht identisch zur ursprünglichen `invert`-Funktion, die ja nur eine Eingabe hat. Um «einfach nur eine Liste umzudrehen», können wir `invert-helper` mit einem leeren Akkumulator aufrufen, wie in diesem Testfall:

```
(check-expect (invert-helper (list 1 2 3) empty)
               (list 3 2 1))
```

#### AUFGABE 10.1

Definiere `invert` um, so dass es `invert-helper` benutzt!

□

Die neue Version von `invert` funktioniert nicht nur korrekt, sondern auch schnell: Sie benutzt keine Hilfsfunktion und macht so viele rekursive Aufrufe wie die Eingabeliste Elemente hat, ihre Laufzeit wächst also *linear*.

Übrigens: Da die Funktion `invert` generell nützlich ist, ist sie unter dem Namen `reverse` fest eingebaut.

## 10.2 SCHABLONEN FÜR FUNKTIONEN MIT AKKUMULATOR

Auch für Funktionen mit Akkumulator entwickeln wir eine Konstruktionsanleitung. Vorher wollen wir aber noch einmal anhand eines weiteren Beispiels Revue passieren lassen, wie der Konstruktionsprozess bei solchen Funktionen eigentlich genau aussieht.

Wir nehmen uns eine Funktion vor, die wir eigentlich schon kennen, nämlich aus Abschnitt 6.1 auf Seite 168:

```
; Summe der Elemente einer Liste von Zahlen berechnen
(: list-sum ((list-of number) -> number))
```

Wir nehmen uns allerdings diesmal vor, mit Akkumulator zu arbeiten. Dazu müssen wir uns erst einmal überlegen, was für Information der Akkumulator eigentlich akkumulieren soll. Das sollte ein *Zwischenergebnis* sein, das ist eine vorläufige Version des gewünschten Endergebnisses. Da es hier das Endergebnis die Summe aller Listenelemente ist, nehmen wir als Zwischenergebnis die Summe aller Listenelemente, die unsere Funktion schon «gesehen» hat. Wir nennen deshalb den Akkumulator `sum` und schreiben folgende Schablone:

```
(define list-sum-helper
  (lambda (list sum)
    (cond
      ((empty? list) ... sum ...)
      ((cons? list)
       (list-sum-helper (rest list)
                         (... (first list) ... sum ...))))))
```

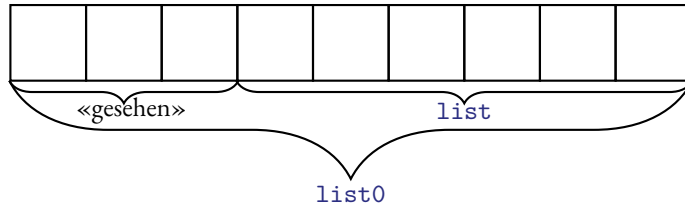
Warum sieht sie gerade so aus, beziehungsweise: Was ist der Unterschied zur ganz normalen Schablone für Listen als Eingabe aus Konstruktionsanleitung 18 auf Seite 175? Hier ist sie zur Erinnerung noch einmal:

```
(define f
  (lambda (... list ...)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ...
       (first list)
       (f (first list))
       ...
      ))))
```

Im `empty`-Zweig steht der Akkumulator `sum`, während in der Schablone aus Konstruktionsanleitung 18 gar nichts steht: Hier ist die Liste am Ende und es ist Zeit, das Endergebnis auszurechnen, das im `empty`-Zweig herauskommen soll. Weil der Akkumulator ein Zwischenergebnis ist, muss er am Schluss zumindest nah am Endergebnis dran sein. Meist *ist* das letzte Zwischenergebnis das Endergebnis.

Im `cons`-Fall steht ein rekursiver Aufruf mit `(rest list)` als Listen-Argument – wie in Konstruktionsanleitung 18 auch. Außerdem steht dort eine Hilfestellung für die Berechnung des Akkumulator-Arguments, also des nächsten Zwischenergebnisses. Da sollte das letzte Zwischenergebnis – hier `sum` – und das nächste Listenelement `(first list)` eingehen. Darum stehen sie in der Schablone. Dieser Teil der Schablone ist also nur eine Erweiterung der ursprünglichen Schablone.

Außerdem fällt Dir vielleicht auf, dass um den rekursiven Aufruf herum keine Ellipsen `...` stehen: Du solltest da nichts drumherum schreiben. Das liegt daran, dass der letzte rekursive Aufruf von `list-sum-helper` am Ende der Liste das Ergebnis produziert – das muss die Funktion einfach unverändert zurückliefern, und darum steht da nichts drumherum.



**Abbildung 10.2:** Gesehene Elemente einer Liste in einer Funktion mit Akkumulator

Um die Funktion zu vervollständigen, müssen wir noch klarer als bisher formulieren, was genau der Akkumulator `sum` repräsentiert. Oben haben wir etwas salopp geschrieben, dass es sich um die Summe aller Listenelemente handelt, welche die Funktion schon «gesehen» hat. Die sind aber für `list-sum-helper` gar nicht mehr sichtbar. Wir können sie sichtbar machen, indem wir die Funktion `list-sum` einbeziehen, die `list-sum-helper` aufruft. Hier ist die Schablone dafür:

```
(define list-sum
  (lambda (list0)
    (list-sum-helper list0 ...)))
```

Wir haben bewusst den Namen `list0` gewählt, damit wir ihn nicht mit dem `list` aus `list-sum-helper` durcheinanderbringen. (Anders als noch bei `reverse` – wir versuchen, es besser zu machen.)

Abbildung 10.2 zeigt die Beziehung zwischen `list0` und `list`: `list0` ist die Liste *aller* Elemente, `list` markiert die Stelle, an der sich `list-sum-helper` gerade befindet, besteht also aus noch nicht gesehenen Elementen.

Bevor wir nun die Schablone ausfüllen, sollten wir überlegen, in welchem Verhältnis `list0`, `list` und `sum` stehen. Hier ist das nämlich noch einfach, aber bei machen Funktionen mit Akkumulator werden wir sehen, dass es schwieriger ist. In diesem Fall könnte das so aussehen:

`sum` ist die Summer aller Elemente in `list0` vor `list`.

Diese Aussage sollten wir als Kommentar in die Funktion schreiben, denn daraus ergibt sich alles weitere. Weil sie so wichtig ist, hat sie einen eigenen Namen: *Invariante*. («Vario» heißt im Lateinischen «sich ändern», entsprechend ist eine In-variante etwas, was sich nicht verändert.)

Um die Funktion fertigzustellen, fangen wir damit an, die Lücke in `list-sum` zu schließen, die `list-sum-helper` mit `list0` aufruf. `list0` und `list` sind also gleich – es gibt keine Elemente «vor `list`». Die Summe dieser leeren Liste und damit das Argument im Aufruf von `list-sum-helper` ist entsprechend 0:

```
(define list-sum
```

```
(lambda (list0)
  (list-sum-helper list0 0)))
```

Als nächstes ist der `empty`-Zweig dran. Hier ist `sum` die Summe aller Elemente vor `list`, und, weil `list` leer ist, sind das *alle* Elemente von `list0`. Deswegen ist `sum` das gewünschte Endergebnis. Zwischenstand:

```
(define list-sum-helper
  ; sum ist die Summe aller Elemente in list0 vor list
  (lambda (list sum)
    (cond
      ((empty? list) sum)
      ((cons? list)
       (list-sum-helper (rest list)
                         (... (first list) ... sum ...))))))
```

Es bleibt der rekursive Aufruf. Hier muss der *neue* Wert von `sum` berechnet werden, also die Summe aller Elemente vor `(rest list)`. Dazu müssen wir auf die bisherige Summe `(first list)` addieren:

```
(define list-sum-helper
  ; sum ist die Summe aller Elemente in list0 vor list
  (lambda (list sum)
    (cond
      ((empty? list) sum)
      ((cons? list)
       (list-sum-helper (rest list) (+ (first list) sum))))))
```

Fertig!

Also na ja – es nervt etwas, immer zwei Funktionen schreiben zu müssen und immer `-helper` dranzuhängen. Wir können die Funktion etwas übersichtlicher machen, indem wir `list-sum-helper` zu einer lokalen Definition innerhalb von `list-sum` machen:

```
(define list-sum
  (lambda (list0)
    (define list-sum-helper
      ; sum ist die Summe aller Elemente in list0 vor list
      (lambda (list sum)
        (cond
```

```

      ((empty? list) sum)
      ((cons? list)
       (list-sum-helper (rest list) (+ (first list) sum))))))
(list-sum-helper list0 0)))

```

Außerdem kannst Du, wenn Dich das `helper` nervt, einen knackigeren Namen wählen, der Dir besser gefällt. Wir nehmen `accumulate`:

```

(define list-sum
  (lambda (list0)
    (define accumulate
      ; sum ist die Summer aller Elemente in list0 vor list
      (lambda (list sum)
        (cond
          ((empty? list) sum)
          ((cons? list)
           (accumulate (rest list) (+ (first list) sum))))))
      (accumulate list0 0)))

```

Aus diesem Beispiel ergibt sich folgende Konstruktionsanleitung:

#### KONSTRUKTIONSANLEITUNG 21 (LISTEN ALS EINGABE, MIT AKKUMULATOR: SCHABLONE)

Wenn Du eine Funktion schreibst, die eine Liste akzeptiert und einen Akkumulator benutzen soll, gehe folgendermaßen vor:

1. Überlege Dir, was für Information der Akkumulator repräsentieren soll. Das ist typischerweise ein Zwischenergebnis – also ein vorläufiger Wert für das Endergebnis. Insbesondere ist die Signatur des Akkumulators die gleiche wie die des Endergebnisses.
2. Konstruiere die Schablone wie folgt:

```

(define f
  (lambda (... list0 ...)
    (define accumulate
      (lambda (list acc)
        (cond
          ((empty? list) ... acc ...)
          ((cons? list)
           (accumulate (rest list)

```

```
(... (first list) ... acc ...))))
(accumulate list0 ...))
```

3. Formuliere eine möglichst konkrete Invariante zwischen *list<sub>0</sub>*, *list* und *acc* und schreibe sie als Kommentar zu `accumulate`.
4. Fülle mit Hilfe der Invariante die Ellipsen in der Funktion aus.

Die Konstruktionsanleitung zeigt auch, warum schwieriger ist, eine Funktion mit Akkumulator zu schreiben, als eine «normale» Funktion, die Listen akzeptiert: Du musst eine Invariante finden, und dafür gibt es nur wenig allgemeingültige Hilfestellung.

#### AUFGABE 10.2

Schreibe die Funktion `list-min-nonempty` aus Abschnitt 6.8 auf Seite 189 noch einmal, diesmal mit Akkumulator.

Überlege Dir, was der Akkumulator repräsentieren sollte sowie eine sinnvolle Invariante!

Genau wie bei `list-min-nonempty` musst Du dafür die Konstruktionsanleitung etwas variieren: Die nichtleere Liste kannst Du schon vorab in erstes Element und Rest aufteilen und daraus die richtigen Eingaben für `accumulate` berechnen. `Accumulate` ist aber wie in der Schablone. □

### 10.3 ÜBER NATÜRLICHE ZAHLEN AKKUMULIEREN

Auch über natürliche Zahlen können wir Funktionen schreiben, die einen Akkumulator sinnvoll benutzen. Wir zeigen das anhand der `power`-Funktion, die wir schonmal «normal» in Abschnitt 7.1 auf Seite 200 implementiert haben.

Kurzbeschreibung, Signatur und Tests sind wie gehabt:

```
; Potenz einer Zahl berechnen
(: power (number natural -> number))
```

```
(check-expect (power 5 0) 1)
(check-expect (power 5 3) 125)
```

Wir benutzen nun schon von vornherein die gleiche Namenskonvention wie in Konstruktionsanleitung 21 und benennen den entschiedenen Parameter mit einer 0 am Ende.

```
(define power
  (lambda (base exponent0)
    ...))
```



Die Schablone entsteht jetzt ganz ähnlich wie bei den Listen: Wir schreiben eine `accumulate`-Funktion mit Parametern `exponent` und `acc`. `Accumulate` macht genau wie in der Schablone aus Konstruktionsanleitung 19 auf Seite 202 eine Verzweigung nach den zwei Fällen natürlicher Zahlen. Der rekursive Aufruf bei `positive?` muss genau wie dort (`- exponent 1`) übergeben:

```
(define power
  (lambda (base exponent0)
    (define accumulate
      (lambda (exponent acc)
        (cond
          ((zero? exponent)
           ... acc ...)
          ((positive? exponent)
           (accumulate (- exponent 1) ... acc ...))))
      (accumulate exponent0 ...)))
```

Jetzt müssen wir uns noch überlegen, was `acc` eigentlich sein soll. Da die Potenz ein wiederholtes Produkt von `base` ist, bietet sich an, dort ebenfalls ein Produkt unterzubringen. Während `accumulate` jedesmal `exponent` um eins herunterzählt, könnten wir den Akkumulator bei jeder Runde ein weiteres Mal mit `base` multiplizieren – ähnlich, wie wir es vielleicht auch auf einem Zettel machen würden. Der Akkumulator ist also eine «Zwischenpotenz» und wir nennen ihn deshalb einfach `power`.

Wir sollten noch eine Invariante formulieren, damit auch alles richtig läuft. Dafür müssen wir klären, *welche* Zwischenpotenz `power` eigentlich ist. Da `exponent` immer kleiner wird, bietet sich die Differenz zwischen `exponent0` und `exponent` an – die fängt bei 0 an und wird immer größer:

```
(define power
  (lambda (base exponent0)
    (define accumulate
      ; power ist base^(exponent0 - exponent)
      (lambda (exponent power)
        (cond
          ((zero? exponent) ...)
          ((positive? exponent)
           (accumulate (- exponent 1) ...))))
      (accumulate exponent0 1)))
```

Das Hütchen `^` heißt hier «hoch».

Mit der Invariante können wir jetzt die Lücken füllen: Im ersten Zweig ist `exponent` o. Das bedeutet, dass `power` gerade «`base` hoch `exponent0`» ist – das gewünschte Endergebnis. Im zweiten Zweig müssen wir entsprechend dem reduzierten `exponent` einmal `base` an das Zwischenergebnis drannmultiplizieren. Es fehlt nur noch der Wert für `power` im ersten Aufruf von `accumulate`. Da  $b^0 = 1$  ist, müssen wir da 1 einsetzen:

```
(define power
  (lambda (base exponent0)
    (define accumulate
      ; power ist base^(exponent0 - exponent)
      (lambda (exponent power)
        (cond
          ((zero? exponent)
           power)
          ((positive? exponent)
           (accumulate (- exponent 1) (* power base))))))
    (accumulate exponent0 1)))
```

Fertig!

Auch für solche Funktionen mit Akkumulator, die natürliche Zahlen konsumieren, können wir eine Schablone formulieren:

#### KONSTRUKTIONSANLEITUNG 22 (NATÜRLICHE ZAHLEN ALS EINGABE, MIT AKKUMULATOR: SCHABLONE)

Wenn Du eine Funktion schreibst, die eine natürliche Zahl akzeptiert und einen Akkumulator benutzen soll, gehe folgendermaßen vor:

1. Überlege Dir, was für Information der Akkumulator repräsentieren soll. Das ist typischerweise ein Zwischenergebnis – also ein vorläufiger Wert für das Endergebnis.
2. Konstruiere die Schablone wie folgt:

```
(define f
  (lambda (... n0 ...)
    (define accumulate
      (lambda (n acc)
        (cond
          ((zero? n) ... acc ...)
          ((positive? n)
```

```

      (accumulate (- n 1)
                  (... acc ...))))))
(accumulate n0 ...)))

```

3. Formuliere eine möglichst konkrete Invariante zwischen  $n_0$ ,  $n$  und  $acc$  und schreibe sie als Kommentar zu `accumulate`.
4. Fülle mit Hilfe der Invariante die Ellipsen in der Funktion aus.

### AUFGABE 10.3

Programmiere eine Funktion, welche die *Fakultät* einer Zahl berechnet (auf englisch «factorial»). Für eine Zahl  $n$  ist deren Fakultät  $n!$  das folgende Produkt:

$$1 \times 2 \times \cdots \times (n - 1) \times n$$

□

Hier noch ein Tipp für Funktionen auf natürlichen Zahlen mit Akkumulator. Bei den Funktionen `exponent` und `list-sum` ist es egal, ob die Funktion von 0 hochzählt oder von einer Zahl  $n$  herunterzählt: Bei Addition und Multiplikation spielt die Reihenfolge keine Rolle. Das ist aber nicht immer so: Häufig macht es einen Unterschied, ob die Funktion bei 0 (oder manchmal auch 1) anfängt und hochzählt oder von einer Zahl  $n$  herunterzählt. (Schau Dir gegebenenfalls noch einmal Abschnitt 7.3 auf Seite 205 an.) Um das zu entscheiden, überlege Dir, ob Du mit Bleistift und Papier «unten» oder «oben» anfangen würdest.

## 10.4 AKTIENKURSE ANALYSIEREN

Es gibt durchaus Funktionen, bei denen mehrere Zwischenergebnisse nötig sind. Um das zu demonstrieren, schreiben wir eine Funktion, die den maximalen Gewinn aus einer Reihe von Aktienkursen berechnet.

Diese Reihe von aufeinanderfolgenden Kursen repräsentieren wir als Liste von Zahlen. Entsprechend sehen Kurzbeschreibung und Signatur so aus:

```

; Bestmöglichen Gewinn durch Kauf und Verkauf ermitteln
(: max-profit ((nonempty-list-of real) -> real))

```

Gewinn erzielt man hier, indem die Aktie zunächst gekauft und dann wieder verkauft wird – und nicht umgekehrt. Die Differenz zwischen Verkaufs- und Kaufpreis ist dann der Gewinn. Der folgende Testfall illustriert dies:

```
(check-expect (max-profit (list 5 2 7 3 9 5 1)) 7)
```

Hier hätte man den maximalen Gewinn erzielt durch Kauf zum Kurs 2 und Verkauf zum Kurs 9.

Das Gerüst sieht so aus:

```
(define max-profit
  (lambda (spots)
    ...)))
```

(«Spot» ist ein englisches Wort für «Kurs».)

Aber wie geht es weiter?

Die einfache Strategie, einfach die Differenz zwischen Maximum und Minimum der Liste als maximalen Gewinn auszuweisen, funktioniert leider nicht: Das Minimum ist 1, liegt aber leider hinter dem Maximum von 9.

Ebensowenig eignet sich dieses Problem für eine naive, «normale» rekursive Funktion. Die Schablone wäre diese hier:

```
(define max-profit
  (lambda (spots)
    (cond
      ((empty? spots) ...)
      ((cons? spots)
       ... (first spots) ...
       ... (max-profit (rest spots) ...))))))
```

Das Problem ist, dass der Profit des Rests der Liste nicht ausreicht, um daraus und dem ersten Element den Profit der Gesamtliste auszurechnen. Wir müssten dafür auch noch den dazugehörigen Verkaufskurs wissen.

Wir versuchen es also mal mit Zwischenergebnissen. Dafür sieht die Schablone so aus:

```
(define max-profit
  (lambda (spots0)
    (define accumulate
      (lambda (spots acc)
        (cond
          ((empty? spots) acc)
```

```

      ((cons? spots)
       (accumulate (rest spots)
                   ... (first spots) ... acc ...))))
    (accumulate spots0 ...))

```

Da die Funktion im ganzen den maximalen Gewinn ausrechnen soll, bietet es sich an, als Zwischenergebnis den maximalen Gewinn der bisher gesehenen Kurse zwischen `spots0` und `spots` als Zwischenergebnis mitzuführen:

```

(define max-profit
  (lambda (spots0)
    (define accumulate
      ; max-profit ist der maximale Gewinn zwischen
      ; spots0 und spots
      (lambda (spots max-profit)
        (cond
          ((empty? spots) max-profit)
          ((cons? spots)
           (accumulate (rest spots)
                       ... (first spots) ... acc ...))))
      (accumulate spots0 0)))

```

Das dumme ist, dass wir auch hier nicht genug Daten haben, um beim rekursiven Aufruf von `accumulate` einen neuen Wert für `max-profit` zu berechnen. Grundsätzlich gibt es aber nur zwei Möglichkeiten:

- `Max-profit` bleibt, wie es ist.
- `Max-profit` wird aktualisiert, weil es besser ist, zum Kurs `(first spots)` zu verkaufen, als zum bisher besten Verkaufskurs.

Wir müssen also herausbekommen, wie hoch der Gewinn wäre, wenn wir zum Kurs `(first spots)` verkaufen würden. Das könnten wir, wenn wir den bestmöglichen *Kaufkurs* wüssten. Etwas Überlegung ergibt, dass dieser Kaufkurs der minimale Kurs unter den vergangenen Kursen ist: Dieses Minimum führen wir als weiteres Zwischenergebnis mit. Aktualisiert wird dieses Minimum mit der eingebauten Funktion `min`:

```

(define max-profit
  (lambda (spots0)
    (define accumulate
      ; min-spot ist das Minimum der Elemente zwischen

```



```
(accumulate (rest spots0) (first spots0) 0)))
```

Fertig!

#### AUFGABE 10.4

Die *Fibonacci-Folge* ist eine Folge natürlicher Zahlen, die mit 0 und 1 anfängt. Jede weitere Zahl ist die Summe der beiden Zahlen davor.

Schreibe eine Funktion, welche für eine Zahl  $n$  die  $n$ -te Zahl aus der Fibonacci-Zahl berechnet. Schreibe dafür eine Funktion mit zwei Akkumulatoren.

Wichtig bei dieser Aufgabe: Überlege Dir vorher, ob Du die Fibonacci-Zahlen mit Papier und Bleistift ausrechnen würdest, indem Du bei 0 anfängst und hochzählst oder stattdessen von  $n$  herunterzählst. Schreibe Deine Funktion entsprechend.  $\square$

### 10.5 KONTEXT UND ENDREKURSION

In diesem Abschnitt werfen wir einen Blick darauf, wie die Auswertung rekursiver Funktionsaufrufe funktioniert. Dabei wird ein wichtiger Unterschied zwischen den Funktionen mit Akkumulator und den «normalen» Funktionen davor sichtbar.

Als Beispiel betrachten wir ein weiteres Mal `list-sum`, zunächst in der Version mit Akkumulator aus Abschnitt 10.2 auf Seite 289. Am besten ist, wenn Du Dir selbst im Stepper die Auswertung von

```
(list-sum (list 1 2 3 4))
```

anschaust. Hier sind die wichtigsten Schritte bei der Auswertung:

```
(list-sum #<list 1 2 3 4>)
↪ (accumulate #<list 1 2 3 4> 0)
↪ (accumulate (rest #<list 1 2 3 4>) (+ (first #<list 1 2 3 4>) 0))
↪ (accumulate #<list 2 3 4> 1)
↪ (accumulate #<list 3 4> 3)
↪ (accumulate #<list 4> 6)
↪ (accumulate #<empty-list> 10)
↪ 10
```

Wir haben den Wert des `sum`-Parameters immer untereinander geschrieben, und man sieht gut, wie sich das Zwischenergebnis von 0 schrittweise auf das Endergebnis 10 zubewegt.

Wenn Du das «alte» `list-sum` in Abschnitt 6.1 auf Seite 6.1 im Stepper laufen lässt, sieht das schon optisch ganz anders aus:

```
(list-sum #<list 1 2 3 4>)
↪ (+ (first #<list 1 2 3 4>) (list-sum (rest #<list 1 2 3 4>)))
↪ (+ 1 (list-sum #<list 2 3 4>))
↪ (+ 1 (+ 2 (list-sum #<list 3 4>)))
↪ (+ 1 (+ 2 (+ 3 (list-sum #<list 4>))))
↪ (+ 1 (+ 2 (+ 3 (+ 4 (list-sum #<empty-list>)))))
↪ (+ 1 (+ 2 (+ 3 (+ 4 0))))
↪ (+ 1 (+ 2 (+ 3 4)))
↪ (+ 1 (+ 2 7))
↪ (+ 1 9)
↪ 10
```

Hier sieht man, dass die rekursiven Aufrufe die einzelnen Additionen «aufstauen», und die eigentliche Arbeit der Addition erst nach dem letzten rekursiven Aufruf stattfindet. Dieses «Aufstauen» kommt daher, dass der rekursive Aufruf in der alten Version innerhalb des Aufrufs von `+` steht:

```
(+ (first list) (list-sum (rest list)))
```

Bei der Version mit Akkumulator ist es genau umgekehrt und der rekursive Aufruf steht um den Aufruf von `+` herum:

```
(accumulate (rest list) (+ (first list) sum))
```

Wenn der Computer in der alten Version einen rekursiven Aufruf auswertet, muss er sich merken, dass nach dem rekursiven Aufruf noch eine Addition passieren muss. Entsprechend wird die Kette von `+`-Aufrufen bei jedem rekursiven Aufruf länger. Dieser Aufruf von `+` heißt der *Kontext* des rekursiven Aufrufs – er ist um ihn herumgewickelt.

Bei der Version mit Akkumulator hat der rekursive Aufruf von `accumulate` keinen Kontext, dementsprechend staut sich bei der Auswertung und im Stepper da auch nichts auf. Ein rekursiver Aufruf ohne Kontext heißt *endrekursiv*, weil nach dem Aufruf nichts mehr passieren muss, der Aufruf also «am Ende» steht.

Der Begriff «Endrekursion» ist etwas unglücklich: Ob ein Funktionsaufruf einen Kontext hat oder nicht, hat eigentlich gar nichts damit zu tun, ob er rekursiv ist oder nicht. Im Englischen gibt es den besseren Begriff *tail call*, der sowohl auf rekursive als auch nicht-rekursive Aufrufe zutrifft.

Dass der Aufruf im alten `list-sum` nicht endrekursiv ist, legt schon die Schablone fest, in der das Ergebnis des rekursiven Aufrufs noch mit dem ersten Element der Liste kombiniert werden



muss. Bei der Schablone für Funktionen mit Akkumulator ist das nicht so, entsprechend haben die entstehenden rekursiven Aufrufe auch keinen Kontext.

Die Auswertungsprozesse, die von endrekursiven Aufrufen generiert werden, gehen in einer geraden Linie voran und heißen auch *iterative* Prozesse. In anderen Programmiersprachen spricht man auch von *Schleifen*; viele Programmiererinnen und Programmierer benutzen deshalb den Namen `loop` statt `accumulate`.

## 10.6 DAS PHÄNOMEN DER UMGEDREHTEN LISTE

Von `list-sum` kennen wir jetzt zwei Varianten: die ursprüngliche «normal rekursive» Version und die iterative. Bei allen Funktionen, die wir bisher auf Listen geschrieben haben, ist es auch möglich, eine endrekursive Fassung zu schreiben. Die endrekursiven Versionen sind nicht besser. Im Gegenteil: sie sind ja schwieriger zu schreiben. In diesem Kapitel ist das deshalb nur eine Fingerübung. Wir werden aber in Kapitel 17 auf Seite 509 zeigen, dass wir in vielen anderen Programmiersprachen die Listenfunktionen iterativ schreiben *müssen*. Darum ist es gut, wenn wir das schonmal gemacht haben.

Bei Funktionen, die Listen als Ergebnis produzieren, ist eine Kleinigkeit zu beachten, wenn wir sie iterativ schreiben. Wir zeigen, was diese Kleinigkeit ist, anhand einer Funktion, die aus einer Liste von ganzen Zahlen die geraden Zahlen extrahiert. Das geht auch, indem wir einfach `filter` aufrufen, aber wir programmieren das zur Übung noch einmal. Kurzbeschreibung, Signatur und Testfall sind wie folgt:

```
; Aus einer Liste gerade Zahlen extrahieren
(: evens ((list-of integer) -> (list-of integer)))

(check-expect (evens (list 1 2 3 4 5 6))
              (list 2 4 6))
```

Die Schablone gibt folgendes her:

```
(define evens
  (lambda (list0)
    (define accumulate
      (lambda (list acc)
        (cond
          ((empty? list) ... acc ...)
```

```

      ((cons? list)
       (accumulate (rest list)
                   ... (first list) ... acc ...))))
(accumulate list0 ...))

```

Wir müssen uns wieder einen geeigneten Akkumulator überlegen: Es bietet sich an, die schon gesehenen geraden Zahlen aufzusammeln. Wir nennen also den Akkumulator genau wie die Funktion `evens` und schreiben eine geeignete Invariante:

```

(define evens
  (lambda (list0)
    (define accumulate
      ; evens enthält die geraden Zahlen zwischen list0 und list
      (lambda (list evens)
        (cond
          ((empty? list) ... evens ...)
          ((cons? list)
           (accumulate (rest list)
                       ... (first list) ... evens ...))))))
    (accumulate list0 ...)))

```

Der erste Wert für den Akkumulator muss gemäß der Invariante `empty` sein. Ebenfalls gemäß der Invariante ist `evens` am Ende das Endergebnis. Außerdem müssen wir im rekursiven Aufruf von `accumulate` noch einen neuen Wert für `evens` berechnen: Wir hängen `(first list)` an das bisherige Zwischenergebnis an, falls es gerade ist:

```

(define evens
  (lambda (list0)
    (define accumulate
      ; evens enthält die geraden Zahlen zwischen list0 und list
      (lambda (list evens)
        (cond
          ((empty? list) evens)
          ((cons? list)
           (accumulate (rest list)
                       (if (even? (first list))
                           (cons (first list) evens)
                           evens))))))
    (accumulate list0 ...)))

```

```
(accumulate list0 empty)))
```

Sieht gut aus, oder? Es gibt allerdings ein kleines Problem: Der Testfall schlägt fehl. Die Funktion liefert `#<list 6 4 2>` statt `#<list 2 4 6>` – die Liste ist verkehrt herum!

Das ist eine natürliche Konsequenz, wenn eine Liste iterativ erzeugt wird: Bei einem rekursiven Aufruf, bei dem die Ergebnisliste wächst, wird ja mit `cons` ein neues Element *vorn* angehängt, das erste Element kommt also zuletzt.

Das bedeutet, dass die Invariante zwar richtig, aber unpräzise ist: Sie sagt, welche Elemente in `evens` drin sind, aber nicht, in welcher Reihenfolge. Wir sollten sie erweitern:

```
; evens enthält die geraden Zahlen zwischen list0 und list
; in umgekehrter Reihenfolge
```

(Vielleicht magst Du `evens` sogar in `inverted-evens` umbenennen.)

Wir müssen die Funktion natürlich noch korrigieren. Glücklicherweise geht das einfach: Wir rufen einfach einfach am Ende `invert` auf dem Ergebnis auf, und das Ergebnis passt. Wir müssen uns keine Sorgen machen, dass dadurch die Funktion wieder quadratisch wird, weil wir `invert` nur am Ende aufrufen, nicht bei jedem rekursiven Aufruf.

#### AUFGABE 10.5

Schreibe eine iterative Variante von `concatenate` aus Abschnitt 6.6.3 auf Seite 183. Vergleiche sie mit der Funktion `invert-helper` auf Seite 288: gibt es eine Gemeinsamkeit? □

## 10.7 LISTEN ITERATIV ZUSAMMENFALTEN

Du erinnerst Dich vielleicht: In Abschnitt 9.4 auf Seite 257 haben wir Mutter aller Listenfunktionen geschrieben, `list-fold`, indem wir aus der Schablone für Listenfunktionen eine Funktion gemacht haben. Alle Listenfunktionen, die nach dieser Schablone programmiert sind, können durch einen einfachen Aufruf von `list-fold` ersetzt werden.

Für die Funktionen mit Zwischenergebnis gibt es eine andere Schablone, und wir machen in diesem Abschnitt für diese Schablone das gleiche, was `list-fold` für die frühere Schablone gemacht hat. Diesmal leiten wir die Funktion von vornherein aus der Schablone her und geben ihr gleich einen Namen, nämlich `list-fold-left`.

Warum `list-fold-left`? Wenn Du Dir die Arbeitsweise von `list-fold` im Stepper anschaut, siehst Du, dass sie die Liste von rechts nach links aufrollt. Eine Funktion mit Zwischenergebnis arbeitet aber von links nach rechts, darum das `-left`.

Hier die Schablone aus Konstruktionsanleitung 21 auf Seite 293 in Form einer Funktion:

```
(define list-fold-left
  (lambda (list0 ...)
    (define accumulate
      (lambda (list acc)
        (cond
          ((empty? list) ... acc ...)
          ((cons? list)
           (accumulate (rest list)
                       ... (first list) ... acc))))))
    (accumulate list0 ...)))
```

Überall, wo Ellipsen stehen, müssen wir Namen hinschreiben: Das sind insgesamt drei Stellen. Wir arbeiten uns von hinten nach vorn, zunächst ist der erste Aufruf von `accumulate` an der Reihe. Die dortige Ellipse steht für den ersten Wert des Akkumulators `acc` – wir wählen dafür den Namen `start-acc` und nehmen ihn in die Parameter von `list-fold-left` auf.

Beim rekursiven Aufruf von `accumulate` müssen `(first list)` und `acc` zum nächsten Wert für den Akkumulator kombiniert werden. Das lassen wir von einer Funktion erledigen, analog zu `for-cons` bei `list-fold`. Schließlich bleibt noch die Ellipse am Ende im `empty?`-Zweig, bei der aus dem letzten Akkumulator das Endergebnis wird. Diesen Übergang lassen wir ebenfalls von einer Funktion erledigen, die wir `final-acc->result` nennen.

Das Ergebnis sieht so aus:

```
(define list-fold-left
  (lambda (list0 start-acc next-acc final-acc->result)
    (define accumulate
      (lambda (list acc)
        (cond
          ((empty? list) (final-acc->result acc))
          ((cons? list)
           (accumulate (rest list)
                       (next-acc (first list) acc))))))
    (accumulate list0 start-acc)))
```

Damit können wir die Aufgaben der bisherigen iterativen Funktionen erledigen. Das machen wir gleich als Testfälle. Hier sind `invert` und `list-sum`:

```
(check-expect (list-fold-left (list 1 2 3)
```

```

empty
cons
(lambda (inverted) inverted))

(list 3 2 1))

(check-expect (list-fold-left (list 1 2 3 4 5)
                              0 +
                              (lambda (sum) sum))
              15)

```

## AUFGABE 10.6

Bilde `evens` mit einem Aufruf von `list-fold-left` nach!



Es fehlt noch eine Signatur für `list-fold-left`. Wir fangen wieder damit an, dass wir aufschreiben, was wir wissen, nämlich dass `list-fold-left` vierstellig und das `list0`-Argument eine Liste ist:

```
(: list-fold-left ((list-of %a) ... -> ...))
```

Beim ersten Testfall sieht man, dass der Akkumulator nicht die gleiche Signatur haben muss wie die Listenelemente. Wir benutzen für den Akkumulator deshalb eine neue Signaturvariable `%acc`:

```
(: list-fold-left ((list-of %a) %acc ... -> ...))
```

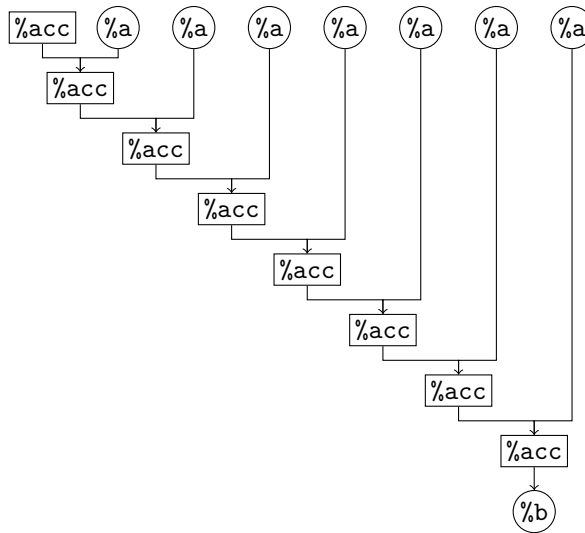
Wir wissen außerdem, dass `next-acc` eine zweistellige und `final-acc->result` eine einstellige Funktion ist:

```
(: list-fold-left
  ((list-of %a) %acc (... -> ...) (... -> ...) -> ...))
```

Die Argumente von `next-acc` sind ein Listenelement und der vorige Akkumulator, heraus kommt der nächste Wert für den Akkumulator. Wir können also folgendermaßen ergänzen:

```
(: list-fold-left
  ((list-of %a) %acc (%acc %a -> %acc) (... -> ...) -> ...))
```

Es bleibt `final-acc->result`, das als Argument den letzten Akkumulator akzeptiert. Das Ergebnis wird dann zum Endergebnis der Funktion, wofür wir die Signaturvariable `%b` wählen. Am Ende steht da folgende Signatur:



**Abbildung 10.3:** Ablauf von `list-fold-left`

```
(: list-fold-left
  ((list-of %a) %acc (%acc %a -> %acc) (%acc -> %b) -> %b))
```

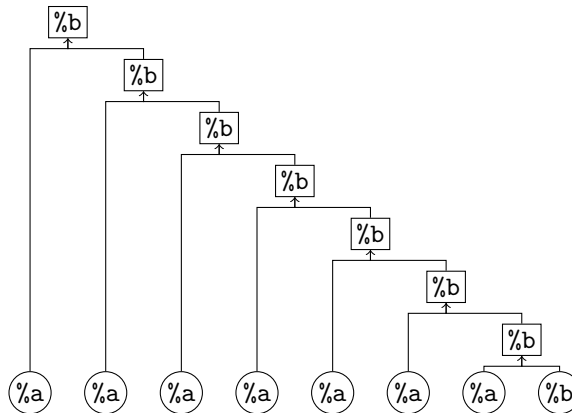
Abbildung 10.3 illustriert den Zusammenhang zwischen dem Ablauf von `list-fold-left` und der Signatur: Die Verarbeitung läuft sichtlich von links nach rechts.

Abbildung 10.4 zeigt im Gegensatz dazu die Arbeitsweise von `list-fold`, das von rechts nach links arbeitet. Zur Erinnerung hier noch einmal die Signatur:

```
(: list-fold (%a (%b %a -> %a) (list-of %b) -> %a))
```

## ANMERKUNGEN

Bei der Auswertung von Programmen durch den Computer wird für die Verwaltung von Kontexten Speicherplatz benötigt: Bei rekursiven Funktionen ohne Akkumulator wächst dieser Speicherplatz mit der Größe der Argumente. Entsprechend wird prinzipiell kein Speicherplatz benötigt, wenn kein Kontext anfällt. In den Lehrsprachen wird auch tatsächlich kein Speicherplatz für endrekursive Aufrufe verbraucht. Leider sind viele andere Programmiersprachen nicht so sparsam und verbrauchen sinnlos Speicherplatz auch für endrekursive Aufrufe. Mehr dazu in Kapitel 17.



**Abbildung 10.4:** Ablauf von `list-fold`

## AUFGABEN

### AUFGABE 10.7

Schreibe eine Funktion `list-sum+product`, die eine Liste von Zahlen akzeptiert und eine zweielementige Liste zurückgibt, deren erstes Element die Summe der Listenelemente und deren zweites Element ihr Produkt ist. Schreibe zwei Varianten der Funktion: eine «normal-rekursive» ohne Akkumulator und eine iterative mit zwei Akkumulatoren.

### AUFGABE 10.8

Von `power` hatten wir ja in Abschnitt 7.4 auf Seite 209 eine Variante `power2` programmiert, die durch eine alternative Sichtweise auf die natürlichen Zahlen viel schneller funktioniert. Schreibe auch davon eine Version mit Akkumulator. Hinweis: Du kannst die gleiche Invariante wie bei `power` verwenden.

### AUFGABE 10.9

Identifiziere die Kontexte der Aufrufe der Funktionen namens `p` in folgenden Ausdrücken:

```
(+ (p (- n 1)) 1)
(p (- n 1) acc)
(* (p (rest lis)) b)
(+ (* 2 (p (- n 1))) 1)
```

```
(p (- n 1) (* acc n))
(f (p n))
(+ (f (p n)) 5)
(p (f (- n 1)) (* n (h n)))
(+ (f (p n)) (h n))
```

Welche Aufrufe sind endrekursiv beziehungsweise *tail calls*?

#### AUFGABE 10.10

Schreibe eine iterative Variante von `list-length`.

#### AUFGABE 10.11

Schreibe eine iterative Variante von `concatenate`.

#### AUFGABE 10.12

Schreibe `list-fold-left` mit Hilfe von `list-fold`, indem Du die Eingabeliste umdrehst.

Schreibe umgekehrt `list-fold` mit Hilfe von `list-fold-left`!

□

#### AUFGABE 10.13

Das Newton-Verfahren dient zur nährungsweisen Berechnung von Nullstellen. Für ein gegebenen Startwert nähert sich die Iteration

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

immer näher an eine Nullstelle an.

Programmiere das Newton-Verfahren!

Hinweis: Die Lösung ist fertig, wenn der Funktionswert nahe bei 0 liegt, also kleiner als eine Toleranz ist. Da das Newton-Verfahren nicht immer eine Lösung liefert, programmiere Deine Funktion so, dass sie nach einer gewissen Anzahl von Schritten automatisch abbricht.



## 11 VIDEOSPIELE PROGRAMMIEREN

In diesem Kapitel programmieren wir zum ersten Mal ein vollständiges, benutzbares Programm, nämlich ein kleines Videospiel. Dafür brauchen wir zwei Zusätze für die Lehrsprache in DrRacket, die Teachpacks `image.rkt` und `universe.rkt`. Ersteres ist für das Anzeigen der Grafik zuständig (wir haben es im ersten Kapitel schon einmal benutzt), und `universe.rkt` ist dafür da, die Grafiken zu bewegen und interaktiv zu machen.

Beide Teachpacks können noch viel mehr als in dieses Kapitel passt. Zum Glück steht im Hilfezentrum von DrRacket umfangreiche Dokumentation, allerdings leider derzeit nur auf Englisch: Für das `image.rkt`-Teachpack muss man dort nach `2htdp/image` suchen, für `universe.rkt` nach `2htdp/universe`.<sup>1</sup> Hier sind direkte Links auf die Dokumentation:

<https://docs.racket-lang.org/teachpack/2htdpimage.html>

<https://docs.racket-lang.org/teachpack/2htdpuniverse.html>

### 11.1 BILDER MIT IMAGE.RKT

Für die Grafikprogrammierung mit DrRacket laden wir ein sogenanntes *Teachpack*, ein kleiner Sprachzusatz, in diesem Fall mit einer Reihe von Funktionen zur Erzeugung von Bildern. Wir sind ihm bereits in Abschnitt 1.4.2 auf Seite 15 begegnet: Um das Teachpack zu laden, musst Du im Menü *Sprache* (oder *Language* in der englischen Ausgabe) den Punkt *Teachpack hinzufügen* (*Add teachpack*) anwählen, und im dann erscheinenden Auswahl-Dialog die Datei `image.rkt` auswählen.

#### 11.1.1 EINFACHE BILDER

Im Teachpack `image.rkt` erzeugen verschiedene Funktionen einfache Bilder. So hat zum Beispiel die Funktion `rectangle` folgende Signatur:

```
(: rectangle (real real mode color -> image))
```

---

<sup>1</sup> `2htdp` bezieht sich auf den Titel des Buchs *How to Design Programs* in der zweiten Auflage [FELLEISEN et al., 2014], für das diese Teachpacks ursprünglich entwickelt wurden.

Dabei sind die ersten beiden Argumente Breite und Höhe eines Rechtecks in Pixeln. Das Argument mit Signatur `mode` ist eine Zeichenkette, die entweder `"solid"` oder `"outline"` sein muss. Sie bestimmt, ob das Rechteck als durchgängiger Klotz oder nur als Umriss gezeichnet wird. Das Argument mit der Signatur `color` ist eine Zeichenkette, die eine Farbe (auf Englisch) bezeichnet, zum Beispiel `"red"`, `"blue"`, `"yellow"`, `"black"`, `"white"` oder `"gray"`. Als Ergebnis liefert `rectangle` ein Bild, das von der DrRacket-REPL entsprechend angezeigt wird wie andere Werte auch. Beispiel:

```
(rectangle 100 30 "outline" "brown")
```



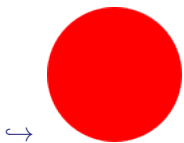
Als Farbe geht übrigens auch `"transparent"`, die das Bild durchsichtig beziehungsweise unsichtbar macht. (Wozu ist ein unsichtbares Bild gut, fragst Du vielleicht. Zum Beispiel, wenn wir ein kleineres Bild nicht ganz mittig platzieren wollen – dann können es auf ein transparentes Bild legen.)

Es gibt es noch weitere Funktionen, die geometrische Figuren zeichnen:

```
(: circle (real mode color -> image))
```

Die `circle`-Funktion liefert einen Kreis, wobei das erste Argument den Radius angibt. Die `mode`- und `color`-Argumente sind wie bei `rectangle`. Beispiel:

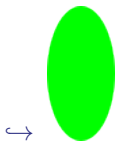
```
(circle 50 "solid" "red")
```



```
(: ellipse (real real mode color -> image))
```

Diese Funktion liefert eine Ellipse, wobei das erste Argument die Breite und das zweite die Höhe angibt. Beispiel:

```
(ellipse 50 100 "solid" "green")
```



```
(: triangle (real mode color -> image))
```

Diese Funktion liefert ein nach oben zeigendes gleichseitiges Dreieck, wobei das erste Argument die Seitenlänge angibt. Beispiel:

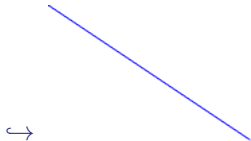
```
(triangle 50 "solid" "gold")
```



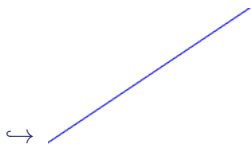
```
(: line (real real color -> image))
```

zeichnet eine Linie. Der Aufruf `(line w h c)` liefert ein Bild mit Breite  $w$  und Höhe  $h$ , in dem die Linie von der linken oberen in die rechte untere Ecke verläuft. Falls die Linie entlang der anderen Diagonale verlaufen soll, kannst Du einfach das Vorzeichen von  $w$  oder von  $h$  umdrehen. Beispiele:

```
(line 150 100 "blue")
```



```
(line -150 100 "blue")
```



Wir können auch ein Bild erzeugen, in dem Text steht, und zwar mit der Funktion `text`, die folgende Signatur hat:

```
(: text (string real color -> image))
```

Die Zahl ist die Höhe der Buchstaben. Beispiel:

```
(text "Schreibe Dein Programm!" 20 "red")
```

↪ 

Manchmal reichen einfache Rechtecke und Kreise nicht aus und wir wollen komplexere Formen erzeugen. Eine Möglichkeit ist die Funktion `polygon`, die ein  $n$ -Eck zeichnet. Sie hat folgende Signatur:

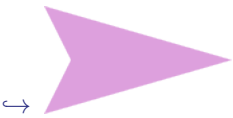
```
(: polygon ((list-of (mixed posn pulled-point)) mode color -> image))
```

Diese Funktion akzeptiert eine Liste von Eckpunkten und die üblichen `mode`- und `color`-Argumente. Ein Eckpunkt muss entweder zur Signatur `posn` oder zu `pulled-point` gehören. Das sind eingebaute Record-Typen, bei denen `posn` so definiert sein könnte:

```
(define-record posn
  make-posn
  posn?
  (posn-x real)
  (posn-y real))
```

Dieser Typ definiert also ganz normale kartesische Koordinaten mit X- und Y-Komponente. Hier ist ein Beispiel für ein Polygon mit solchen Eckpunkten:

```
(polygon (list (make-posn 0 0)
               (make-posn -20 40)
               (make-posn 120 0)
               (make-posn -20 -40))
  "solid"
  "plum")
```

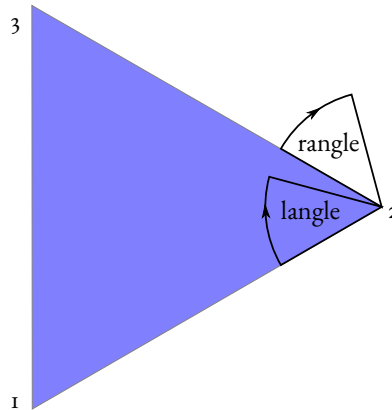


Bei `posn`-Eckpunkten sind die Kanten also alles gerade Linien. Mit `pulled-point`-Eckpunkten ist es möglich, die Kanten zu Kurven zu machen. Auch `pulled-point` ist ein Record-Typ, der so definiert sein könnte:

```
(define-record pulled-point
  make-pulled-point
  pulled-point?
  (pulled-point-lpull real)
  (pulled-point-langle real)
  (pulled-point-x real)
  (pulled-point-y real)
  (pulled-point-rpull real)
  (pulled-point-rangle real))
```

Wie `posn` hat auch `pulled-point` eine X- und eine Y-Koordinate. Außerdem gibt es zwei «Pull»- und zwei «Angle»-Komponenten, die spezifizieren, wie die Kanten gebogen werden. Abbildung 11.1 zeigt, wie das funktioniert: Für jeden Eckpunkt gibt es eine Kante vom vorigen Punkt – die *eingehende* Kante – und die Kante zum nächsten Punkt – die *ausgehende* Kante.

Die `pulled-point-langle`-Komponente gibt den Winkel an, mit dem die eingehende Kante gebogen wird, die `pulled-point-rangle`-Komponente den Winkel ist für die ausgehende

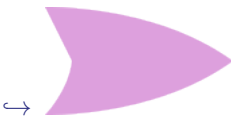


**Abbildung 11.1:** Funktionsweise von `pulled-point`

Kante zuständig. Abbildung 11.1 zeigt, in welche Richtung `langle` und `rangle` die beiden Kanten von Punkt Nummer 2 biegen: Die eingehende Kante kommt von Punkt 1 her, die ausgehende Kante geht zu Punkt 3. Die Komponenten `pulled-point-lpull` und `pulled-point-rpull` geben wann, wieviel die Kanten verbogen werden – das sind typischerweise Zahlen zwischen 0 und 1.

```
(polygon (list (make-pulled-point 1/2 0 0 0 1/2 -20)
               (make-posn -20 40)
               (make-pulled-point 1/2 -20 120 0 1/2 20)
               (make-posn -20 -40)))
```

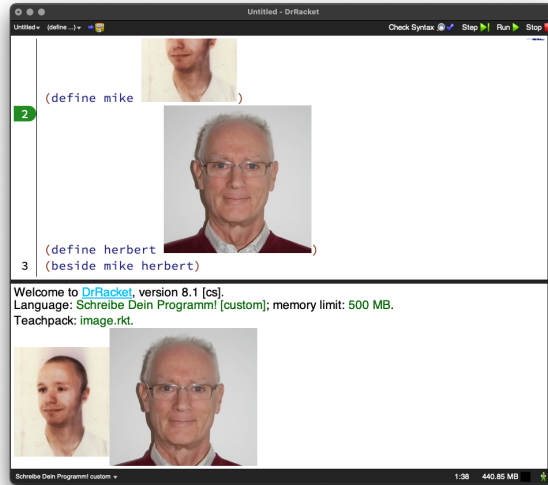
```
"solid"
"plum")
```



Es ist auch möglich, externe Bilder-Dateien in `image.rkt`-Bilder zu verwandeln. Dazu dient der Menüpunkt `Bild einfügen` im Spezial-Menü: DrRacket fragt nach dem Namen einer Bild-datei, die dann in den Programmtext da eingefügt wird, wo der Cursor steht. Alternativ ist es möglich, Bilder in anderen Applikationen auszuwählen und zu kopieren und dann in DrRacket einzufügen. Die eingefügten Bilder dienen dann als Literale für Bild-Objekte. Abbildung 11.2 zeigt ein Beispiel.

Schließlich ermitteln die folgenden Funktionen Breite und Höhe eines Bildes:

```
(: image-width (image -> natural))
```



**Abbildung 11.2:** Bilddatei in Programm einbetten

```
(: image-height (image -> natural))
```

### 11.1.2 BILDER ZUSAMMENSETZEN UND VERÄNDERN

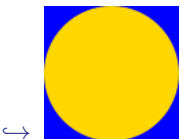
Da diese geometrischen Formen für sich genommen langweilig sind, können wir mehrere Bilder zu einem zusammensetzen.

Zum Aufeinanderlegen gibt es die Funktion `overlay`, die ein Bild mittig auf ein anderes Bild drauflegt. Das Bild, das bei `overlay` herauskommt, ist groß genug, dass beide Eingabebilder genau hineinpassen. Signatur:

```
(: overlay (image image -> image))
```

Beispiel:

```
(overlay
  (circle 50 "solid" "gold")
  (rectangle 100 100 "solid" "blue"))
```

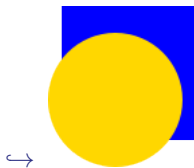


Mittig ist nicht immer das richtige Arrangement für zwei Bilder, die übereinandergelegt werden. Die Funktion `overlay/xy` erlaubt, das untere Bild gegenüber dem oberen in X- und Y-Richtung zu verschieben. Hier ist die Signatur:

```
(: overlay/xy (image real real image -> image))
```

Beispiel:

```
(overlay/xy
  (circle 50 "solid" "gold")
  10 -20
  (rectangle 100 100 "solid" "blue"))
```



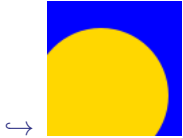
Das Beispiel zeigt, dass auch `overlay/xy` das Ergebnisbild gerade so groß macht, dass die beiden Eingabebilder hineinpassen.

Manchmal macht auch `overlay/xy` nicht das richtige, nämlich wenn einfach nur ein Objekt in eine «Szene» platziert werden soll, zum Beispiel das Gürteltier auf einem Straßenabschnitt. In dem Fall wollen wir genau wie bei `overlay/xy` kontrollieren, an welchen Koordinaten das daraufgelegte Bild erscheint, aber das Bild soll nicht größer werden, wenn zum Beispiel ein Teil des Gürteltiers über den Straßenabschnitt hinausragt. Diese Aufgabe erledigt die Funktion `place-image`, deren Signatur der von `overlay/xy` entspricht:

```
(: place-image (image real real image -> image))
```

Die beiden Koordinaten werden allerdings anders als bei `overlay/xy` nicht relativ zur mittigen Anordnung interpretiert. Stattdessen geben die beiden Koordinaten die Position im zweiten Bild an, wo die Mitte des ersten Bilds platziert wird, ausgehend von der oberen linken Ecke des ersten Bildes. Beispiel:

```
(place-image
  (circle 50 "solid" "gold")
  40 70
  (rectangle 100 100 "solid" "blue"))
```



Die Funktion `place-image` hat noch eine große Schwester `place-image/align`. Während `place-image` den Bezugspunkt immer in der Mitte des ersten Bilds setzt, erlaubt `place-image` auch andere Bezugspunkte. Hier die Signatur:

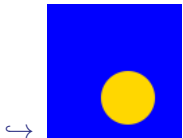
```
(: place-image/align (image real real x-place y-place image -> image))
```

Die Signaturen `x-place` und `y-place` könnten folgendermaßen definiert sein:

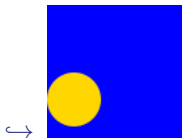
```
(define x-place (signature (enum "left" "right" "center")))
(define y-place (signature (enum "top" "bottom" "center")))
```

Diese Argumente geben also an, ob der Bezugspunkt horizontal am linken oder rechten Rand oder in der Mitte liegt und vertikal oben, unten oder in der Mitte:

```
(place-image/align
  (circle 20 "solid" "gold")
  40 70 "left" "center"
  (rectangle 100 100 "solid" "blue"))
```

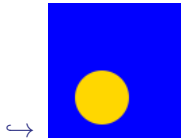


```
(place-image/align
  (circle 20 "solid" "gold")
  40 70 "right" "center"
  (rectangle 100 100 "solid" "blue"))
```



```
(place-image/align
  (circle 20 "solid" "gold")
  40 70 "center" "center"
  (rectangle 100 100 "solid" "blue"))
```

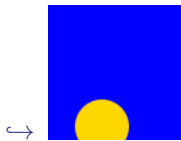




```

↔
(place-image/align
  (circle 20 "solid" "gold")
  40 70 "center" "top"
  (rectangle 100 100 "solid" "blue"))

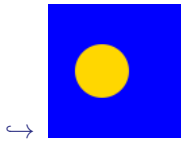
```



```

↔
(place-image/align
  (circle 20 "solid" "gold")
  40 70 "center" "bottom"
  (rectangle 100 100 "solid" "blue"))

```



Die folgenden Funktionen `above` und `beside` kennst Du vielleicht noch aus dem ersten Kapitel auf Seite 16. Sie setzen zwei Bilder übereinander respektive nebeneinander:

```

(: above (image image -> image))
(: beside (image image -> image))

```

Manchmal sieht ein Bild gut aus, ist aber zu klein oder zu groß geraten. Die Funktion `scale` kann es größer oder kleiner machen. Signatur:

```

(: scale (real image -> image))

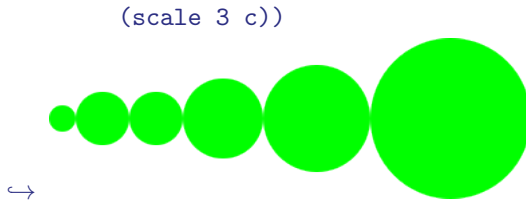
```

Beispiel:

```

(define c (circle 20 "solid" "green"))
(beside (scale 0.5 c)
  c
  (scale 1 c)
  (scale 1.5 c)
  (scale 2 c))

```



## 11.2 MODELLE UND ANSICHTEN

In unserem Videospiel geht es um Gürteltiere auf dem texanischen Highway. Vielleicht erinnerst Du Dich, wir sind ihnen schon in Abschnitt 3.4 auf Seite 94 begegnet. Zur Erinnerung, hier ist der relevante Code dazu:

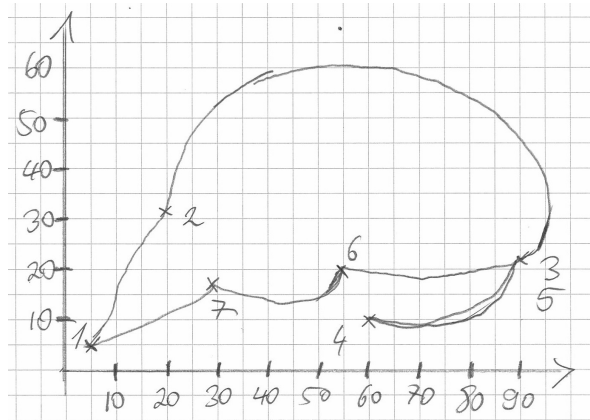
```
; Ein Gürteltier hat folgende Eigenschaften:
; - Gewicht (in g)
; - lebendig oder tot
(define-record dillo
  make-dillo
  dillo?
  (dillo-weight natural)
  (dillo-alive? boolean))

(: make-dillo (natural boolean -> dillo))
(: dillo-weight (dillo -> natural))
(: dillo-alive? (dillo -> boolean))

(define dillo1 (make-dillo 55000 #t)) ; 55 kg, lebendig
(define dillo2 (make-dillo 58000 #f)) ; 58 kg, tot
(define dillo3 (make-dillo 60000 #t)) ; 60 kg, lebendig
(define dillo4 (make-dillo 63000 #f)) ; 63 kg, tot

; Gürteltier überfahren
(: run-over-dillo (dillo -> dillo))

(define run-over-dillo
```



**Abbildung 11.3:** Umriss eines Gürteltiers

```
(lambda (dillo)
  (make-dillo (dillo-weight dillo)
              #f)))
```

Die Datendefinition mit dem zugehörigen Code hält zwar einige Eigenschaften von Gürteltieren fest, für ein Videospiel fehlt aber noch eine grafische Darstellung. Ebenso definiert `run-over-dillo` eine Operation auf Gürteltier-Objekten fest, aber auch die sagt nichts zu einer grafischen Darstellung. Eine solche Repräsentation mit zugehörigen Operationen heißt in der Softwarearchitektur auch *Modell* oder auch *Domänenlogik*, und bei größeren Softwaresystemen ist es eine gute Idee, das Modell von der grafischen Darstellung zu trennen, damit das Modell stabil bleiben kann, auch wenn die grafische Darstellung geändert wird.

Die grafische Darstellung eines Modells nennen wir dessen *Ansicht* (auf englisch *View*). Wenn wir uns also mit Gürteltieren als Modell für ein Videospiel beschäftigen, müssen wir daraus eine Ansicht in Form eines Bildes berechnen.

Wir warnen Dich schonmal vor: Der grafische Gestaltung in diesem Kapitel fehlt es vorsichtig gesagt an Finesse. Wir hoffen, Du kannst es besser machen!

Um also aus einem `dillo`-Wert ein Bild zu machen, werden wir solch eine Funktion benötigen:

```
; Gürteltier-Bild erzeugen
(: dillo-image (dillo -> image))
```

Als Basis dafür machen wir erstmal ein Bild mit dem Körper des Gürteltiers. Zu diesem Zweck haben wir zunächst von Hand eine Zeichnung mit dem Umriss des Gürteltiers in ein Koordinatensystem gezeichnet und die Punkte markiert, an denen sich die Richtung des Strichs abrupt ändert –

also quasi die Ecken. Diese Ecken haben wir durchnummeriert und daraus haben wir dieses Polygon gemacht:

```
(define dillo-body
  (overlay/xy (polygon
    (list (make-pulled-point
      0.3 30
      5 (- 60 5)
      0.2 5)
      (make-pulled-point
      0.4 20
      20 (- 60 32)
      0.5 -70)
      (make-pulled-point
      0.5 120
      90 (- 60 22)
      0.2 -30)
      (make-pulled-point
      0.5 90
      60 (- 60 10)
      0.5 90)
      (make-pulled-point
      0 0
      90 (- 60 22)
      0.2 -30)
      (make-pulled-point
      0 0
      55 (- 60 20)
      0.3 -30)
      (make-pulled-point
      0 0
      29 (- 60 17)
      0.5 20))
    "solid"
    "brown")
  0 -30)
```

```
(rectangle 100 60
         "solid" "transparent"))
```

Du siehst, da haben wir ganz schön viel herumprobiert, bis es einigermaßen aussah, insbesondere bei den Zahlen für die `pulled-points`. Immerhin haben wir die Koordinaten aus der Zeichnung direkt übertragen, dabei ist uns aber aufgefallen, dass die Koordinaten in der Mathematik von unten nach oben, im Teachpack aber von oben nach unten laufen. Darum steht da zum Beispiel `(- 60 5)` für die Y-Koordinate 5, weil der obere Rand der Zeichnung bei 60 liegt.

Außerdem ist das Gürteltier vor allem nach oben und rechts nach außen gewölbt. DrRacket schneidet das Polygon bei der Anzeige ab, weswegen wir es mit `overlay/xy` noch auf einen transparenten Hintergrund montiert haben. So sieht das Ergebnis aus:

dillo-body



Jetzt wollen wir ja nicht jedes Gürteltier genau gleich darstellen, sondern wir wollen dessen Eigenschaften bei der Darstellung berücksichtigen. Dafür schreiben wir nun die Funktion `dillo-image`, von der wir ja schon Kurzbeschreibung und Signatur gesehen haben. Hier ist das Gerüst mit der Schablone für zusammengesetzte Daten als Eingabe:

```
(define dillo-image
  (lambda (dillo)
    ...
    (dillo-weight dillo)
    (dillo-alive? dillo)
    ...))
```

Wir sollten also noch berücksichtigen, wieviel das Gürteltier wiegt und außerdem, ob es noch lebt. Das Gewicht berücksichtigen wir, indem wir das Gürteltier größer oder kleiner machen:

```
(define dillo-image
  (lambda (dillo)
    (scale
      (+ 1
        (/ (- (dillo-weight dillo) 50000)
            15000))
      dillo-body)
    ...))
```

```
(dillo-alive? dillo)
...))
```

Den Faktor bei `scale` haben wir durch Herumprobieren ermittelt. Aber `(dillo-alive? dillo)` steht noch da – wir bauen das ein, indem wir einem toten Gürteltier «tote Augen» ins Gesicht setzen:

```
(define dillo-image
  (lambda (dillo)
    (scale
      (+ 1
        (/ (- (dillo-weight dillo) 50000)
           15000))
      (if (dillo-alive? dillo)
          dillo-body
          (overlay/xy dead-eyes
                      -25 -25
                      dillo-body))))))

(define dead-eyes
  (overlay (line 10 10 "green")
           (line -10 10 "green")))
```

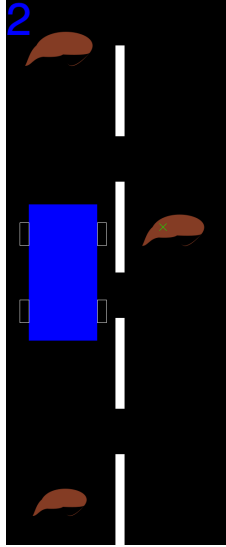
Fertig ist die Gürteltier-Darstellung!

## 11.3 DEN HIGHWAY MODELLIEREN

Ein Gürteltier allein macht noch kein Videospiel: Wir setzen gleich mehrere Gürteltiere auf die Straße und – natürlich – überfahren sie dann. Das sieht dann so aus wie in Abbildung 11.4. Das Auto bewegt sich auf der Straße und kann entweder auf die linke oder rechte Straßenseite wechseln, wo es dann gegebenenfalls über Gürteltiere fährt. Oben links steht ein Punktestand – wieviele Gürteltiere schon vom lebenden Zustand in den toten überführt wurden.

Damit wir das alles darstellen können, müssen für alles Datendefinition schreiben, was auf dem Bild der Straße zu sehen sind:

- die Position des Autos
- die Positionen der Gürteltiere zusammen mit ihren Zuständen



**Abbildung 11.4:** Gürteltiere auf der Straße

- der Punktestand

Für diese Aspekte des Spiels entwickeln wir Datendefinitionen.

Man sieht schon, dass «Positionen» eine wichtige Rolle spielen, sowohl für das Auto als auch für die Gürteltiere. In dem Bild kann man sehen, dass dazu gehört, auf welcher Seite der Straße sich etwas aufhält und bei welchem «Straßenmeter». Dabei kommt folgende beiden Datendefinitionen für die Straßenseite und eine Position heraus:

```
; Straßenseite
(define side
  (signature (enum "left" "right")))

; Eine Position auf der Straße besteht aus:
; - Straßenmeter (Abstand vom Straßenanfang in Meter)
; - Seite
(define-record position
  make-position
  position?
  (position-m-from-start real)
  (position-side side))
```

Als nächstes müssen wir die Idee «Positionen der Gürteltiere zusammen mit ihren Zuständen» in

eine Datendefinition umwandeln. Wir machen das erstmal für ein einzelnes Gürteltier und nennen das Konzept «Gürteltier auf der Straße»:

```
; Ein Gürteltier auf der Straße hat folgende Eigenschaften:
; - Gürteltier-Zustand
; - Position auf der Straße
(define-record dillo-on-road
  make-dillo-on-road
  dillo-on-road?
  (dillo-on-road-state dillo)
  (dillo-on-road-position position))
```

## 11.4 STRASSENANSICHTEN

Jetzt wo wir die Welt modelliert haben können wir sie auch auf den Bildschirm bringen. Wir fangen erstmal mit den einfachen und offensichtlichen Elementen an: der Straße Straße, inklusive der Markierung auf dem Mittelstreifen und dem Auto. Die Gürteltiere haben wir ja schon.

Damit wir einigermaßen realistisch über die Proportionen von allem nachdenken können, messen wir alles in Metern ab und wandeln zwischen Metern und Pixeln hin und her. Auf unserem Bildschirm sieht ein Verhältnis von 1m:100 Pixel gut aus. Diese Funktionen konvertieren hin und her:

```
; Meter in Pixel umwandeln
(: meters->pixels (real -> real))

(define meters->pixels
  (lambda (meters)
    (* meters 100)))

; Pixel in Meter umwandeln
(: pixels->meters (real -> real))

(define pixels->meters
  (lambda (pixels)
    (/ pixels 100)))
```



Die Straße selbst ist nur ein schwarzes Rechteck, das ist einfach. Schwieriger sind die Straßenmarkierungen, also abwechselnd ein weißer und ein schwarzer Streifen. Damit das realistisch aussieht, definieren wir erst einmal die Länge der Markierungen und die der Lücken:

```
(define marking-height 2) ; Höhe der Streifen
(define gap-height 1) ; Höhe der Lücken
```

Die Markierungen und die Lücken müssen wir abwechseln aneinanderkleben. Wie oft fragst Du? Nun, der texanische Highway ist unendlich lang, das wäre schwierig. Aber wir zeigen ja nur einen Ausschnitt, und darum machen wir nur so viele Streifen, wie in den Ausschnitt passen. Wieviele das sind, hängt von der Höhe des Ausschnitts ab. Wir könnten natürlich von Hand nachzählen, das würde aber bedeuten, dass wir uns frühzeitig auf die Höhe des Ausschnitts festlegen müssten. Um das zu vermeiden, schreiben wir erst einmal eine Funktion, die als Eingabe die Anzahl der Streifen akzeptiert. Wir folgen der Schablone für Funktionen auf natürlichen Zahlen:

```
; Straßenmarkierung mit bestimmter Anzahl von Streifen malen
(: markings (natural -> image))
```

```
(define markings
  (lambda (n)
    (cond
      ((zero? n) empty-image)
      ((positive? n)
       (above (rectangle (meters->pixels .20)
                          (meters->pixels marking-height)
                          "solid"
                          "white")
               (rectangle (meters->pixels .20)
                          (meters->pixels gap-height)
                          "solid"
                          "black")
               (markings (- n 1)))))))
```

Hier jetzt die tatsächliche Höhe des Straßenabschnitts:

```
(define road-window-height 12) ; Höhe des Straßenausschnitts
```

Die Idee, dass Du später nur diese eine Definition ändern musst, falls Dir die Ausschnittshöhe nicht passt, und sich dann alles andere automatisch anpasst.

Um die Anzahl der nötigen Markierungen zu berechnen, teilen wir die Höhe des Straßenausschnitts durch die Länge einer Markierung plus Lücke. Zur Sicherheit addieren noch eins drauf: Die Division könnte nicht ganz aufgehen. Außerdem wird möglicherweise am Ende nur ein Teil eines Streifens zu sehen sein, auch dafür ist es gut gut, wenn im Zweifelsfall einer mehr da ist. Aus der Zahl machen wir schließlich das Bild der Streifen:

```
; Anzahl der nötigen Markierungen
(define marking-count
  (+ 1
    (quotient road-window-height
      (+ marking-height gap-height))))

; sichtbare Markierungen
(define visible-markings
  (markings marking-count))
```

Zur Erinnerung: Die Funktion `quotient` haben wir auf Seite 90 eingeführt, sie teilt ganzzahlig.

Nun können wir das Bild des Straßenausschnitts zusammensetzen. Wir brauchen neben der Höhe auch noch die Breite und machen erstmal eine leere Szene nur aus schwarzem Asphalt:

```
; in Meter
(define road-width 5)

(define blank-road-window
  (empty-scene (meters->pixels road-width)
    (meters->pixels road-window-height)
    "black"))
```

Wie Mittelstreifen auf dem Straßenausschnitt platziert werden, hängt davon ab, an welchem Straßenmeter sich der Bildausschnitt befindet. Wir schreiben also eine Funktion, welche die Straßenmeter akzeptiert und ein passendes Bild liefert:

```
; Straßenausschnitt anzeigen
(: road-window (real -> image))
```

Wir benutzen `place-image/align` (siehe Seite 318), um die Mittelstreifen auf den Asphalt zu setzen:

```
(define road-window
  (lambda (meters)
```

```
(place-image/align visible-markings
  ...
  ...
  ... ...)
blank-road-window)))
```

Wir setzen zunächst den Bezugspunkt beim Mittelstreifen nach oben in die Mitte:

```
(define road-window
  (lambda (meters)
    (place-image/align visible-markings
      ...
      ...
      "center" "top"
      blank-road-window)))
```

In X-Richtung muss der Mittelstreifen in die Mitte, da können wir einfach die Breite der Straße halbieren.

Bei der Y-Koordinate ist es schwieriger: Die hängt von den Straßenmetern ab. Die Straßenmeter sind aber (außer ganz am Anfang) mehr als der Ausschnitt hoch beziehungsweise das Mittelstreifenbild lang ist. Das Mittelstreifenbild wiederholt sich aber immer, wenn ein Streifen plus Lücke vorbeigezogen ist. Wir rechnen also die Anzahl der Pixel dafür aus:

```
(define marking-segment-pixels
  (meters->pixels (+ marking-height gap-height)))
```

Durch diese Zahl teilen wir und nehmen den Divisionsrest, um die Y-Position zu bekommen. Das heißt nicht ganz – der Divisionsrest ist immer positiv, und damit ist der Mittelstreifen zu weit unten: Wir ziehen das Bild noch um einen Streifen nach oben, indem wir `marking-segment-pixels` abziehen. Das kommt dabei heraus:

```
(define road-window
  (lambda (meters)
    (place-image/align visible-markings
      (/ (image-width blank-road-window) 2)
      (- (remainder (meters->pixels meters)
                    marking-segment-pixels)
         marking-segment-pixels)
      "center" "top"
      blank-road-window)))
```

## AUFGABE 11.1

Probiere `road-window` aus. Entferne die Substraktion von `marking-segment-pixels` und untersucht, wie sich das auswirkt! □

11.5 BEWEGTE BILDER MIT `UNIVERSE.RKT`

Bisher haben wir nur statische, langweilige Bilder. Es ist Zeit, sie in Bewegung zu setzen. Dafür brauchen wir ein weiteres Teachpack namens `universe.rkt`. Um es zu aktivieren, musst Du nochmal ins Sprache/Language-Menü und dort auf Teachpack hinzufügen/Add Teachpack drücken und in dem dann auftauchenden Dialog auf `universe.rkt` und schließlich auf OK drücken.

Zum `universe.rkt`-Teachpack gehört eine Funktion namens `animate`. Was sie macht, verdeutlicht am einfachsten ein Beispiel, dass Du in die REPL eintippen kannst:

```
(animate
  (lambda (ticks)
    (place-image/align (circle 5 "solid" "red")
      100 ticks "center" "top"
      (square 200 "solid" "black")))))
```

## AUFGABE 11.2

Probier es aus! □

Du siehst einen kleinen roten Kreis sich in einem quadratischen Bild von oben nach unten bewegen.

Die neue Funktion `animate` akzeptiert eine Funktion mit einer Eingabe namens `ticks`, die ein Bild produziert:

```
(: animate ((natural -> image) -> natural))
```

`Animate` ruft die Funktion 28mal in der Sekunde auf und zeigt die entstehenden Bilder hintereinander als Animation an. Dabei übergibt `animate` als `ticks` eine natürliche Zahl, die bei Null anfängt, und jedesmal um eins größer wird: Sie misst also die Zeit, in sogenannten *Ticks*.

Wir benutzen jetzt `animate`, um das Bild des Straßenausschnitts in Bewegung zu setzen – so, dass das Wagen genau in der Mitte ist. Wir legen zuerst fest, wie schnell der Wagen fährt. Wir

haben uns für ein recht gemächliches Tempo entschieden, damit wir auch wirklich alle Gürteltiere erwischen:

```
; Meter pro Tick
(define meters-per-tick 0.1)
```

Diese Funktion ist so einfach, dass sich ausnahmsweise separates Testen nicht lohnt:

```
; Ticks in Meter umwandeln
(: ticks->meters (natural -> rational))
```

```
(define ticks->meters
  (lambda (ticks)
    (* meters-per-tick ticks)))
```

Damit können wir eine Funktion schreiben, die aus Ticks Straßenmetern macht und dann `road-window` aufruft:

```
; Straßenausschnitt zu Zeitpunkt anzeigen
(: road-window-at-ticks (natural -> image))
```

```
(define road-window-at-ticks
  (lambda (ticks)
    (road-window (ticks->meters ticks))))
```

also folgendes ausprobieren:

```
(animate road-window-at-ticks)
```

### AUFGABE 11.3

Schreibe eine andere Funktion Deiner Wahl mit der für `animate` passenden Signatur und probiere `animate` damit aus! □

## 11.6 AUTOS UND GÜRTELTIERE AUF DEM HIGHWAY

Die Straße ist fertig, als nächstes ist das Auto dran. Wir fangen mit einem Rad an:

```
; Rad
(define wheel
```

```
(rectangle (meters->pixels 0.2) (meters->pixels .5)
  "outline" "white"))
```

Als nächstes setzen wir zwei Räder zu einem Bild zusammen, das dann links und rechts von der Karosserie platziert werden:

```
; Zwei Räder auf einer Seite des Autos
```

```
(define wheels-on-one-side
  (above wheel
    (rectangle 0 (meters->pixels 1.2) "solid" "black")
    wheel))
```

```
; Breite des Autos
```

```
(define car-width 1.5)
```

```
; Länge des Autos
```

```
(define car-length 3.0)
```

```
; Bild des Autos
```

```
(define car
  (beside
    wheels-on-one-side
    (rectangle (meters->pixels car-width) (meters->pixels car-length)
      "solid" "blue")
    wheels-on-one-side))
```

Schließlich müssen wir Auto und auch die Gürteltiere noch auf der Straße platzieren. Wir müssen erstmal überlegen, welche Signatur die Funktion überhaupt hat. Klar ist, dass ein Bild herauskommen soll. Als Eingaben stehen schon einmal fest:

- das Bild, das auf der Straße platziert werden soll
- die Position des Bildes

Wir benötigen außerdem noch:

- die Zeit, weil sie bestimmt, welcher Ausschnitt der Straße angezeigt wird
- das Bild der Straße

Letzteres überrascht Dich vielleicht. Ist das Bild der Straße nicht immer das gleiche? Nein, ist es nicht! Zunächst einmal ist die Position des Mittelstreifens je nach Zeit immer unterschiedlich. Außerdem wollen wir ja nicht nur ein Objekt auf der Straße platzieren, sondern viele. Dafür füttern wir das Ergebnis der Funktion wieder in die Funktion hinein, um ein weiteres Objekt zu platzieren.

Kurzbeschreibung, Signatur und Gerüst sehen entsprechend aus wie folgt:

```
; Bild auf der Straße platzieren
(: place-image-on-road (image natural image position -> image))

(define place-image-on-road
  (lambda (road-image ticks image position)
    ...))
```

Da es sich bei `position` um zusammengesetzte Daten handelt, stehen in der Schablone Aufrufe der Selektoren:

```
(define place-image-on-road
  (lambda (road-image ticks image position)
    ...
    (position-m-from-start position)
    (position-side position)
    ...))
```

Wir müssen jetzt die Pixelkoordinaten des Bildes relativ zum Straßenausschnitt berechnen. Bei der X-Koordinate ist das verhältnismäßig einfach. Wir legen dafür eine Variable namens `pixels-from-left` an, die sich aus `(position-side position)` ergibt, was ja schon in der Schablone steht. Die Seite ist relativ zum Mittelstreifen, für dessen X-Koordinate wir ebenfalls eine Zwischenvariable anlegen:

```
; X-Koordinate der Mitte der Straße, in Pixeln
(define middle-pixels (/ (image-width road-image) 2))
```

Daraus ergibt sich `pixels-from-left` so:

```
; X-Koordinate des Mittelpunkts des Bilds
(define pixels-from-left
  (cond
    ((string=? (position-side position) "left")
     (* middle-pixels 0.5)) ; Mitte der linken Spur
    ((string=? (position-side position) "right")
     (* middle-pixels 1.5)))) ; Mitte der rechten Spur
```

Als nächstes brauchen wir die Y-Koordinate, da ist es etwas komplizierter: Wir benötigen die Y-Koordinate von `image` *relativ* zum unteren Rand des Bildes. (Oder des oberen, das geht auch.) Das wäre folgender Ausdruck:

```
(- (position-m-from-start position)
   (ticks->meters ticks))
```

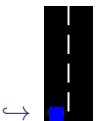
Diese Koordinate ist noch in Metern, die müssen wir noch in Pixel umrechnen. Außerdem müssen wir berücksichtigen, dass die Pixel-Koordinaten von oben nach unten laufen, die Straßenmeter aber von unten nach oben – wir ziehen also diese Zahl von der Höhe des Bildes ab. Damit können wir wieder einmal `place-image/align` (siehe Seite 318) bemühen, um das Bild auf die Straße zu setzen. Die Funktion sieht vollständig so aus:

```
(define place-image-on-road
  (lambda (road-image ticks image position)
    ; X-Koordinate der Mitte der Straße, in Pixeln
    (define middle-pixels (/ (image-width road-image) 2))
    ; X-Koordinate des Mittelpunkts des Bilds
    (define pixels-from-left
      (cond
        ((string=? (position-side position) "left")
         (* middle-pixels 0.5)) ; Mitte der linken Spur
        ((string=? (position-side position) "right")
         (* middle-pixels 1.5)))) ; Mitte der rechten Spur

    (place-image/align
     image
     pixels-from-left
     (- (image-height road-image)
        (meters->pixels (- (position-m-from-start position)
                           (ticks->meters ticks))))
     "center" "center"
     road-image)))
```

Die Funktion ist kompliziert genug, dass wir sie testen sollten. Einen `check-expect`-Test vorab zu konstruieren ist schier unmöglich, aber wir können die Funktion zumindest in der REPL ausprobieren, am einfachsten bei 0 Ticks, also am Anfang des Spiels:

```
(place-image-on-road (road-window 0) 0 car (make-position 0 "left"))
```





Da ist noch ein Problem! Das Auto ist am unteren Rand und nur zur Hälfte zu sehen, wir wollen es aber gern in der Mitte. Wir müssen also noch die Hälfte der Höhe des Straßenabschnitts abziehen, entsprechend den Aufruf von `place-image/align` korrigieren:

```
(place-image/align
  image
  pixels-from-left
  (- (image-height road-image)
     (meters->pixels (+ (- (position-m-from-start position)
                           (ticks->meters ticks))
                        (/ road-window-height 2))))
  "center" "center"
  road-image)
```

Jetzt ist alles bei 0 in der Mitte!

Diese Funktion können wir jetzt benutzen, um zunächst das Autobild auf dem Straßenabschnitt zu platzieren:

```
; Auto auf die Straße setzen
(: place-car-on-road (natural position image -> image))

(define place-car-on-road
  (lambda (ticks car-position road-image)
    (place-image-on-road road-image
                        ticks
                        car car-position)))
```

#### AUFGABE 11.4

Konstruiere auf Basis von `place-car-on-road` eine Funktion, die Du an `animate` übergeben kannst und probiere sie aus! □

Ähnlich machen wir das für das Gürteltier-Bild. Die nötigen Angaben, zu Position und Zustand des Gürteltiers ziehen wir aus dem `dillo-on-road`-Record:

```
; Ein Tier auf die Straße malen
(: place-dillo-on-road (natural dillo-on-road image -> image))
```

```
(define place-dillo-on-road
  (lambda (ticks dillo-on-road road-image)
    (place-image-on-road road-image
                          ticks
                          (dillo-image (dillo-on-road-state dillo-on-
road))
                          (dillo-on-road-position dillo-on-road))))
```

Jetzt gibt es ja auf der Straße nicht nur ein Gürteltier sondern viele, die als Liste im `world`-Record stecken. Dafür brauchen wir eine Funktion mit folgender Kurzbeschreibung und Signatur:

```
; Alle Tiere auf die Straße malen
(: place-dillos-on-road
  (natural (list-of dillo-on-road) image -> image))
```

Wir könnten das nach Konstruktionsanleitung für Listen programmieren. Einfacher ist aber, die universelle Funktion `fold` zu benutzen. (Siehe Abschnitt 9.4 auf Seite 257.) Das sieht dann so aus:

```
(define place-dillos-on-road
  (lambda (ticks dillos-on-road road-image)
    (fold road-image
          (lambda (dillo-on-road image)
            (place-dillo-on-road ticks dillo-on-road image))
          dillos-on-road)))
```

Neben Auto und Gürteltieren fehlt noch die Punktzahl oben links in der Ecke. Das erledigt folgende Funktion:

```
; Punktzahl anzeigen
(: place-score (natural image -> image))

(define place-score
  (lambda (score image)
    (place-image/align
      (text (number->string score) 100 "blue")
      0 0
      "left" "top"
      image)))
```

## 11.7 DIE WELT EIN HIGHWAY

Es wird Zeit, dass wir alles zusammensetzen zu einem Spiel, das sich bewegt und dass auf Knopfdruck reagiert. Das `universe.rkt`-Teachpack fordert, dass wir eine Datendefinition erstellen, in der der gesamte Zustand des Spiels steckt. In Abschnitt 11.3 auf Seite 324 hatten wir schon folgende Aspekte aufgelistet:

- die Position des Autos
- die Positionen der Gürteltiere zusammen mit ihren Zuständen
- der Punktestand

Wir wissen, dass uns das `universe.rkt`-Teachpack später die Zeit als Ticks liefern wird: Daraus können wir die Straßenmeter der Position des Autors berechnen. Nur noch die Seite müssen wir explizit repräsentieren. Es entsteht also folgende etwas angepasste Datendefinition:

```
; Die Welt des Spiels besteht aus:
; - Ticks seit Spielanfang
; - Seite, auf der das Auto fährt
; - Tiere auf der Straße
; - Punktzahl
(define-record world
  make-world
  world?
  (world-ticks natural)
  (world-car-side side)
  (world-dillos-on-road (list-of dillo-on-road))
  (world-score natural))
```

Hier ein Beispiel dafür:

```
; Vier Gürteltiere auf der Straße
(define dillos-on-road
  (list (make-dillo-on-road dillo1 (make-position 20 "left"))
        (make-dillo-on-road dillo2 (make-position 26 "right"))
        (make-dillo-on-road dillo3 (make-position 30 "left"))
        (make-dillo-on-road dillo4 (make-position 42 "left"))))

; Welt am Anfang, Auto steht links
(define initial-world
```

```
(make-world 0 "left" dillos-on-road 0))
```

Wir werden für `place-car-on-road` die Position des Autos als `position`-Record benötigen, die noch nicht direkt im `world`-Record vorhanden ist. Die Seite steht immerhin schon da. Die Straßenmeter berechnen wir aus den Ticks. Die Funktion dafür hat Kurzbeschreibung, Signatur und Testfälle wie folgt:

```
; Position des Autos
(: world-car-position (world -> position))

(check-expect (world-car-position (make-world 0 "left" empty 0))
              (make-position 0 "left"))
(check-expect (world-car-position (make-world 100 "right" empty 0))
              (make-position (ticks->meters 100) "right"))
```

Die Schablone für die Funktion berücksichtigt, dass es sich sowohl bei Ein- als auch bei der Ausgabe jeweils um zusammengesetzte Daten handelt:

```
(define world-car-position
  (lambda (world)
    (make-position ... ...)
    ...
    (world-ticks world)
    (world-car-side world)
    (world-dillos-on-road world)
    (world-score world)
    ...))
```

Nur die ersten beiden Felder von `world` sind relevant für die Position des Autos – Gürteltiere und Punktzahl sind sichtlich nicht relevant. Für die Position brauchen wir die Straßenmeter, die wir aus den Ticks mit Hilfe von `ticks->meters` berechnen können:

```
(define world-car-position
  (lambda (world)
    (make-position (ticks->meters (world-ticks world))
                  (world-car-side world))))
```

Jetzt haben wir endlich alle Funktionen, die wir benötigen, um die ganze Welt anzuzeigen. Wir müssen sie nur kombinieren, um alles in einem `world`-Record als Straßenabschnitt anzuzeigen:

```
; Spiel anzeigen
(: world->image (world -> image))

(define world->image
  (lambda (world)
    (define ticks (world-ticks world))
    (place-score
      (world-score world)
      (place-car-on-road
        ticks
        (world-car-position world)
        (place-dillos-on-road
          ticks
          (world-dillos-on-road world)
          (road-window ticks))))))
```

Du kannst jetzt natürlich

```
(world->image initial-world)
```

ausprobieren, aber da ist nur der anfängliche, leere Straßenabschnitt zu sehen.

Besser wird es mit `animate`. Dafür brauchen wir eine Funktion mit Signatur `(natural -> image)`. Da wir schon `world->image` mit der Signatur `(world -> image)` haben, brauchen wir noch eine Funktion mit Signatur `(natural -> world)`, die ein passendes `world`-Objekt erzeugt:

```
; Statische Welt zu einem bestimmten Zeitpunkt darstellen
(: world-at-ticks (natural -> world))

(define world-at-ticks
  (lambda (ticks)
    (make-world ticks "left" dillos-on-road 0)))
```

Nun können wir `animate` mit einer Kombination aus den beiden aufrufen:

```
(animate (lambda (ticks) (world->image (world-at-ticks ticks))))
```

Da sieht man die Straße vorbeiziehen, und schließlich tauchen auch Gürteltiere auf. Aber das Auto ist links festgeklebt und die Gürteltiere sterben auch nicht, wenn das Auto drüberfährt.

## 11.8 ÜBERFAHREN KONKRET

Was fehlt noch, damit aus dem `world`-Record alle Aspekte des Spiels berechnet werden können, mithin der texanische Highway simuliert:

1. Wir müssen feststellen ob das Auto die Position eines Gürteltiers berührt und dieses damit überfährt.
2. Wir müssen zählen, wieviele Tiere vor einem gegebenen Zeitpunkt vom Auto touchiert, mithin überfahren werden, damit wir die Punktzahl entsprechend erhöhen können.
3. Schließlich müssen wir alle Tiere überfahren, deren Position unter das Auto gerät.

Diese Liste arbeiten wir in diesem Abschnitt ab.

**Berührung eines Gürteltiers** | Als nächstes auf der Liste auf Seite 340 müssen wir feststellen, ob das Auto die Position eines Gürteltiers berührt.

Wir brauchen eine Funktion mit Kurzbeschreibung und Signatur wie folgt:

```
; Berührt das Auto eine Position?
(: car-on-position? (position position -> boolean))
```

Das Auto berührt eine Position, wenn die Position auf der gleichen Straßenseite nah dran ist. Hier die Testfälle dazu:

```
(check-expect (car-on-position? (make-position 10 "left")
                                (make-position 10 "right"))
              #f)
(check-expect (car-on-position? (make-position 10 "left")
                                (make-position 10 "left"))
              #t)
(check-expect (car-on-position? (make-position 11 "left")
                                (make-position 10 "left"))
              #t)
(check-expect (car-on-position? (make-position 10 "left")
                                (make-position 11 "left"))
              #t)
```

Für die Schablone haben wir es bei beiden Eingaben mit zusammengesetzten Daten zu tun:

```
(define car-on-position?
  (lambda (car-position position)
    ...
```

```
(position-side car-position)
(position-side position)
(position-m-from-start car-position)
(position-m-from-start position)
...))
```

Die beiden Straßenseiten müssen gleich sein. Bei den Straßenmetern der beiden Positionen ist relevant, dass die Straßenmeter des Autos genau in der Mitte des Autos sind – die Position darf nicht mehr als die Hälfte davon entfernt sein. Um den Vergleich zu erleichtern, egal ob die Position vor oder hinter dem Mittelpunkt ist, benutzen wir die eingebaute Funktion `abs`. Diese Funktion berechnet den *Absolutbetrag* einer Zahl: Sie macht aus negativen Zahlen die entsprechenden positiven Zahlen:

```
(abs 5)
↪ 5
(abs -5)
↪ 5
```

Hier die fertige Funktion:

```
(define car-on-position?
  (lambda (car-position position)
    (and (string=? (position-side car-position)
                   (position-side position))
         (<= (abs (- (position-m-from-start car-position)
                    (position-m-from-start position)))
              (/ car-length 2)))))
```

**Wieviele Tiere sind gerade unter dem Auto?** | Als nächstes schreiben wir eine Funktion, welche die Tiere unter dem Auto zählt, damit wir die Punktzahl entsprechend erhöhen können. Genauer gesagt zählen wir nur die *lebendigen* Gürteltiere: Für das nochmalige Überfahren eines toten Gürteltiers gibt es keinen Punkt. Hier Kurzbeschreibung und Signatur:

```
; Wieviele Tiere werden vom Auto berührt?
(: live-dillos-under-car-count
  (position (list-of dillo-on-road) -> natural))
```

Für den Test sollten wir in der Eingabe sowohl lebendige als auch tote Gürteltiere unterbringen, ebenso wie solche, die unter dem Auto liegen, und solche, die weiter weg sind:

```
(check-expect
  (live-dillos-under-car-count
    (make-position 10 "left")
    (list (make-dillo-on-road dillo1 (make-position 10 "left"))
          (make-dillo-on-road dillo2 (make-position 10 "right"))
          (make-dillo-on-road dillo2 (make-position 11 "left"))
          (make-dillo-on-road dillo1 (make-position 9 "left"))
          (make-dillo-on-road dillo1 (make-position 12 "left")))))
  2)
```

Der Test zählt das erste und das vorletzte Gürteltier.

Für die Definition könnten wir die Schablone für Funktionen auf Listen heranziehen. Da es aber darum geht, aus einer Liste einige Elemente auszuwählen, die einem bestimmten Kriterium entsprechen, können wir auch die Funktion `filter` von Seite 251 benutzen, die eingebaute Version von `extract-list`. Damit extrahieren wir die zu zählenden Gürteltiere. Die müssen wir nur noch mit der ebenfalls eingebauten Funktion `length` zählen. (`length` ist die eingebaute Version von `list-length`, siehe Seite 178.)

```
(define live-dillos-under-car-count
  (lambda (car-position dillos-on-road)
    (length
      (filter (lambda (dillo-on-road)
                (and (car-on-position?
                     car-position
                     (dillo-on-road-position dillo-on-road))
                     (dillo-alive?
                      (dillo-on-road-state dillo-on-road)))))
              dillos-on-road))))
```

**Gürteltiere tatsächlich überfahren** | Als letzter Arbeitsauftrag bleibt noch, die Gürteltiere, die vom Auto touchiert werden, vom lebenden in den toten Zustand zu überführen. Die Funktion akzeptiert wie `live-dillos-under-car-count` die Position des Autos und die Gürteltiere:

```
; Alle Tiere überfahren, die das Auto berührt
(: run-over-dillos-on-road
  (position (list-of dillo-on-road) -> (list-of dillo-on-road)))
```



Beim Testen gibt es jede Menge tote Gürteltiere, darum definieren wir den dazugehörigen Zustand zwecks Wiederverwendung:

```
(define dead-dillo1 (run-over-dillo dillo1))

(check-expect
  (run-over-dillos-on-road
    (make-position 10 "left")
    (list (make-dillo-on-road dillo1 (make-position 10 "left"))
          (make-dillo-on-road dillo1 (make-position 10 "right"))
          (make-dillo-on-road dillo1 (make-position 11 "left"))
          (make-dillo-on-road dillo1 (make-position 9 "left"))
          (make-dillo-on-road dillo1 (make-position 12 "left"))))
    (list (make-dillo-on-road dead-dillo1 (make-position 10 "left"))
          (make-dillo-on-road dillo1 (make-position 10 "right"))
          (make-dillo-on-road dead-dillo1 (make-position 11 "left"))
          (make-dillo-on-road dead-dillo1 (make-position 9 "left"))
          (make-dillo-on-road dillo1 (make-position 12 "left"))))
```

Die Funktion wendet also eine Operation auf jedes Element der Liste an und gibt uns die Ergebnisse in einer Liste zurück: Dafür können wir wieder eine Higher-Order-Funktion aus Kapitel 9 verwenden, nämlich `map`, die eingebaute Version von `list-map`, siehe Seite 257.

```
(define run-over-dillos-on-road
  (lambda (car-position dillos-on-road)
    (map (lambda (dillo-on-road)
          ...)
         dillos-on-road)))
```

Bei dem Argument der `lambda`-Funktion handelt es sich um zusammengesetzte Daten. Es kommen also Selektoraufrufe in die Schablone:

```
(define run-over-dillos-on-road
  (lambda (car-position dillos-on-road)
    (map (lambda (dillo-on-road)
          ...
          (dillo-on-road-position dillo-on-road)
          (dillo-on-road-state dillo-on-road)
          ...)
         dillos-on-road)))
```

```
dillos-on-road)))
```

Die Gürteltiere fallen in zwei Klassen: Die, die vom Auto gerade touchiert werden, und alle anderen. Wir können die Funktion `car-on-position?` benutzen, um eine binäre Verzweigung zu bilden:

```
(define run-over-dillos-on-road
  (lambda (car-position dillos-on-road)
    (map (lambda (dillo-on-road)
          (if (car-on-position?
                car-position
                (dillo-on-road-position dillo-on-road))
              ...
              ...))
         dillos-on-road)))
```

Im ersten Fall sollten wir das Gürteltier überfahren, das machen wir mit `run-over-dillo`. Im anderen Fall geben wir `dillo-on-road` unverletzt zurück:

```
(define run-over-dillos-on-road
  (lambda (car-position dillos-on-road)
    (map (lambda (dillo-on-road)
          (if (car-on-position?
                car-position
                (dillo-on-road-position dillo-on-road))
              (make-dillo-on-road
                (run-over-dillo (dillo-on-road-state dillo-on-road))
                (dillo-on-road-position dillo-on-road))
              dillo-on-road))
         dillos-on-road)))
```

#### AUFGABE 11.5

Was ist die Konsequenz, wenn auch das Überfahren toter Gürteltiere Punkte geben würde? Versuche, das hier schon einmal durch reines Nachdenken zu klären und probiere es später aus! ☐

## 11.9 REAKTIVE ANIMATION

An diesem Punkt haben wir alle Grundaspekte des Spiels durch Funktionen beschrieben. Wir müssen sie noch zu einem Spiel zusammensetzen. Das sollte dann insbesondere auch auf Tastendrucke reagieren. Für solche reaktiven Animationen brauchen wir noch Unterstützung vom `universe.rkt`-Teachpack, das dafür extra ein neues Programmiersprachenelement mitbringt, das heißt `big-bang`. Es startet das Spiel und hat folgende Form:

```
(big-bang world clause ...)
```

Der erste Operand `world` ist die Welt am Anfang des Spiels, danach kommen einige Klauseln, die beschreiben, was das Spiel machen soll, wenn bestimmte Sachen passieren. Das kann sein:

- eine Taste auf der Tastatur wird gedrückt
- eine Taste auf der Tastatur wird losgelassen
- die Maus wird bewegt oder eine Maustaste wird gedrückt

Außerdem gibt es eine Klausel, die festlegt, wie aus der Welt ein Bild generiert wird. Die sieht so aus:

```
(to-draw function)
```

... und die Funktion muss die Signatur `(world -> image)` erfüllen. Damit reicht gerade so aus, um `big-bang` ein erstes Mal auszuprobieren:

```
(big-bang initial-world  
  (to-draw world->image))
```

Da wird immerhin die Welt am Anfang angezeigt, die sich aber noch nicht bewegt. Dafür könnte sie auf Tastendruck reagieren, das geht mit einer Klausel folgender Form:

```
(on-key function)
```

Diese Funktion muss folgende Signatur haben:

```
(world string -> world)
```

Sie wird immer dann aufgerufen, wenn eine Taste gedrückt wird – und zwar mit der Welt vor dem Tastendruck und einer Zeichenkette, die der Taste entspricht. Also zum Beispiel `"a"` für die A-Taste, `"+"` für die +-Taste undsoweiter. Für die «Sondertasten» sieht das so aus:

"left"	Linkspfeil
"right"	Rechtspfeil
"up"	Obenpfeil
"down"	Untenpfeil
"shift"	linke Shift-Taste
"rshift"	rechte Shift-Taste

Für unser Spiel brauchen wir nur die Links- und Rechts-Pfeiltasten, die das Auto auf die linke oder rechte Straßenseite befördern. Wir brauchen also eine Funktion mit folgender Signatur:

```
; Auf Tastendruck reagieren
(: react-to-key (world string -> world))
```

Wir testen beide Pfeiltasten – alle anderen Tasten sollten die Welt unverändert lassen:

```
(check-expect
  (react-to-key (make-world 12 "right" dillos-on-road 5) "left")
  (make-world 12 "left" dillos-on-road 5))

(check-expect
  (react-to-key (make-world 12 "left" dillos-on-road 5) "right")
  (make-world 12 "right" dillos-on-road 5))

(check-expect
  (react-to-key (make-world 12 "left" dillos-on-road 5) "a")
  (make-world 12 "left" dillos-on-road 5))
```

Hier das Gerüst:

```
(define react-to-key
  (lambda (world key)
    ...))
```

Bei `key` handelt es sich effektiv um eine Aufzählung: links, rechts, alles andere. Die Schablone sieht also so aus:

```
(define react-to-key
  (lambda (world key)
    (cond
      ((string=? key "left") ...)
      ((string=? key "right") ...)
```

```
(else ...)))
```

Außerdem steuern die Konstruktionsanleitungen Schablonen für zusammengesetzte Daten als Ein- und als Ausgabe bei:

```
(define react-to-key
  (lambda (world key)
    (cond
      ((string=? key "left")
       (make-world ...
                   ...
                   ...
                   ...))
      ((string=? key "right")
       (make-world ...
                   ...
                   ...
                   ...))
      (else ...))
    ...
    (world-ticks world)
    (world-car-side world)
    (world-dillos-on-road world)
    (world-score-road world)
    ...))
```

Bei Links- und Rechtspfeil ändert sich jeweils nur die Seite des Autos. Bei allen anderen Tasten ändert sich nichts:

```
(define react-to-key
  (lambda (world key)
    (cond
      ((string=? key "left")
       (make-world (world-ticks world)
                   "left"
                   (world-dillos-on-road world)
                   (world-score world)))
      ((string=? key "right")
```

```

(make-world (world-ticks world)
            "right"
            (world-dillos-on-road world)
            (world-score world)))
(else world)))

```

Damit können wir den **big-bang**-Aufruf erweitern:

```

(big-bang initial-world
  (to-draw world->image)
  (on-key react-to-key))

```

Die Straße steht zwar immer noch permanent auf Anfang, aber wir können mit den Pfeiltasten das Auto nach links und rechts bewegen.

Dass die Straße sich bewegt, hatten wir schonmal, aber da haben wir **animate** benutzt. Bei **big-bang** funktioniert das ein wenig anders, da gibt es eine Klausel dieser Form:

```
(on-tick function)
```

Die Funktion bei **on-tick** wird immer aufgerufen, wenn ein Tick verstreicht, und muss die folgende Signatur haben:

```
(world -> world)
```

Das heißt, dass die Funktion nicht wie bei **animate** die Anzahl der Ticks geliefert bekommt sondern selbst zählen muss. Wir fangen mal an, eine solche Funktion zu schreiben:

```

; Wie verändert sich die Welt, wenn ein Tick Zeit vergeht?
(: next-world (world -> world))

```

Gerüst und Schablone nutzen aus, dass sowohl Ein- als auch Ausgabe zusammengesetzte Daten sind:

```

(define next-world
  (lambda (world)
    (make-world ... ..)
    ...
    (world-ticks world)
    (world-car-side world)
    (world-dillos-on-road world)
    (world-score world)
    ...))

```

Wir schreiben erst einmal eine ganz einfache Funktion, bei der nur die Zeit vergeht:

```
(define next-world
  (lambda (world)
    (make-world (+ 1 (world-ticks world))
                (world-car-side world)
                (world-dillos-on-road world)
                (world-score world))))
```

Die Funktion binden wir jetzt in den **big-bang**-Ausdruck ein:

```
(big-bang initial-world
  (to-draw world->image)
  (on-tick next-world)
  (on-key react-to-key))
```

Das sieht schon besser aus: Die Straße bewegt sich, und man kann das Auto nach links und rechts steuern. Aber überfahrene Gürteltiere sterben noch nicht, und die Punktzahl geht frustrierenderweise auch noch nicht hoch. Das liegt daran, dass wir die Funktionen dafür – **run-over-dillos-on-road** und **live-dillos-under-car-count** – noch nicht eingesetzt haben. Das holen wir nach:

```
(define next-world
  (lambda (world)
    (define car-position (world-car-position world))
    (define dillos-on-road (world-dillos-on-road world))
    (make-world (+ 1 (world-ticks world))
                (world-car-side world)
                (run-over-dillos-on-road car-position dillos-on-road)
                (+ (live-dillos-under-car-count
                    car-position dillos-on-road)
                   (world-score world)))))
```

Voilà, viel Spaß beim Gürteltiere-überfahren!

#### AUFGABE 11.6

Ändere das Spiel so, dass auch das Überfahren toter Gürteltiere Punkte gibt! Ist das sinnvoll? ☐

Hier sind noch einige weitere nützliche Klauseln für **big-bang**, um etwas anspruchsvollere Spiele zu schreiben::

(on-mouse *function*)

Die Funktion *function* wird immer dann aufgerufen, wenn die Maus bedient wird – also wenn sie bewegt wird, oder wenn ein Knopf gedrückt wird. Sie muss folgende Signatur haben:

```
(world integer integer mouse-event -> world)
```

Die beiden *integer*-Argumente sind die X- und Y-Koordinaten der Mausposition.

Die Signatur *mouse-event* ist folgendermaßen definiert:

```
(define mouse-event
  (signature
    (enum "button-down" "button-up" "drag" "move" "enter" "leave")))
```

Die einzelnen Fälle haben folgende Bedeutungen:

"button-down"	Der Maus-Knopf wurde gedrückt.
"button-up"	Der Maus-Knopf wurde losgelassen.
"drag"	Die Maus wird gerade gezogen – die Maus bewegt sich also, während der Knopf gedrückt ist.
"move"	Die Maus wurde bewegt.
"enter"	Der Maus-Cursor wurde ins !universe.rkt!-Fenster bewegt.
"leave"	Der Maus-Cursor wurde aus dem universe.rkt-Fenster herausbewegt.

Es gibt eine weitere coole Klausel namens *on-pad*:

(on-pad *function*)

Wenn die spezifiziert ist, wird ein «Game-Pad» angezeigt, das so aussieht:



Die Funktion hat folgende Signatur:

```
(world pad-event -> world)
```

```
(define pad-event
  (signature
    (enum "left" "right" "up" "down" "w" "s" "a" "d" " " "shift" "
    rshift")))
```



Die `pad-event`-Werte entsprechen gerade den Tasten auf dem Pad.

#### AUFGABE 11.7

Erweitere den Aufruf von `big-bang` so, dass man das Spiel auch mit dem Pad spielen kann! ☐

## AUFGABEN

#### AUFGABE 11.8

Mache das Gürteltier-Spiel besser!

#### AUFGABE 11.9

Schreibe ein Telespiel Deiner Wahl.



## 12 BÄUME

Bäume sind eine Form von Daten, die (wie Listen) besonders oft in der Informatik vorkommt. Oft ergeben sich baumförmige Datendefinitionen aus der Problemstellung. Wenn wir über diese Datendefinitionen abstrahieren, entsteht eine universell verwendbare Form von Daten, der *Binärbaum*. Diese Binärbäume sind ähnlich vielseitig wie Listen und erlauben uns außerdem, Daten so in Bäumen zu organisieren, dass wir sie schnell wiederfinden können.

### 12.1 STAMMBÄUME

Die Idee, den Baum als Metapher für eine bestimmte Form von Daten zu benutzen, findet sich bereits in der Bibel, die Wörter wie «Baumstumpf» und «Spross» benutzt, um Abstammung zu beschreiben. Erste bildliche Darstellungen von Stammbäumen sind aus diesen Beschreibungen ab dem 11. Jahrhundert abgeleitet worden.

Das Stammbaum-Beispiel ist etwas ausgetreten, aber dennoch instruktiv. In Stammbäumen sind in der Regel für eine Person ihr Name sowie Verbindungen zu den beiden Eltern vermerkt. Das führt zu folgender Datendefinition.

```
; Eine Person hat folgende Eigenschaften:  
; - Name  
; - Elternteil #1  
; - Elternteil #2
```

Diese Definition lässt offen, um was für Daten es sich bei den beiden Elternteilen handelt. Natürlich sind es auch Personen, aber wenn wir in einem Stammbaum weit genug nach oben gehen, sind diese irgendwann unbekannt: Jeder konkrete Stammbaum endet irgendwo. Wir brauchen also auch noch eine Repräsentation für einen «unbekannten Elternteil» ohne (bekannte) Eigenschaften:

```
; Ein unbekannter Elternteil hat keine Eigenschaften  
(define-record unknown-parent  
  make-unknown-parent  
  unknown-parent?)
```

Daraus entsteht eine Datendefinition für «Elternteil»:

```
; Ein Elternteil ist eins der folgenden:
```

```
; - eine Person
; - ein unbekannter Elternteil
```

Diese – und die Datendefinition für «Person» können wir nun in Code übersetzen:

```
(define-record person
  make-person
  person?
  (person-name string)
  (person-parent-1 parent)
  (person-parent-2 parent))

(define parent
  (signature
    (mixed person unknown-parent)))
```

Für das unbekannte Elternteil stellen wir gleich mal einen Wert her:

```
(define an-unknown-parent (make-unknown-parent))
```

Hier ein kleiner Stammbaum als Beispiel:

```
(define slash
  (make-person "Slash"
    (make-person "Ola Hudson"
      an-unknown-parent
      an-unknown-parent)
    (make-person "Anthony Hudson"
      an-unknown-parent
      an-unknown-parent)))

(define london-hudson
  (make-person "London Hudson"
    slash
    (make-person "Perla Ferrar"
      an-unknown-parent
      an-unknown-parent)))
```

Wir schreiben nun eine Funktion, die feststellen soll, ob jemand der Vorfahr einer Person ist, so etwa:

```
; Ist jemand Vorfahr:in einer Person?
(: ancestor? (string person -> boolean))
```

```
(check-expect (ancestor? "Slash" london-hudson) #t)
(check-expect (ancestor? "Axl" london-hudson) #f)
```

Die Schablone für diese Funktion sieht folgendermaßen aus:

```
(define ancestor?
  (lambda (name person)
    ...
    (person-name person)
    (person-parent-1 person)
    (person-parent-2 person)
    ...))
```

Was können wir mit diesen Bestandteilen anfangen? Den Namen der Person könnten wir mit dem gesuchten Namen vergleichen – wenn ja, handelt es sich um einen Vorfahren:

```
(define ancestor?
  (lambda (name person)
    (if (string=? name (person-name person))
        #t
        ...))
    (person-parent-1 person)
    (person-parent-2 person)
    ...))
```

Bei `(person-parent-1 person)` und `(person-parent-2 person)` handelt es sich um gemischte Daten. Wir könnten die nötige Verzweigung direkt in `ancestor?` einbauen. Genauso können wir eine separate Funktion schreiben, welche die Frage beantwortet, ob ein Elternteil Vorfahr ist. Da es zwei Elternteile gibt, lohnt sich tendenziell eine solche separate Funktion mit Kurzbeschreibung und Signatur wie folgt:

```
; Ist jemand Vorfahr:in eines Elternteils?
(: parent-ancestor? (string parent -> boolean))
```

Diese Funktion schreiben wir gleich im Anschluss. Aber ihre Signatur ist genug, damit wir die Schablone in Funktion `ancestor?` weiter ausfüllen, indem wir überprüfen, ob Elternteil Nr. 1 oder Nr. 2 Vorfahr ist:

```
(define ancestor?
  (lambda (name person)
    (if (string=? name (person-name person))
        #t
        (if (or (parent-ancestor? name (person-parent-1 person))
                (parent-ancestor? name (person-parent-2 person)))
            #t
            #f))))
```

Diese Funktion ist schon korrekt, aber sie könnte noch etwas eleganter sein. Der zweite `if`-Ausdruck liefert `#t`, falls die Bedingung `#t` und `#f`, falls die Bedingung `#f` liefert: Es kommt also immer das Ergebnis der Bedingung heraus. Das ist eine allgemein anwendbare Regel:

```
(if b #t #f) = b
```

Wir können `ancestor?` also verkürzen auf:

```
(define ancestor?
  (lambda (name person)
    (if (string=? name (person-name person))
        #t
        (or (parent-ancestor? name (person-parent-1 person))
            (parent-ancestor? name (person-parent-2 person))))))
```

Auch den verbleibenden `if`-Ausdruck können wir noch loswerden, weil er `#t` ergibt, wenn die Bedingung `#t` ergibt oder wenn der `or`-Ausdruck `#t` liefert. Wir können deshalb die Funktion mit einem großen `or` schreiben:

```
(define ancestor?
  (lambda (name person)
    (or (string=? name (person-name person))
        (parent-ancestor? name (person-parent-1 person))
        (parent-ancestor? name (person-parent-2 person)))))
```

Notwendig war diese Vereinfachung nicht, aber schöner sieht das Resultat schon aus, finden wir!

Es fehlt noch die Hilfsfunktion `parent-ancestor?`. Hier sind ein paar Tests:

```
(check-expect (parent-ancestor? "Slash" london-hudson) #t)
(check-expect (parent-ancestor? "Axl" london-hudson) #f)
(check-expect (parent-ancestor? "Slash" an-unknown-parent) #f)
```

Gerüst und Schablone ergeben sich – wie immer – aus der Datendefinition von `parent`:

```
(define parent-ancestor?
  (lambda (name parent)
    (cond
      ((person? parent) ...)
      ((unknown-parent? parent) ...))))
```

Für den ersten Fall können wir `ancestor?` benutzen, im zweiten Fall können wir mit `#f` antworten:

```
(define parent-ancestor?
  (lambda (name parent)
    (cond
      ((person? parent)
       (ancestor? name parent))
      ((unknown-parent? parent)
       #f))))
```

Fertig!

#### AUFGABE 12.1

Ändere die Funktion `ancestor?` dahingehend, dass eine Person nicht ihr eigener Vorfahr ist. Achte darauf, dass ansonsten die Funktion noch richtig arbeitet! Wird die Funktion einfacher? ☐

## 12.2 BINÄRBÄUME

Wir schauen uns nochmal die Record-Definition von `person` an:

```
(define-record person
  make-person
  person?
  (person-name string)
  (person-parent-1 parent)
  (person-parent-2 parent))
```

Vielleicht hat Dich das an eine Record-Definition aus Kapitel 5 erinnert:

```
(define-record confluence
  make-confluence
  confluence?
  (confluence-location string)
  (confluence-main-stem river)
  (confluence-tributary river))
```

Die Struktur ist bei beiden Definitionen gleich. Insbesondere enthalten beide Definitionen jeweils zwei Selbstreferenzen. Bei `person` ist die Selbstreferenz auf `parent`, das so definiert ist:

```
(define parent
  (signature
    (mixed person unknown-parent)))
```

Bei `confluence` ist die Selbstreferenz auf `river`:

```
(define river
  (signature (mixed creek confluence)))
```

Die jeweils anderen Fälle von `parent` und `person` unterscheiden sich leicht:

```
(define-record unknown-parent
  make-unknown-parent
  unknown-parent?)
```

```
(define-record creek
  make-creek
  creek?
  (creek-origin string))
```

In beiden steckt selbst aber keine Selbstreferenz mehr. Beide Datendefinitionen bilden baumartige Strukturen ab: Ein `person`- oder `confluence`-Record bildet einen Ast, der zweifach verzweigt. Ein Baum endet jeweils bei `unknown-parent`- oder `creek`-Records. Weil die «inneren» Äste immer zweifach verzweigen, handelt es sich in beiden Fällen um *Binärbäume*.

Über diese beiden Sätze von Definitionen können wir abstrahieren. Fangen wir mit `person` und `confluence` an. Der gängige Name für die Verzweigungen innerhalb eines Binärbaums ist *Knoten* oder *innerer Knoten*, auf Englisch *node*. Wir brauchen außerdem einen Namen für die «Namensdaten», die bei beiden noch dabei sind. Üblich ist *Markierung*, auf Englisch *label*. Die Signatur für den Selbstbezug nennen wir einfach `tree`:



```
(define-record node
  make-node node?
  (node-label string)
  (node-left-branch tree)
  (node-right-branch tree))
```

Das Wort «branch» heißt wörtlich übersetzt «Zweig», wir verwenden aber die Begriffe «linker Teilbaum» und «rechter Teilbaum», was im Deutschen üblicher ist.

Bei der Definition für `tree` brauchen wir noch einen Namen für die Werte an den Rändern des Baums – genannt *Blätter*, auf Englisch *leaf*.

```
(define tree
  (signature (mixed leaf node)))
```

Es fehlt noch die Definition von `leaf`. Hier ist es nicht ganz so einfach, weil `creek` noch einen Namen enthält, `unknown-parent` aber nicht. Wir müssen also über beide abstrahieren. Einen Namen haben wir ja schon – `leaf` – es fehlt noch das `lambda`:

```
(define tree-of
  (lambda (leaf)
    (signature (mixed leaf node))))
```

Das zieht noch eine weitere Änderung nach sich, weil `tree` ja in der Definition von `node` verwendet wird. Wir müssen da entsprechend den `leaf`-Parameter mit durchziehen:

```
(define-record (node-of leaf)
  make-node node?
  (node-label string)
  (node-left-branch (tree-of leaf))
  (node-right-branch (tree-of leaf)))
```

Die Notation für die Abstraktion der Record-Signatur mit den Extra-Klammern um `(node-of leaf)` haben wir bisher erst einmal gesehen, bei der Definition von `cons-list-of` in Abschnitt 6.2 auf Seite 173.

Wir könnten an dieser Stelle fertig sein. Wir nehmen aber noch eine Verallgemeinerung vor: Wie wir sehen werden, müssen die Markierungen in Bäumen nicht unbedingt Zeichenketten sein – wir werden da noch andere Arten von Werten ablegen wollen. Darum abstrahieren wir auch über die Signatur der Markierungen noch. Außerdem reichen wir noch die Datendefinitionen nach:

```
; Ein Knoten besteht aus
```

```

; - Markierung
; - linken Ast
; - rechter Ast
(define-record (node-of leaf label)
  make-node node?
  (node-label label)
  (node-left-branch (tree-of leaf label))
  (node-right-branch (tree-of leaf label)))

; Ein Binärbaum ist entweder ein Blatt oder ein Knoten
(define tree-of
  (lambda (leaf label)
    (signature (mixed leaf (node-of leaf label))))))

```

Das nun ist die Definition eines Binärbaums in Reinform. Hier sind zwei Beispiele, bei denen wir einfach den Wert `#f` als Blatt verwendet haben:

```

(: tree1 (tree-of false number))
(define tree1 (make-node 3 (make-node 4 #f (make-node 7 #f #f)) #f))
(: tree2 (tree-of false number))
(define tree2 (make-node 17 (make-node 3 #f tree1) #f))

```

Hier ist noch ein weiteres Beispiel, bei dem die Blätter Zahlen sind und die Markierungen Zeichenketten

```

(: tree3 (tree-of number string))
(define tree3 (make-node "Axl"
  (make-node "Slash" 17
    (make-node "Duff" 14 23))
  12))

```

Abbildung 12.1 stellt die drei Bäume `tree1`, `tree2` und `tree3` grafisch dar. Dort hat jeder Baum erkennbar einen «obersten» Knoten, die sogenannte *Wurzel* – ein Begriff, der im Zusammenhang mit Bäumen häufig verwendet wird. Bei dem Begriff muss immer mitgesagt werden, *wovon* die Wurzel gemeint ist. So ist `tree1` ein Teil von `tree2`. Entsprechend steckt die Wurzel von `tree1` (der Knoten mit der 3) auch in `tree2` drin, aber die Wurzel von `tree2` ist eben der Knoten mit der 17.

Mit der Wurzel hängt auch der Begriff *Pfad* zusammen: Der Pfad eines Knotens oder Blattes ist die Liste der Knoten von der Wurzel zu diesem Knoten oder Blatt. Der Pfad der 14 in `tree3` in Abbildung 12.1 besteht zum Beispiel aus den Knoten Axl, Slash und Duff.

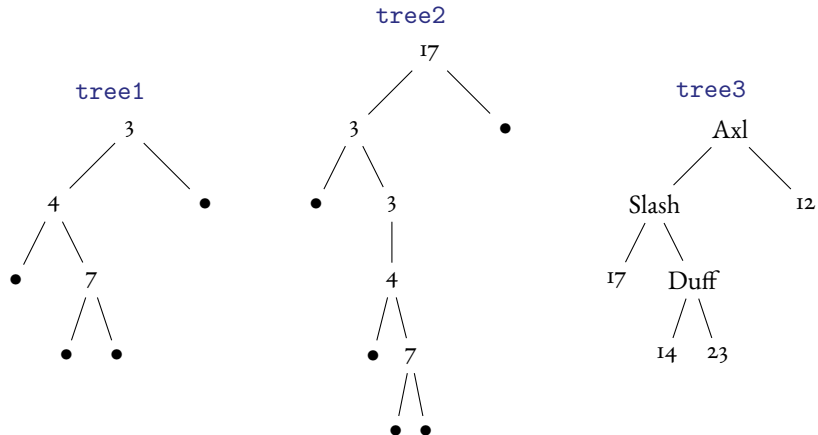


Abbildung 12.1: Beispielbäume

Wir können `tree-of` benutzen, um die Definitionen von `river` und `confluence` zu vereinfachen:

```
(define river (tree-of creek string))
(define river? node?)
(define make-confluence make-node)
(define confluence-location node-label)
(define confluence-main-stem node-left-branch)
(define confluence-tributary node-right-branch)
```

## AUFGABE 12.2

Definiere `person` mit Hilfe von `tree-of`!

□

Auf Bäumen kann man alle möglichen Sachen berechnen. Ein Beispiel ist die *Tiefe*, also die maximale Anzahl Knoten auf dem Weg zu einem Blatt. (Manchmal heißt diese Größe auch die *Höhe* eines Baums.) Für die Tiefe des Baums sind die Signaturen der Blätter und Markierungen egal:

```
; Tiefe eines Baums berechnen
(: depth ((tree-of %leaf %label) -> natural))
```

Hier sind zwei Testfälle:

```
(check-expect (depth tree1) 3)
(check-expect (depth tree2) 5)
```

Bei `tree1` sind es die Knoten mit den Markierungen 3, 4 und 7, die den maximal langen Weg zu einem Blatt bilden. Bei `tree2` sind es 17, 3, 3, 4, 7.

Es geht wieder los mit der Konstruktionsanleitung. Wir brauchen die Schablone für gemischte Daten als Eingabe. Da die Datendefinition für Binärbäume zwei Fälle hat, brauchen wir ein `cond` mit zwei Zweigen. Beim ersten können wir Bedingungen mit `node?` bilden. Die Blätter haben kein festes Prädikat, aber das sind einfach alle Bäume, die keine Knoten sind – wir können also statt einer Bedingung `else` schreiben:

```
(define depth
  (lambda (tree)
    (cond
      ((node? tree) ...)
      (else ... 0))))
```

In den Knoten stecken zwei Selbstbezüge, wir brauchen also zwei rekursive Aufrufe:

```
(define depth
  (lambda (tree)
    (cond
      ((node? tree)
       ...
       (depth (node-left-branch tree))
       (depth (node-right-branch tree))
       ...)
      (else ...))))
```

Für die Tiefe zählt nur der Weg mit der maximalen Anzahl von Knoten. Außerdem müssen wir den Knoten in `tree` noch mitzählen. Blätter zählen überhaupt nicht:

```
(define depth
  (lambda (tree)
    (cond
      ((node? tree)
       (+ 1
          (max (depth (node-left-branch tree))
                (depth (node-right-branch tree)))))
      (else 0))))
```

Fertig!

## AUFGABE 12.3

Schreibe eine Funktion, die alle Knoten eines Baums zählt!



## AUFGABE 12.4

Schreibe eine Funktion, die für einen Baum eine Liste aller Blätter des Baums liefert!



## 12.3 BÄUME FÜR'S SUCHEN

Viele Probleme bei der Programmierung sind «Suchprobleme»: Einen Namen, eine Telefonnummer, eine Bestellnummer aus einer Liste heraussuchen. Darum geht es in diesem Abschnitt und wir fangen damit an, dass wir das Wort «Liste» wörtlich nehmen und eine Funktion wie folgt schreiben:

```
; ist Wert Element einer Liste?
(: member? (%a (list-of %a) -> boolean))
```

Wir haben eine Signaturvariable verwendet, weil es sich bei den Listenelementen mal um Zahlen, mal um Zeichenketten, mal um etwas anderes handeln kann. Hier sind ein paar Testfälle:

```
(check-expect (member? 5 empty) #f)
(check-expect (member? 2 (list 1 2 3)) #t)
(check-expect (member? "Slash" (list "Ax1" "Slash")) #t)
(check-expect (member? "Buckethead" (list "Ax1" "Slash")) #f)
```

Hier Gerüst und Schablone für Funktionen auf Listen:

```
(define member?
  (lambda (element list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ...
       (first list)
       (member? element (rest list))
       ...))))
```

Bei der leeren Liste kann die Funktion nur `#f` zurückgeben. Bei der Cons-Liste legt die Schablone nahe, dass die Funktion erst einmal prüfen sollte, ob `(first list)` das gesuchte Element ist.

Klingt einfach, oder? Aber *wie* prüfen wir das? Wir könnten das hier hinschreiben:

```
(= element (first list))
```

... aber das würde die Funktion auf Zahlen beschränken, weil `=` nur auf Zahlen funktioniert. Für die beiden Testfälle mit Zeichenketten müssten wir `string=?` verwenden. Wir müssen also `member ?` über `=` respektive `string=?` abstrahieren, noch bevor die Funktion überhaupt fertig ist. Wir brauchen wie immer einen weiteren Parameter und nennen ihn `equals?`:

```
(define member?
  (lambda (equals? element list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ...
       (first list)
       (member? equals? element (rest list))
       ...))))
```

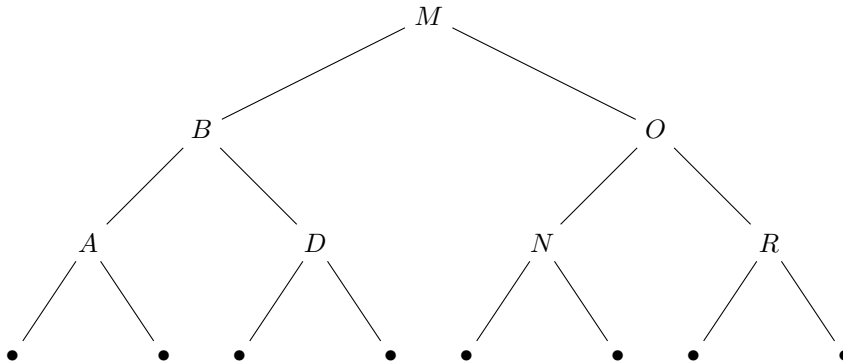
(Aufpassen: Der rekursive Aufruf muss – wie immer – auch durch den neuen Parameter erweitert werden.)

Jetzt können wir den Vergleich mit Hilfe von `equals?` durchführen und diesen mit einer binären Verzweigung verarbeiten:

```
(define member?
  (lambda (equals? element list)
    (cond
      ((empty? list) #f)
      ((cons? list)
       (if (equals? element (first list))
           #t
           (member? equals? element (rest list))))))
```

Signatur und Testfälle haben von dem neuen Parameter noch nichts mitbekommen. Die `equals?`-Funktion akzeptiert zwei Listenelemente und liefert ein boolesches Ergebnis. Da die Listenelemente die Signatur `%a` haben, sieht die Signaturdeklaration für `member?` so aus:

```
(: member? ((%a %a -> boolean) %a (list-of %a) -> boolean))
```



**Abbildung 12.2:** Sortierter Baum über Buchstaben

Bei den Testfällen müssen wir jeweils noch die richtige Vergleichsfunktion übergeben. Das ist `=` für Zahlen und `equals?` für Zeichenketten.

```
(check-expect (member? = 5 empty) #f)
(check-expect (member? = 2 (list 1 2 3)) #t)
(check-expect (member? string=? "Slash" (list "Ax1" "Slash")) #t)
(check-expect (member? string=? "Buckethead" (list "Ax1" "Slash")) #f)
```

Fertig!

Allerdings hat `member?` einen Nachteil: Bei kurzen Listen oder wenn das gesuchte Element am Anfang der Liste steht, wird `member?` ziemlich schnell fertig. Aber stell Dir vor, die Liste hat ein paar Millionen Elemente und das gesuchte Element ist am Ende. Oder gar nicht drin: Dann muss `member?` die gesamte Liste abklappern.

#### AUFGABE 12.5

Schreibe mit Hilfe von `member?` eine Funktion, die von zwei Listen alle Elemente liefert, die in beiden Listen stehen, also deren Schnittmenge. Wie lange braucht diese Funktion im ungünstigsten Fall? □

Kann irgendwie schneller herausfinden, ob ein Wert Element einer Menge ist oder nicht? In der Tat ist das möglich, aber nicht mit Listen: Wir brauchen eine andere Struktur, um das Suchen zu beschleunigen – Bäume.

Schau Dir mal Abbildung 12.2 an. In diesem Baum musst Du nicht alle Elemente anschauen, um ein bestimmtes Element zu finden. Das liegt daran, dass die Buchstaben in dem Baum auf bestimmte Art nach dem Alphabet sortiert sind:

Die Wurzel mit der Markierung  $M$  hat zwei Teilbäume – die Markierungen des linken Teilbaums liegen allesamt *vor*  $M$  im Alphabet, alle Markierungen des rechten Teilbaums *nach*  $M$ . Wenn Du also nach einem Buchstaben suchst – nehmen wir mal  $D$  – dann weißt Du, wenn Du die Wurzel mit  $M$  siehst, dass  $D$  im linken Teilbaum von  $M$  – mit der Markierung  $B$  und von da aus im rechten Teilbaum von  $B$  stehen muss. Die Knoten  $A$ ,  $O$ ,  $N$ ,  $R$  kannst Du ignorieren.

Die Suche braucht also höchstens so viele Schritte wie der Baum tief ist. Das ist schonmal besser, als in der Liste zu suchen, wo wir potenziell alle Elemente anschauen müssen.

Programmieren wir das also!

Wir fangen mit einem sortierten Baum über reellen Zahlen an. (Reelle Zahlen deshalb, weil wir sie einfach mit  $=$  und  $<$  vergleichen können. Wir verallgemeinern das später.) Die Zahlen kleben als Markierungen an den Knoten. An den Blättern steht nichts relevantes, wir benutzen deshalb immer `#f`. Entsprechend sehen Kurzbeschreibung und Signatur so aus:

```
; Ist eine Zahl in einem sortierten Baum vorhanden?
(: tree-member? (real (tree-of false real) -> boolean))
```

Die Signatur `false` ist neu: Sie beschreibt nur den Wert `#f`. Entsprechend gibt es natürlich auch eine Signatur `true` für `#t`.

Als Beispielbaum für die Tests benutzen wir diesen hier:

```
(define tree4
  (make-node 5
    (make-node 3 #f #f)
    (make-node 17
      (make-node 10 #f (make-node 12 #f #f))
      #f)))
```

Hier sind ein paar Tests:

```
(check-expect (tree-member? 5 tree4) #t)
(check-expect (tree-member? 17 tree4) #t)
(check-expect (tree-member? 3 tree4) #t)
(check-expect (tree-member? 10 tree4) #t)
(check-expect (tree-member? 2 tree4) #f)
```

Hier ist das Gerüst:

```
(define tree-member?
  (lambda (value tree)
    ...))
```



In die Schablone für Bäume tragen wir gleich den zweiten Fall ein: Wenn die Funktion ein Blatt erreicht, dann ist das `value` definitiv nicht im Baum, das Ergebnis dann `#f`:

```
(define tree-member?
  (lambda (value tree)
    (cond
      ((node? tree)
       ...
       (node-label tree)
       (tree-member? value (node-left-branch tree))
       (tree-member? value (node-right-branch tree))
       ...
      (else #f)))))
```

Bei Knoten können wir drei Fälle unterscheiden: Wenn die Markierung gerade das gesuchte `value` ist, wenn `value` kleiner ist als die Markierung (also im linken Teilbaum stehen muss) und wenn sie größer ist. Daraus ergibt sich folgende Weiterentwicklung:

```
(define tree-member?
  (lambda (value tree)
    (cond
      ((node? tree)
       ...
       (tree-member? value (node-left-branch tree))
       (tree-member? value (node-right-branch tree))
       ...
      (cond
        ((= value (node-label tree)) #t)
        ((< value (node-label tree)) ...
         (else ...)))
      (else #f)))))
```

Da `(node-label tree)` zweimal vorkommt, machen wir dafür eine Definition und setzen die Bestandteile der Schablone so zusammen:

```
(define tree-member?
  (lambda (value tree)
    (cond
      ((node? tree)
```

```

(define label (node-label tree))
(cond
  ((= value label) #t)
  ((< value label)
   (tree-member? value (node-left-branch tree)))
  (else
   (tree-member? value (node-right-branch tree))))
(else #f))))

```

Fertig!

## 12.4 SORTIERTE BÄUME HERSTELLEN

In den Testfällen für `tree-member?` haben wir immer den Baum `tree4` verwendet, den wir direkt mit `make-node` konstruiert haben. Dabei mussten wir selbst darauf achten, dass er auch sortiert ist. In diesem Abschnitt automatisieren wir diese Konstruktion, dann kann dabei auch kein Fehler passieren. (Wenn wir korrekt programmieren zumindest.)

Wir schreiben dafür eine Funktion, die ein neues Element in einen bestehenden sortierten Baum einfügt. Kurzbeschreibung und Signatur sind wie folgt:

```

; Zahl in sortierten Baum einfügen
(: tree-insert (real (tree-of false real) -> (tree-of false real)))

```

Testfälle brauchen wir als nächstes. Wir könnten das so machen wie immer: Wir schreiben einen Aufruf von `tree-member?` hin und den Ergebniswert, den wir uns erhoffen. In diesem Fall aber ist es gar nicht so wichtig, was der Ergebniswert genau ist. Wichtig ist, dass ein eingefügtes Element im Ergebnisbaum auch drin ist. Außerdem ist es mühsam, immer den ganzen Baum hinzuschreiben. Darum benutzen wir `tree-member?`, um `tree-insert` zu testen.

```

(check-expect (tree-member? 5 (tree-insert 5 tree4)) #t)
(check-expect (tree-member? 11 (tree-insert 11 tree4)) #t)

```

Später werden wir feststellen, dass `tree-insert` unterschiedliche sortierte Bäume liefern kann, die allesamt korrekt sind.

### AUFGABE 12.6

Die beiden Tests erwarten jeweils, dass `#t` bei `tree-member?` herauskommt. Wäre es sinnvoll, auch noch welche mit `#f` zu schreiben? □

Wenn Du ein mulmiges Gefühl bei den spärlichen beiden Tests hast: richtig! In Kapitel 13 auf Seite 397 werden wir zeigen, wie man Funktionen wie `tree-insert` besser testet.

Aber jetzt geht's erst einmal mit der Funktion selbst los. Gerüst und Schablone sind genau wie bei `tree-member?`:

```
(define tree-insert
  (lambda (value tree)
    (cond
      ((node? tree)
       ...
       (tree-insert value (node-left-branch tree))
       (tree-insert value (node-right-branch tree))
       ...)
      (else ...))))
```

Die Fallunterscheidung bei Knoten ist ebenfalls wie in `tree-member?`, darum können wir auch die Verzweigung aus der dortigen Schablone übernehmen:

```
(define tree-insert
  (lambda (value tree)
    (cond
      ((node? tree)
       ...
       (tree-insert value (node-left-branch tree))
       (tree-insert value (node-right-branch tree))
       ...
       (cond
        ((= value (node-label tree)) ...)
        (< value (node-label tree)) ...)
        (else ...)))
      (else ...))))
```

Auch hier ist der erste Fall einfach: Wenn `value` gerade die Markierung eines Knotens ist, dann enthält der Baum den Wert bereits, die Funktion muss nichts einfügen und kann einfach `tree` liefern. Auch für den Fall, dass `tree` ein Blatt ist (das letzte `else`), ist es recht einfach: Wir konstruieren einen neuen, einelementigen Baum:

```
(define tree-insert
  (lambda (value tree)
```

```

(cond
  ((node? tree)
    ...
    (tree-insert value (node-left-branch tree))
    (tree-insert value (node-right-branch tree))
    ...
    (cond
      ((= value (node-label tree)) tree)
      ((< value (node-label tree)) ...)
      (else ...)))
  (else (make-node value #f #f))))

```

Es bleiben noch zwei Fälle, in denen der einzufügende Wert links beziehungsweise rechts von der Knotenmarkierung liegt. Er muss entsprechend im linken oder rechten Teilbaum eingefügt werden: Genau das erledigen die beiden rekursiven Aufrufe aus der Schablone. Der jeweils andere Teilbaum bleibt so wie er ist:

```

(define tree-insert
  (lambda (value tree)
    (cond
      ((node? tree)
        (cond
          ((= value (node-label tree)) tree)
          ((< value (node-label tree))
            (make-node (node-label tree)
              (tree-insert value (node-left-branch tree))
              (node-right-branch tree)))
          (else
            (make-node (node-label tree)
              (node-left-branch tree)
              (tree-insert value (node-right-branch tree))))))
      (else
        (make-node value #f #f))))

```

Fertig!

## 12.5 SUCHBÄUME

Unsere Funktionen `tree-member?` und `tree-insert` funktionieren nur auf Zahlen. Die Bäume aus den Abbildungen 12.2 und 12.4 enthalten aber beide Buchstaben. Wenn wir andere Werte als Zahlen zulassen wollen, müssen wir wieder einmal abstrahieren über alles, was mit Zahlen zu tun hat.

Schau Dir nochmal die Definition von `tree-member?` an: Es gibt zwei Stellen, die «zahlen-spezifisch» sind, nämlich `=` und `<`. Wenn Zeichenketten in einem sortierten Baum unterbringen wollten, müssten wir da `string=?` und `string<?` hinschreiben.

### AUFGABE 12.7

Abstrahiere `tree-member?` und `tree-insert` über `=` und `<`.

Übergib mal statt `<` die Funktion `>`. Funktionieren `tree-member?` und `tree-insert` dann noch korrekt? Wie unterscheiden sich die Bäume, die mit `<` aus `tree-insert` herauskommen von denen mit `>`? □

Die abstrahierten Versionen von `tree-member?` und `tree-insert` haben einen Nachteil: Bei jedem Aufruf dieser Funktionen müssen wir die beiden Argumente für `=` und `<` hinschreiben. Das nervt und ist jedesmal eine Gelegenheit für einen Fehler, weil wir das vollkommen konsistent machen müssen.

Wir wollen also versuchen, ohne diese zusätzlichen Parameter auszukommen. Dazu benutzen wir einen Trick und packen die Funktionen für `=` und `<` zusammen mit `tree`. Das Ergebnis heißt *Suchbaum*. Heraus kommt folgende Datendefinition:

```
; Ein Suchbaum besteht aus
; - Funktion für =
; - Funktion für <
; - Binärbaum
```

Um die Definition in Code umzusetzen, benutzen wir eine Record-Definition. Diesmal abstrahieren wir über die Signatur der Elemente mit einem Signatur-Parameter namens `element`:

```
(define-record (search-tree-of element)
  make-search-tree search-tree?
  (search-tree-label=?-function (element element -> boolean))
  (search-tree-label-<?-function (element element -> boolean))
  (search-tree-tree (tree-of false element)))
```

Vergleiche die Signatur von `search-tree-tree` mit der Signatur der Bäume bei `tree-member?` und `tree-insert!`

Hier ist der Suchbaum aus Abbildung 12.2 auf Seite 365:

```
(define search-tree1
  (make-search-tree
    string=? string<?
    (make-node "M"
      (make-node "B"
        (make-node "A" #f #f)
        (make-node "D" #f #f))
      (make-node "O"
        (make-node "N" #f #f)
        (make-node "R" #f #f)))))
```

Wir schreiben nun eine Variante von `tree-member?`, die Suchbäume statt «normale» Bäume akzeptieren. Hier Kurzbeschreibung, Signatur und Tests für die Funktion für das Suchen:

```
; feststellen, ob Element in Suchbaum vorhanden ist
(: search-tree-member? (%a (search-tree-of %a) -> boolean))
(check-expect (search-tree-member? "M" search-tree1) #t)
(check-expect (search-tree-member? "D" search-tree1) #t)
(check-expect (search-tree-member? "N" search-tree1) #t)
(check-expect (search-tree-member? "R" search-tree1) #t)
(check-expect (search-tree-member? "Z" search-tree1) #f)
```

Hier ist das Gerüst für die Funktionsdefinition, zusammen mit der Schablone für `search-tree`:

```
(define search-tree-member?
  (lambda (value search-tree)
    ...
    (search-tree-label=?-function search-tree)
    (search-tree-label-<?-function search-tree)
    (search-tree-tree search-tree)
    ...))
```

Jetzt könnten wir natürlich alles wie immer ausprogrammieren. Viel einfacher aber ist es, wenn wir die existierende Funktion `tree-member?` wiederverwenden: Wir haben ja genau das gleiche vor, was die Funktion schon macht. Dazu kopieren wir sie einfach in den Rumpf:

```
(define search-tree-member?
  (lambda (value search-tree)
    ...
    (search-tree-label=?-function search-tree)
    (search-tree-label<?-function search-tree)
    (search-tree-tree search-tree)
    (define tree-member?
      (lambda (value tree)
        (cond
          ((node? tree)
           (define label (node-label tree))
           (cond
             ((= value label) #t)
             ((< value label)
              (tree-member? value (node-left-branch tree)))
             (else
              (tree-member? value (node-right-branch tree))))))
        (else #f))))
    ...))
```

Wir müssen allerdings noch über = und < abstrahieren. Dazu denken wir uns erstmal nur neue Namen aus, nämlich `label=?` für = und `label<?` für <:

```
(define search-tree-member?
  (lambda (value search-tree)
    ...
    (search-tree-label=?-function search-tree)
    (search-tree-label<?-function search-tree)
    (search-tree-tree search-tree)
    (define tree-member?
      (lambda (value tree)
        (cond
          ((node? tree)
           (define label (node-label tree))
           (cond
             ((label=? value label) #t)
             ((label<? value label)
```

```

        (tree-member? value (node-left-branch tree)))
      (else
       (tree-member? value (node-right-branch tree))))
    (else #f))))
  ...))

```

Anders als sonst legen wir aber keine Parameter für `label=?` und `label<?` an: Die Funktionen dafür stehen ja schon in der Schablone, wir müssen ihnen nur die richtigen Namen geben mit Hilfe von lokalen Definitionen:

```

(define search-tree-member?
  (lambda (value search-tree)
    (define label=? (search-tree-label==?-function search-tree))
    (define label<? (search-tree-label-<?-function search-tree))
    ...
    (search-tree-tree search-tree)
    ...
    (define tree-member? ...)
    ...)))

```

Zu guter letzt brauchen wir noch einen Aufruf von `tree-member?` im Rumpf von `search-tree-member?`, damit es auch losgeht. Dafür verbrauchen wir den letzten Baustein aus der Schablone:

```

(define search-tree-member?
  (lambda (value search-tree)
    (define label=? (search-tree-label==?-function search-tree))
    (define label<? (search-tree-label-<?-function search-tree))
    (define tree-member?
      (lambda (value tree)
        (cond
          ((node? tree)
           (define label (node-label tree))
           (cond
             ((label=? value label) #t)
             ((label<? value label)
              (tree-member? value (node-left-branch tree)))
             (else

```



```

      (tree-member? value (node-right-branch tree))))))
    (else #f))))
(tree-member? value (search-tree-tree search-tree))))

```

Fertig!

#### AUFGABE 12.8

Schreibe entsprechend zu `search-tree-member?` eine Funktion `search-tree-insert` auf Basis von `tree-insert`! □

## 12.6 SORTIERTE BÄUME SIND EFFIZIENTER ALS LISTEN

Intuitiv ist die Idee mit dem sortierten Baum beim Suchen effizienter. Aber hält diese Intuition auch einer mathematischen Untersuchung stand? Betrachte dazu den Baum in «Ebenen» – erste Ebene ist die Wurzel, die zweite Ebene deren Teilbäume, die dritte Ebene wiederum deren Teilbäume undso weiter. Dann passen in jede Ebene doppelt soviele Knoten wie in die Ebene darüber.

In einen Baum der Tiefe 1 passt auch nur  $1 = 2^0$  Knoten, in einen der Tiefe 2 passen  $2^0 + 2^1 = 1 + 2 = 3$  Knoten, dann 7, dann 15 undso weiter. Dir fällt vielleicht auf, dass die Zahlen 1, 3, 7, 15 allesamt Vorgänger einer Zweierpotenz sind. Wir können deshalb versuchen, das zu einer Formel zu verallgemeinern. Die Tiefe des Baums heißt dabei  $t$ . Dann nehmen wir an (oder hoffen zumindest), dass für für alle

$$2^0 + \dots + 2^{t-1} = \sum_{i=0}^{i=t-1} 2^i = 2^t - 1$$

Da es sich bei  $t$  um eine natürliche Zahl handelt, können wir vollständige Induktion anwenden nach der Anleitung in Abschnitt 8.2 auf Seite 217.

Wir müssen also beweisen, dass für alle  $t \in \mathbb{N}$  gilt:

$$\sum_{i=0}^{i=t-1} 2^i = 2^t - 1$$

$t = 0$ :

$$\sum_{i=0}^{i=0-1} 2^i = 2^0 - 1$$

Beweis:

$$\begin{aligned}
 \sum_{i=0}^{i=-1} 2^i &= 0 \\
 &= 1 - 1 \\
 &= 2^0 - 1
 \end{aligned}$$

(Vielleicht kommt Dir komisch vor, dass die Summe von 0 bis  $-1$  läuft. Da es keine Zahl zwischen 0 bis  $-1$  gibt, ist die Summe leer und deshalb das neutrale Element bezüglich der Addition.)

Induktionsvoraussetzung:

$$\sum_{i=0}^{i=t-1} 2^i = 2^t - 1$$

Induktionsschluss (zu zeigen):

$$\sum_{i=0}^{i=(t+1)-1} 2^i = 2^{t+1} - 1$$

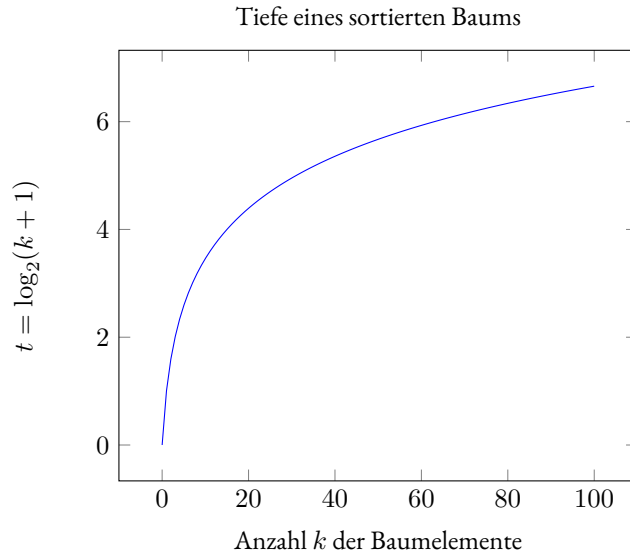
Beweis:

$$\begin{aligned}
 \sum_{i=0}^{i=(t+1)-1} 2^i &= \sum_{i=0}^{i=t} 2^i \\
 &= \sum_{i=0}^{i=t-1} 2^i + 2^t \\
 &= 2^t - 1 + 2^t \quad \text{Induktionsvoraussetzung} \\
 &= 2^t \times 2 - 1 \\
 &= 2^{t+1} - 1
 \end{aligned}$$

Wozu ist diese Formel gut, fragst Du Dich vielleicht. Nun, die rechte Seite können wir umdrehen. (Bei der linken Seite ist das schwieriger.) Wenn die Anzahl der Knoten  $k$  ist, dann gilt:

$$\begin{aligned}
 k &= 2^t - 1 \\
 \iff k + 1 &= 2^t \\
 \iff \log_2(k + 1) &= t
 \end{aligned}$$

Die  $\log_2$  ist der sogenannte *Logarithmus* zur Basis 2, auch genannte *Zweierlogarithmus*. Das ist die Umkehrfunktion zur Exponentialfunktion mit der Basis 2.



**Abbildung 12.3:** Beziehung zwischen Anzahl von Knoten und Tiefe eines Binärbaums

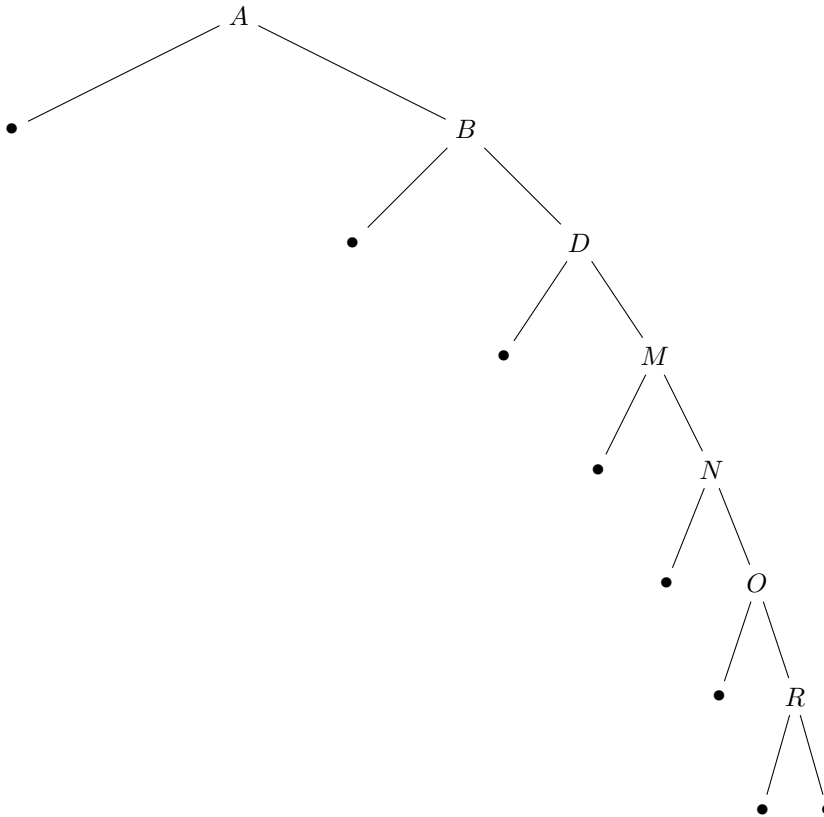
Schau Dir Abbildung 12.3 an: Da siehst Du, dass die Tiefe – mithin der Logarithmus – viel langsamer wächst als die Anzahl der Knoten, und die Kurve immer flacher wird. Diese Kurve erklärt, warum das mit dem sortierten Baum eine gute Idee ist: Die Tiefe des Baums ist ja die Anzahl der Elemente des Baums, die man abklappen muss, um das gewünschte Element zu finden. (Beziehungsweise herauszubekommen, dass es nicht im Suchbaum ist.) Das gilt allerdings nur, wenn der Suchbaum «voll besetzt» ist. Wir müssen uns also irgendwann Gedanken machen, wie wir dafür sorgen, dass Suchbäume immer möglichst voll besetzt sind. Dazu kommen wir später.

## 12.7 SUCHBÄUME BALANCIEREN

Abbildung 12.4 zeigt einen Suchbaum, der nicht voll besetzt ist – einen sogenannten *entarteten Suchbaum*. Bei diesem Suchbaum dauert die Suche genau so lang wie in einer Liste.

### AUFGABE 12.9

Schreibe einen Ausdruck aus Aufrufen von `search-tree-insert`, um den Suchbaum in Abbildung 12.4 zu erzeugen! □



**Abbildung 12.4:** Entarteter Suchbaum

In diesem Abschnitt wollen wir eine Variante von `search-tree-insert` schreiben, bei der niemals so ein entarteter Suchbaum herauskommen kann und die stattdessen den Baum immer *balanciert*.

Spätestens jetzt betreten wir einen Bereich der Programmierung, der in der Informatik meist *Algorithmen und Datenstrukturen* heißt [BIRD and GIBBONS, 2020]. (Manchmal auch nur *Algorithmen* oder nur *Datenstrukturen*.) Ein Algorithmus ist in der Regel eine Funktion, die in vielen Programmen verwendet werden können. Entsprechend ist eine Datenstruktur eine Datendefinition, auf der ein Algorithmus aufbaut: Die beiden gehören oft zusammen.<sup>1</sup>

<sup>1</sup> Der Begriff «Algorithmus» stammt vom Namen des persischen Rechenmeisters *Al-Chwarizmi*. Die deutsche Schul informatik kennt noch detaillierte Definitionen des Begriffs, die aber für die praktische Programmierung weitgehend irrelevant sind.

Viele Algorithmen basieren auf cleveren Ideen, von denen die Schöpferinnen und Schöpfer selbst gar nicht so genau erklären können, wie sie auf sie gekommen sind. Dementsprechend erwarten wir nicht von Dir (und auch nicht von uns), auf das Material dieses Abschnitts selbst zu kommen. Glücklicherweise gibt es in vielen Bereichen der Programmierung schon fertige Algorithmen und Datenstrukturen, die Du nur benutzen musst und nicht selbst programmieren. In diesem Abschnitt zeigen wir aber trotzdem beispielhaft, wie und warum so eine Datenstruktur funktioniert. Vielleicht findest Du es interessant – andernfalls reicht es aber, wenn Du diesen Abschnitt überfliegst oder überspringst.

Wie also können wir einen Suchbaum balancieren? Dabei sollten wir nicht nur dafür sorgen, dass das Suchen effizient funktioniert. Auch das Balancieren sollte nicht zuviel Arbeit machen – sonst werden die Vorteile beim Suchen durch den Aufwand beim Balancieren wieder zunichte gemacht. Das bedeutet ganz praktisch, dass wir einen Kompromiss suchen müssen zwischen den Einsparungen beim Suchen und dem Aufwand für das Balancieren. Solche Kompromisse gibt es oft bei Algorithmen, und oft werden Messungen eingesetzt, um sie auszuwählen und schrittweise zu verbessern. So ist es auch hier.

Wir zeigen Dir hier eine Balancier-Funktion, die in einem wissenschaftlichen Aufsatz von Stephen Adams [ADAMS, 1993] aus dem Jahr 1993 beschrieben ist. Dieser beruht auf einem Programmierwettbewerb, bei dem Adams den zweiten Platz belegte, was die Effizienz betrifft. Gleichzeitig ist seine Lösung aber deutlich eleganter als die erstplatzierte, weswegen der Aufsatz bis heute oft eingesetzt wird, um effiziente Suchbäume zu programmieren.

Die Idee von Adams beruht auf einer Variante von `search-tree-insert`, die bei der Konstruktion eines neuen Knotens darauf achtet, dass der Baum nicht zu sehr aus dem Gleichgewicht gerät. Falls dies doch droht zu passieren, wird der Baum *rotiert* und so balanciert.

### 12.7.1 GRÖSSENANNOTIERTE BÄUME

Um festzustellen, ob der Baum aus dem Gleichgewicht gerät, vergleicht der Algorithmus die Anzahl der Knoten in den beiden Teilbäumen eines Knotens. Nun wäre es blöd, wenn er dafür jedesmal nachzählen müsste, wieviele das denn sind. Viel effizienter ist es, die Größe eines Baums im Baum selbst mitzuführen, als Teil der Markierung eines Knotens. Die definieren wir erstmal, bevor es mit dem Ausbalancieren losgeht.

Bäume, bei denen die Größe an jedem Knoten dransteht, nennen wir *größenannotiert*. Im Englischen ist der Begriff nicht ganz so sperrig: *sized*. Neben der Größe brauchen wir an jedem Knoten auch noch die «eigentliche» Markierung. Zusammen ergibt das folgende Daten- und Record-Definition:

```

; Die Markierung an einem größenannotierten Baum besteht aus:
; - Anzahl der Knoten
; - "eigentliche" Markierung
(define-record (sized-label-of label)
  make-sized-label
  sized-label
  (sized-label-size natural)
  (sized-label-label label))

```

Hier ist nochmal der Baum aus Abbildung 12.2 auf Seite 365, aber diesmal größenannotiert:

```

(define sized-tree1
  (make-node (make-sized-label 7 "M")
    (make-node (make-sized-label 3 "B")
      (make-node (make-sized-label 1 "A") #f #f)
      (make-node (make-sized-label 1 "D") #f #f))
    (make-node (make-sized-label 3 "Q")
      (make-node (make-sized-label 1 "N") #f #f)
      (make-node (make-sized-label 1 "R") #f #f))))

```

Ganz schön umständlich, jedesmal von Hand die richtige Größe hinzuschreiben. Das können wir auch gut den Computer erledigen lassen, indem wir eine Hilfsfunktion schreiben, welche die Größe bei der Konstruktion eines Knotens ausrechnet. Hier sind Kurzbeschreibung und Signatur:

```

; Größenannotierten Knoten konstruieren
(: make-sized-node (%label (tree-of %leaf (sized-label-of %label))
  (tree-of %leaf (sized-label-of %label))
  -> (node-of %leaf (sized-label-of %label))))

```

Diese Signatur entspricht der von `make-node` – hier ist diese zum Vergleich:

```

(: make-node (%label (tree-of %leaf %label)
  (tree-of %leaf %label)
  -> (node-of %leaf %label)))

```

Für einen Test nehmen wir den Ausdruck von `sized-tree1` ersetzen die Aufrufe von `make-node` durch welche von `make-sized-node`. Es sollte das gleiche herauskommen:

```

(check-expect (make-sized-node
  "M"

```

```

(make-sized-node "B"
  (make-sized-node "A" #f #f)
  (make-sized-node "D" #f #f))
(make-sized-node "O"
  (make-sized-node "N" #f #f)
  (make-sized-node "R" #f #f)))
sized-tree1)

```

Hier ist Gerüst mit Schablone für die Funktion. Die Signatur besagt, dass die Funktion einen `node`-Record liefert, entsprechend besteht die Schablone aus einem Aufruf des Konstruktors:

```

(define make-sized-node
  (lambda (label left-branch right-branch)
    (make-node ... ... ...)))

```

Die beiden Argumente für `left-branch` und `right-branch` können wir direkt an den Konstruktor weiterreichen. Allein die Größe müssen wir noch ausrechnen: Diese ist die Summe der beiden Teilbäume, plus 1, weil das `make-node` noch einen neuen Knoten erzeugt:

```

(define make-sized-node
  (lambda (label left-branch right-branch)
    (make-node
      (make-sized-label (+ 1
                          (sized-tree-size left-branch)
                          (sized-tree-size right-branch))
                        label)
      left-branch right-branch)))

```

Schau nochmal den Teilausdruck an, der die Größe berechnet: Da wird eine Funktion `sized-tree-size` benutzt, welche die Größe eines Baums berechnet. Die gib's aber noch gar nicht – reines Wunschdenken! Das müssen wir noch nachholen. Hier sind Kurzbeschreibung, Signatur und zwei Tests:

```

; Größe eines größenannotierten Baums liefern
(: sized-tree-size ((tree-of %leaf (sized-label-of %label)) ->
  Natural))

(check-expect (sized-tree-size "A") 0)
(check-expect (sized-tree-size sized-tree1) 7)

```

Hier sind Gerüst und Schablone. Die Funktion verarbeitet gemischte Daten, also brauchen wir eine Verzweigung:

```
(define sized-tree-size
  (lambda (tree)
    (cond
      ((node? tree) ...)
      (else ...))))
```

Im `node`-Fall wissen wir schon, wo die Größe steht, nämlich in der `sized-label`-Markierung. Im anderen Fall handelt es sich um ein Blatt, das hat Größe 0:

```
(define sized-tree-size
  (lambda (tree)
    (cond
      ((node? tree)
       (sized-label-size (node-label tree)))
      (else 0))))
```

Wenn Du Dir die Signaturdeklarationen von `make-sized-node` und `sized-tree-size` anschaut, dann siehst Du, dass folgende Signaturen mehrfach vorkommen:

```
(node-of %leaf (sized-label-of %label))
(tree-of %leaf (sized-label-of %label))
```

Das wird noch öfter passieren, darum abstrahieren wir:

```
; Signatur für größenannotierte Knoten
(define sized-node-of
  (lambda (leaf label)
    (node-of leaf (sized-label-of label))))

; Signatur für größenannotierte Bäume
(define sized-tree-of
  (lambda (leaf label)
    (tree-of leaf (sized-label-of label))))
```

#### AUFGABE 12.10

Schreibe bessere Signaturdeklarationen für `make-sized-node` und `sized-tree-size` mit Hilfe dieser Funktionen! □



Wir können außerdem schon absehen, dass wir noch häufig die Markierung eines größenannotierten Knotens benötigen werden. Dafür müssen wir im Moment zwei Funktionen aufrufen:

```
(sized-node-label (node-label node))
```

Wir definieren deshalb eine Hilfsfunktion:

```
; aus größenannotiertem Knoten die Markierung extrahieren
(: sized-node-label ((sized-node-of %leaf %label) -> %label))
```

```
(define sized-node-label
  (lambda (node)
    (sized-label-label (node-label node))))
```

## 12.7.2 BALANCIERTE SUCHBÄUME

Für die Entwicklung von Algorithmen gibt es leider keine übergreifende Strategie analog zu den Konstruktionsanleitungen. Häufig steht am Anfang eine Idee, die sich aus einer bestimmten Sichtweise des Problems entsteht, verbunden mit Variationen und Experimenten.

Der Ausgangspunkt von Stephen Adams' Algorithmus ist die Funktion `search-tree-insert`. Die platziert das neue Element an die erstbeste Stelle im Suchbaum, die passt: Das ist der minimal mögliche Aufwand. Adams' Idee ist nun, `search-tree-insert` im wesentlichen beizubehalten, aber mit einer Ergänzung: Wenn das Einfügen eines neuen Elements den Baum aus dem Gleichgewicht bringt, wird er ein bisschen neu ausbalanciert.

So, als würde man immer, wenn man was neues kauft, den neuen Gegenstand im Zimmer ins Regal legen und dann ein kleines bisschen aufräumen. Die Kunst ist, gerade genug aufzuräumen, dass das Zimmer auf Dauer trotzdem nicht im Chaos versinkt. Dabei hilft wiederum die Annotation mit der Größe.

Wir brauchen also erstmal eine Variante der Datendefinition für Suchbäume mit größenannotierten Bäumen:

```
; Ein größenannotierter Suchbaum besteht aus
; - Funktion für =
; - Funktion für <
; - größenannotierter Binärbaum
(define-record (sized-search-tree-of element)
  make-sized-search-tree sized-search-tree?)
```

```
(sized-search-tree-label-=?-function (element element -> boolean))
(sized-search-tree-label-<?-function (element element -> boolean))
(sized-search-tree-tree (sized-tree-of false element)))
```

Hier ist ein Beispiel:

```
(define sized-search-tree1
  (make-sized-search-tree
    string=? string<?
    (make-sized-node "M"
      (make-sized-node "B"
        (make-sized-node "A" #f #f)
        (make-sized-node "D" #f #f))
      (make-sized-node "O"
        (make-sized-node "N" #f #f)
        (make-sized-node "R" #f #f)))))
```

#### AUFGABE 12.11

Warum müssen wir einen neuen Record-Typ definieren? Könnten wir `sized-search-tree-of` nicht folgendermaßen definieren?

```
(define sized-search-tree-of
  (lambda (element)
    (search-tree-of (sized-label-of element))))
```

□

Wir machen als nächstes aus der Funktion `search-tree-member?` die Funktion `sized-search-tree-member?`. Dafür müssen wir die Funktion nur kopieren, umbenennen und auf die Benutzung von `sized-search-tree` umstellen.

```
; feststellen, ob Element in Suchbaum vorhanden ist
(: sized-search-tree-member? (%a (sized-search-tree-of %a) ->
  boolean))
(check-expect (sized-search-tree-member? "M" sized-search-tree1) #t)
(check-expect (sized-search-tree-member? "D" sized-search-tree1) #t)
(check-expect (sized-search-tree-member? "N" sized-search-tree1) #t)
(check-expect (sized-search-tree-member? "R" sized-search-tree1) #t)
```

```
(check-expect (sized-search-tree-member? "Z" sized-search-tree1) #f)

(define sized-search-tree-member?
  (lambda (value search-tree)
    (define label=? (sized-search-tree-label=?-function search-tree))
    (define label<? (sized-search-tree-label-<?-function search-tree))
    (define tree-member?
      (lambda (value tree)
        (cond
          ((node? tree)
           (define label (sized-node-label tree))
           (cond
             ((label=? value label) #t)
             ((label<? value label)
              (tree-member? value (node-left-branch tree)))
             (else
              (tree-member? value (node-right-branch tree))))))
          (else #f))))
    (tree-member? value (sized-search-tree-tree search-tree))))
```

Vergleiche sie mit `search-tree-insert`: Es gibt nur zwei Unterschiede:

- Die Funktion benutzt größenannotierte Suchbäume und Bäume anstatt der «normalen».
- Die Funktion ruft statt `node-label` `sized-node-label` auf.

Nun widmen wir uns der Konstruktion von balancierten Suchbäumen und schreiben eine Funktion `balanced-search-tree-insert`. Auch sie entsteht aus `search-tree-insert` durch Kopieren und Umstellen auf `sized-search-tree`:

```
; neues Element in größenannotierten Suchbaum einfügen
(: balanced-search-tree-insert
  (%a (sized-search-tree-of %a) -> (sized-search-tree-of %a)))

(define balanced-search-tree-insert
  (lambda (value search-tree)
    (define label=? (sized-search-tree-label=?-function search-tree))
    (define label<? (sized-search-tree-label-<?-function search-tree))
    (define tree-insert
      (lambda (value tree)
```

```

(cond
  ((node? tree)
    (cond
      ((label=? value (sized-node-label tree))
        tree)
      ((label<? value (sized-node-label tree))
        (make-balanced-node
          (sized-node-label tree)
          (tree-insert value (node-left-branch tree))
          (node-right-branch tree)))
      (else
        (make-balanced-node
          (sized-node-label tree)
          (node-left-branch tree)
          (tree-insert value (node-right-branch tree))))))
    (else
      (make-sized-node value #f #f))))
(make-sized-search-tree
 label=? label<?
 (tree-insert value (sized-search-tree-tree search-tree))))

```

Für die Konstruktion eines neuen Knotens benutzt die Funktion nicht `make-node` (beziehungsweise `make-sized-node`), sondern eine noch zu schreibende Funktion `make-balanced-node`. Die muss sich dann um das oben beschriebene «bisschen Aufräumen» durch Ausbalancieren kümmern. Sie hat die gleiche Signatur wie `make-sized-node`:

```

; neuen Knoten herstellen, dabei neu ausbalancieren
(: make-balanced-node (%label (sized-tree-of false %label)
                           (sized-tree-of false %label)
                           -> (sized-node-of false %label)))

```

Hier das Gerüst:

```

(define make-balanced-node
  (lambda (label left-branch right-branch)
    ...))

```

Wir könnten es uns einfach machen und schreiben:

```
(define make-balanced-node
  (lambda (label left-branch right-branch)
    (make-sized-node label left-branch right-branch)))
```

Das entspricht aber der Strategie eines gewissen 12jährigen Kindes eines der Autoren: Einfach das neue Element aufs Bett werfen, Tür zu, fertig. Aber die Idee ist ja: Wenn die Unordnung zu groß ist, wollen wir ein bisschen aufräumen. Wobei Unordnung hier heißt «Ungleichgewicht zwischen `left-branch` und `right-branch`». Und das Ungleichgewicht definieren wir (genauer gesagt: Stephen Adams) über die Größe der Teilbäume.

Naiv könnten wir es folgendermaßen machen:

```
(define make-balanced-node
  (lambda (label left-branch right-branch)
    (cond
      ((= (sized-tree-size left-branch)
          (sized-tree-size right-branch))
       (make-sized-node label left-branch right-branch))
      (else ...))) ; ausbalancieren
```

... aber das wäre zu hart: Das würde heißen, dass die Funktion *jedesmal* ausbalancieren müsste, wenn auch nur das kleinste bisschen Ungleichgewicht entstehen würde. Abgesehen davon ist das auch unmöglich: Ein Baum der Größe 3 kann gar nicht perfekt balanciert sein zum Beispiel.

Wir brauchen also ein «weicheres» Kriterium. Stephen Adams hatte nun folgende Idee: Wenn ein Teilbaum mehr als  $n$  mal so groß ist wie der andere, dann wird neu ausbalanciert. Wie ist er darauf gekommen und wie groß sollte  $n$  sein? Wie er darauf gekommen ist, ist nicht überliefert, aber wir gehen davon aus, dass er unterschiedliche Kriterien ausprobiert hat und dieses hier dazu führte, das die entstehende Funktion sowohl einfach als auch im Schnitt schnell funktioniert. Dabei hat er auch ermittelt, wie groß  $n$  sinnvollerweise ist, nämlich 5. Es entsteht folgende Schablone:

```
(define ratio 5)

(define make-balanced-node
  (lambda (label left-branch right-branch)
    (define left-size (sized-tree-size left-branch))
    (define right-size (sized-tree-size right-branch))
    (cond
      ((> right-size (* ratio left-size)) ...) ; rechts hat Ü
      bergewicht
```

```

(> left-size (* ratio right-size)) ...) ; links hat Übergewicht
(else
  (make-sized-node label left-branch right-branch))))

```

Du siehst, wir haben die `5` in eine separate Definition verlagert, weil sie zweimal benutzt wird und so einfach geändert werden kann. Adams hat die Funktion noch etwas weiter verbessert, indem er ein weiteres Kriterium hinzufügte, bei dem kein Ausbalancieren notwendig ist:

```

(define make-balanced-node
  (lambda (label left-branch right-branch)
    (define left-size (sized-tree-size left-branch))
    (define right-size (sized-tree-size right-branch))
    (cond
      ((< (+ left-size right-size) 2)
       (make-sized-node label left-branch right-branch))
      (> right-size (* ratio left-size)) ...) ; rechts hat Ü
bergewicht
      (> left-size (* ratio right-size)) ...) ; links hat Übergewicht
      (else
       (make-sized-node label left-branch right-branch))))

```

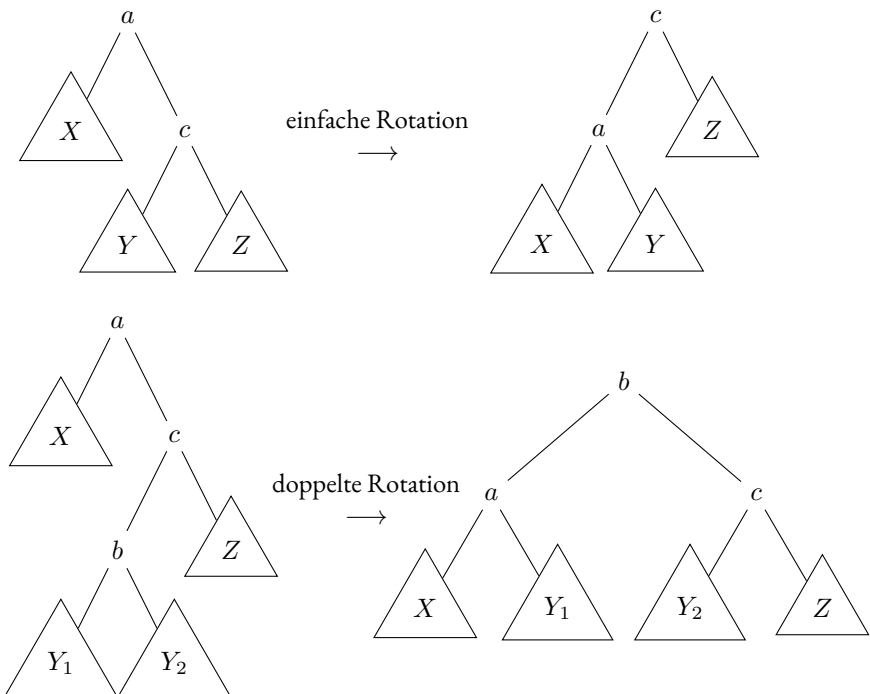
Der erste Zweig kommt immer dann zum Zug, wenn in den beiden Teilbäumen zusammen nur ein Knoten ist: Dann macht Ausbalancieren auch keinen Sinn.

## 12.8 AUSBALANCIEREN DURCH ROTATION

Um den Baum auszubalancieren, schauen wir uns an, wie so ein Ungleichgewicht überhaupt aussehen kann. Schau Dir Abbildung 12.5 an. Links oben siehst Du einen Baum mit Markierungen *a* und *c* sowie dreieckigen Teilbäumen *X*, *Y* und *Z*. Nun stell Dir vor, der Baum hat rechts deutlich mehr Knoten als links. Dann ist er aus dem Gleichgewicht und sollte neu ausbalanciert werden.

Das erledigt eine *Rotation*, die aus diesem Baum den Baum oben rechts macht: Dabei wird der Teilbaum *Y* von rechts nach links geschoben, was das Ungleichgewicht zumindest etwas lindern kann.

Unten links siehst Du ebenfalls einen Baum, der aus dem Gleichgewicht gekommen ist: Auch hier wird ein Teilbaum von rechts nach links verschoben, der sich aber noch eine Ebene tiefer befin-



**Abbildung 12.5:** Ausbalancieren von Bäumen durch Rotation

det. Diese Rotation ist etwas komplizierter, darum heißt sie «doppelte» Rotation, die erste heißt entsprechend «einfach».

#### AUFGABE 12.12

Warum bewirken die Rotationen in Abbildung 12.5 nicht, dass die Bäume auf der rechten Seite nicht mehr sortiert sind? □

Wann sollte der Algorithmus nun einfach und wann doppelt rotieren? Schau Dir dazu noch einmal in Abbildung 12.5 den oberen linken Baum an. Von den Teilbäumen  $X$  und  $Y$  ist einer größer als der andere. Nehmen wir mal an, das ist  $X$ . Dann vermindert eine einfache Rotation das Ungleichgewicht zwischen links und rechts.

Wenn aber  $Y$  größer ist  $Z$ , dann macht die einfache Rotation das Ungleichgewicht *größer*: Der eh größere Brocken  $Y$  tut sich dann auch noch mit  $X$  zusammen, während der kleine Brocken  $Z$  rechts ganz allein zurückbleibt. In diesem Fall ist die doppelte Rotation angezeigt, die nur einen Teilbaum von  $Y$ , nämlich  $Y_1$ , nach links verlagert.

## AUFGABE 12.13

Nimm Dir ein Stück Papier und male die beiden Rotationen aus 12.5 so auf, dass jeweils links Übergewicht herrscht. □

Diese Rotationen übertragen wir jetzt in den Rumpf von `make-balanced-node`. Damit wir dabei möglichst keine Fehler machen, verwenden wir Namen genau wie in Abbildung 12.5:

```
(define make-balanced-node
  (lambda (label left-branch right-branch)
    (define left-size (sized-tree-size left-branch))
    (define right-size (sized-tree-size right-branch))
    (cond
      ((< (+ left-size right-size) 2)
        (make-sized-node label left-branch right-branch))
      (> right-size (* ratio left-size))
      (define a label)
      (define X left-branch)
      (define c (sized-node-label right-branch))
      (define Y (node-left-branch right-branch))
      (define Z (node-right-branch right-branch))
      ...)
    ...)))
```

## AUFGABE 12.14

Die Definitionen von `c`, `Y` und `Y` überprüfen gar nicht, ob `right-branch` wirklich ein Knoten ist. Könnte es nicht auch ein Blatt sein? □

Als nächstes ist die Rotation dran: Die Funktion muss noch entscheiden, ob sie eine einfache oder eine doppelte Rotation vornimmt. Das hängt davon ab, ob der `Y` oder `Y` größer ist:

```
(if (< (sized-tree-size Y) (sized-tree-size Z))
    ... ; einfach
    ...) ; doppelt
```

Bei der einfachen Rotation müssen wir die rechte Seite von Abbildung 12.5 in Code übersetzen:



```
(if (< (sized-tree-size Y) (sized-tree-size Z))
    (make-sized-node c
                     (make-sized-node a X Y)
                     Z)
    ...)
```

Bei der doppelten Rotation stehen in der Abbildung noch weitere Namen für die Teilbäume von *Y*. Es ist sinnvoll, die ebenfalls zu binden. Leider erlaubt **if** nicht die Bindung von lokalen Variablen, wir müssen daraus ein **cond** machen:

```
(cond
  ((< (sized-tree-size Y) (sized-tree-size Z))
   (make-sized-node c
                    (make-sized-node a X Y)
                    Z))
  (else
   (define b (sized-node-label Y))
   (define Y1 (node-left-branch Y))
   (define Y2 (node-right-branch Y))
   (make-sized-node b
                    (make-sized-node a X Y1)
                    (make-sized-node c Y2 Z))))
```

#### AUFGABE 12.15

Bei der doppelten Rotation nimmt die Funktion an, dass *Y* ein Knoten ist. Könnte es nicht auch ein Blatt sein? □

Wir benötigen in der Funktion nun noch den umgekehrten Fall, dass also das Übergewicht links und nichts rechts ist. Der Code dafür ist gegenüber dem gerade behandelten Fall «gespiegelt». Auch der letzte Fall fehlt noch: Der tritt dann ein, wenn der Baum kein exzessives Ungleichgewicht aufweist. Dann könnten wir einfach **make-sized-node** benutzen. Hier die ganze Funktion:

```
(define make-balanced-node
  (lambda (label left-branch right-branch)
    (define left-size (sized-tree-size left-branch))
    (define right-size (sized-tree-size right-branch))
    (cond
```

---

```

((< (+ left-size right-size) 2)
 (make-sized-node label left-branch right-branch))
(> right-size (* ratio left-size))
(define a label)
(define X left-branch)
(define c (sized-node-label right-branch))
(define Y (node-left-branch right-branch))
(define Z (node-right-branch right-branch))
(cond
  ((< (sized-tree-size Y) (sized-tree-size Z))
   (make-sized-node c
                     (make-sized-node a X Y)
                     Z))
  (else
   (define b (sized-node-label Y))
   (define Y1 (node-left-branch Y))
   (define Y2 (node-right-branch Y))
   (make-sized-node b
                     (make-sized-node a X Y1)
                     (make-sized-node c Y2 Z))))))
(> left-size (* ratio right-size))
(define c label)
(define a (sized-node-label left-branch))
(define X (node-left-branch left-branch))
(define Y (node-right-branch left-branch))
(define Z right-branch)
(cond
  ((< (sized-tree-size Y) (sized-tree-size X))
   (make-sized-node a
                     X (make-sized-node c Y Z)))
  (else
   (define b (sized-node-label Y))
   (define Y1 (node-left-branch Y))
   (define Y2 (node-right-branch Y))
   (make-sized-node b
                     (make-sized-node a X Y1)
                     (make-sized-node c Y2 Z))))))

```

```

                                (make-sized-node a X Y1)
                                (make-sized-node c Y2 Z))))))
    (else
      (make-sized-node label left-branch right-branch))))))

```

Es gibt bisher noch keine Tests für `balanced-search-tree-insert` oder `make-balanced-node`. Das ist bedenklich, insbesondere weil die Funktionen ziemlich kompliziert sind und viele Fälle abdecken müssen. Es lohnt sich, ein weiteres Mal an Mantra 7 auf Seite 38 zu erinnern:

#### MANTRA 7

Wenn Deine Funktion Dir kompliziert scheint, ist es wahrscheinlich, dass sie Fehler enthält: Entweder, weil sie eigentlich einfacher sein sollte, oder weil die Aufgabe kompliziert und damit ihre Lösung fehleranfällig ist.

So richtig testen werden wir sie im nächsten Kapitel. Hier schon einmal eine Aufgabe zur Vorbereitung:

#### AUFGABE 12.16

Mach Dir schomal ein paar Gedanken darüber, *was* eigentlich getestet werden müsste. Reicht es, die Funktion so zu testen wie `search-tree-insert`? ☐

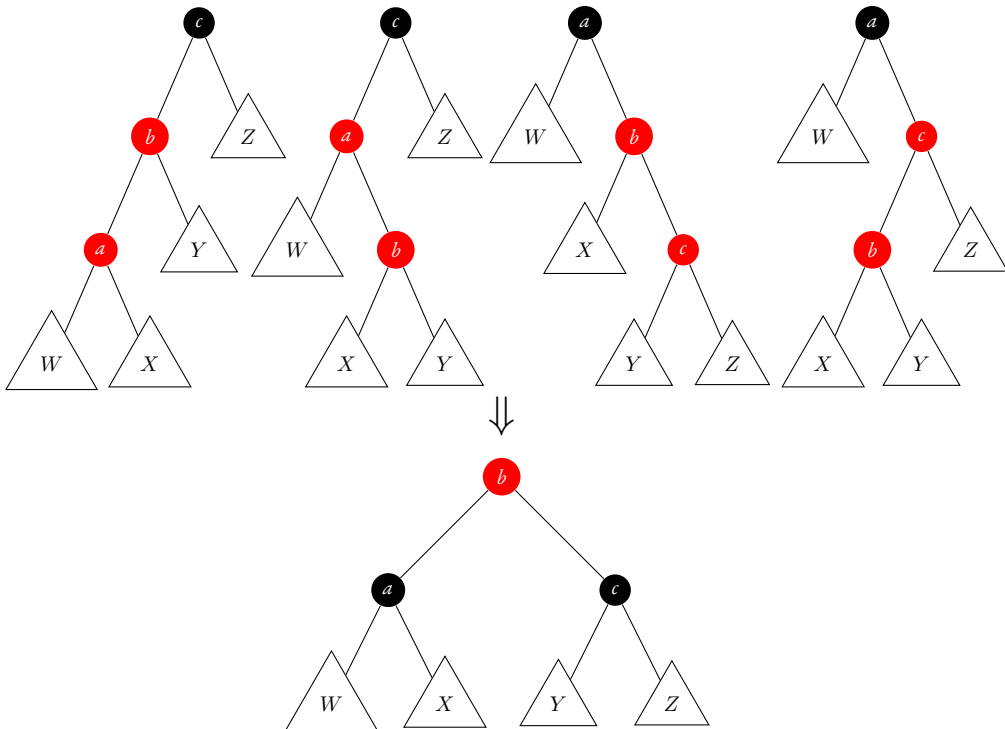
## AUFGABEN

#### AUFGABE 12.17

Schreibe eine Funktion, die einen Binärbaum akzeptiert und eine Liste aller Markierungen in dem Baum zurückgibt.

#### AUFGABE 12.18

Die Implementierung von Suchbäumen ist für viele Suchprobleme nicht mächtig genug, da `search-tree-member` nur überprüft, *ob* ein Element in einem Suchbaum vorhanden ist. Das hilft nicht zum Beispiel beim Suchen von Telefonnummern zu gegebenen Namen. Erweitere die Implementierung so, dass sie zum Beispiel auch zum Suchen von Telefonnummern verwendet werden kann. Realisiere exemplarisch das Suchen nach Telefonnummern!



**Abbildung 12.6:** Red-Black-Trees

Hinweis: Benutze als Markierungen im Suchbaum sogenannte *Einträge*, die aus einer *Schlüssel* (zum Beispiel dem Namen) und dem *Wert* bestehen. Schreibe dazu Daten-, Record- und Signatur-Definitionen. Ändere `search-tree-insert` dahingehend, dass es Schlüssel und Element akzeptiert. Schreibe eine Funktion `search-tree-find`, die zu einem Schlüssel den zugehörigen Wert findet.

#### AUFGABE 12.19

In Abbildung 12.6 ist eine alternative Strategie abgebildet, um Suchbäume auszubalancieren. Diese Strategie benutzt nicht größenannotierte Bäume sondern auf solchen, wo an jedem Knoten eine *Farbe* Teil der Markierung ist. Es gibt nur die Farben Rot und Schwarz, entsprechend heißen diese Bäume *Red-Black-Trees*. (Falls Du die Abbildung in Schwarz-Weiß anschaut: Die grauen Markierungen sind rot, die schwarzen schwarz.) Blätter haben immer die Farbe Schwarz.

Für einen Red-Black-Tree gelten die folgenden Bedingungen:

- Kein roter Knoten hat einen roten Teilbaum.

- Auf jedem von der Wurzel ausgehenden Weg zu den Blättern ist die Anzahl der schwarzen Knoten gleich.

Durch diese Bedingungen wird garantiert, dass der längste mögliche Weg, also ein Weg bei dem sich rote und schwarze Knoten abwechseln, weniger als doppelt so lang wie der kürzestmögliche Weg ist. (Der kürzestmögliche Pfad besteht nur aus schwarzen Knoten.) Soll ein Red-Black-Tree um einen Knoten erweitert werden, muss nach dem Einfügen diese Bedingung immer noch gelten, er muss also entsprechend ausbalanciert werden.

Programmieren Red-Black-Trees mit Hilfe der folgenden Anleitung!

- Schreibe zunächst Daten- und Record-Definition für Red-Black-Trees analog zu größenanno-tierten Bäumen und Suchbäumen.
- Schreibe Funktionen `red-tree?` und `black-tree?`, die überprüfen, ob ein (Teil)baum rot oder schwarz ist.
- Schreibe eine Funktion `red-black-tree-member?` analog zu `sized-tree-member?`.
- Schreibe eine Funktion `red-black-tree-insert` analog zu `balanced-search-tree-insert`.

#### AUFGABE 12.20

Schreibe eine Funktion `search-tree-delete`, die ein Element aus einem Suchbaum entfernt.

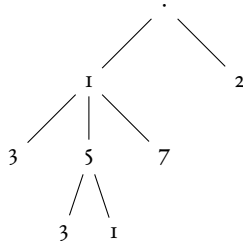
#### AUFGABE 12.21

Entwickle eine zu Listen alternative Repräsentation für endliche Folgen auf Basis folgender Daten-  
definition:

```
; Eine AListe ist eins der folgenden:
; - die leere AListe
; - eine einelementige AListe
; - eine Aneinanderhängung von zwei AListen
```

Programmiere einige der Funktionen aus Kapitel 6 auf dieser neuen Repräsentation, zum Beispiel `list-sum`, `concat`, `list-map` und `list-filter`. Welche Vor- und Nachteile hat diese Reprä-  
sentation gegenüber Listen?

#### AUFGABE 12.22



**Abbildung 12.7:** Rosen-Suchbaum

Bäume müssen nicht immer binär sein: Es ist auch möglich, Bäume so zu definieren, dass ein Knoten statt immer zwei eine beliebig lange Liste von Teilbäumen mit sich führt. Die entstehenden Bäume heißen «Rosenbäume» oder auf Englisch «rose tree».

1. Schreibe eine Datendefinition für Rosenbäume und übersetze sie in Code!
2. Benötigst Du noch eine separate Definition für Blätter?
3. Man kann Rosenbäume verwenden, um eine spezielle Form von Suchbaum zu implementieren, bei dem die Einträge Listen sind

Der Suchbaum hat für jeden Eintrag einen Pfad im Baum, bei dem die Knoten entlang des Pfades die Elemente der Liste sind. Die Markierung an der Wurzel wird ignoriert.

Abbildung 12.7 zeigt ein Beispiel für einen Rosen-Suchbaum, in dem die Listen `#list<1 3>`, `#list<1 5 3>`, `#list<1 5 1>`, `#list<1 7>` und `#list<2>` stehen.

Schreibe Funktionen analog zu `search-tree-insert` und `search-tree-member?` für Rosen-Suchbäume!

## 13 EIGENSCHAFTEN VON FUNKTIONEN

Woher wissen wir eigentlich, dass die Funktion, die wir geschrieben haben, auch richtig funktioniert? Wenn wir mit Konstruktionsanleitungen arbeiten, stehen die Chancen nicht schlecht. Die systematische Konstruktion hilft, von vornherein die Funktion richtig zu schreiben. Aber Kontrolle ist besser: Die Tests helfen, etwaige Fehler zu finden. Leider funktioniert das nicht immer, weil jeder Tests nur ein einzelnes Beispiel überprüft. Zumindest war das bisher so. In diesem Kapitel ändern wir das, in dem wir statt einzelner Beispiele allgemeine *Eigenschaften* von Funktionen formulieren. Aus diesen Eigenschaften können wir automatisch Tests machen, die effektiver sind als die bisherigen, auf `check-expect` basierenden Tests. Für hundertprozentige Sicherheit können wir gelegentlich auch Eigenschaften mathematisch beweisen. Dieses Kapitel zeigt, wie es geht.

### 13.1 KORREKTHEIT UND TESTS

Erinnerst Du Dich noch an die Funktion `heat-water` aus Abschnitt 2.5 auf Seite 51? Die hatte ziemlich komplizierte Verzweigungen. Dabei sind auch ein paar Fehler passiert. Zum Schluss hatten wir `heat-water` in zwei Funktionen aufgeteilt, `heat->temperature` und `temperature->heat`. Hier ist die erste davon, wobei wir die Bedingungen in der Verzweigung etwas vereinfacht haben:

```
; Aus Wärme Temperatur berechnen
(: heat->temperature (real -> real))

(define heat->temperature
  (lambda (heat)
    (cond
      ((<= heat 0) heat)
      ((<= heat 80) 0)
      ((<= heat 180) (- heat 80))
      (else 100))))
```

Da waren außerdem ziemlich viele Testfälle:

```
(check-expect (heat->temperature -50) -50)
(check-expect (heat->temperature 0) 0)
```

```
(check-expect (heat->temperature 20) 0)
(check-expect (heat->temperature 80) 0)
(check-expect (heat->temperature 81) 1)
(check-expect (heat->temperature 180) 100)
(check-expect (heat->temperature 200) 100)
```

Die Testfälle haben eigentlich zwei Aufgaben:

1. Sie sollen als Beispiele die Funktion *dokumentieren*.
2. Sie sollen außerdem sicherstellen, dass die Funktion *korrekt* programmiert ist.

Bei dieser Funktion allerdings gibt es bei beiden Aspekten Probleme. Es sind so viele Testfälle und sie sind willkürlich und scheinbar zufällig ausgewählt. Das macht es Leserinnen und Lesern schwer, das Wirkprinzip dahinter zu erkennen. Das führt dazu, dass uns (und vielleicht auch Dich) das Gefühl nicht loslässt, dass bei den Testfällen noch Fehler durchschlüpfen könnten.

#### AUFGABE 13.1

Ändere die Funktion absichtlich so, dass sie einen Fehler enthält, und zwar so, dass trotzdem noch alle Testfälle erfolgreich laufen. □

Ein `check-expect`-Test ist eben leider immer nur ein einzelnes Beispiel, was seine Aussagekraft einschränkt.<sup>1</sup> Oft ist eine allgemeine Aussage die bessere Dokumentation.

In diesem Fall könnten wir zum Beispiel aussagen, dass es um Wasser geht, die Temperatur also niemals größer als 100°C sein kann. Das geht nicht nur als Text, sondern auch als ein Stück Code:

```
(<= (heat->temperature heat) 100)
```

Für sich genommen ergibt dieser Ausdruck keinen Sinn: `heat` ist nicht gebunden. Wir brauchen noch den Zusatz, dass die Aussage *für alle* Werte von `heat` gilt. Also streng genommen auch nicht für wirklich *alle* Werte, nur alle Werte der Signatur `real`. Das können wir tatsächlich hinschreiben, mit Hilfe einer neuen Form namens `for-all`:

```
(for-all ((heat real))
  (<= (heat->temperature heat) 100))
```

Das `for-all` bedeutet, wie der Name schon sagt, «für alle». Da steht:

Für *alle* reellen Zahlen namens `heat` muss das Ergebnis von `(heat->temperature heat)` kleiner oder gleich 100 sein.

<sup>1</sup> In der professionellen Entwicklung heißt ein solcher Test *Unit-Test*..



**For-all** ermöglicht das Formulieren von *Eigenschaften*. Ein **for-all**-Ausdruck hat die folgende allgemeine Form:

```
(for-all ((var1 sig1) ... (varn sign)) b)
```

Dabei müssen die  $var_i$  Variablen sein, die  $_i$  Signaturen und  $b$  (der Rumpf) ein Ausdruck, der entweder einen booleschen Wert oder eine Eigenschaft liefert. Der **for-all**-Ausdruck hat als Wert eine Eigenschaft, die besagt, dass  $b$  gilt für *alle* Werte der  $var_i$ , welche die Signaturen  $sig_i$  erfüllen.

**Abbildung 13.1:** **for-all**

«Warum sind da doppelte Klammern um **heat real**?» wunderst Du Dich vielleicht. Das liegt daran, dass in dem äußeren Klammernpaar mehrere Variablen vorkommen können, jede davon mit Signatur in einem inneren Klammernpaar. Dafür wird es noch Beispiele geben. Abbildung 13.1 beschreibt genauer, wie **for-all**-Ausdrücke im allgemeinen aufgebaut sind.

Das Ergebnis des **for-all**-Ausdrucks wird in der REPL etwas undurchsichtig angezeigt:

```
(for-all ((heat real))
  (<= (heat->temperature heat) 100))
↪ #<:property>
```

Auf deutsch heißt «property» «Eigenschaft», denn es handelt sich um eine Eigenschaft von **heat-temperature**.

Diese Eigenschaft ersetzt nicht (immer) die Unit-Tests, ist aber eine wertvolle Ergänzung der Dokumentation.

Sie kann außerdem auch dazu beitragen, die Korrektheit sicherzustellen. Dazu wickeln wir um die Eigenschaft noch ein **check-property**:

```
(check-property
  (for-all ((heat real))
    (<= (heat->temperature heat) 100)))
```

**Check-property** macht wie **check-expect** oder **check-within** einen Testfall.

Da **heat->temperature** korrekt programmiert ist, meldet **check-property** auch nur einen bestandenen Test. Wozu **check-property** fähig ist, sieht man erst, wenn die Funktion einen Fehler enthält. Wenn wir zum Beispiel aus der 180 eine 280 machen, dann erscheint folgende Meldung:

Eigenschaft falsifizierbar mit heat = 206

Wichtig: Wenn Du das bei Dir ausprobierst, kann die konkrete Zahl eine andere sein.

Dickes Wort, «falsifizierbar»: Es heißt, das Racket ein *Gegenbeispiel* für die Eigenschaft gefunden hat. Wir können das nachprüfen:

`check-property` testet eine Eigenschaft analog zu `check-expect`. Eine `check-property`-Form sieht so aus:

```
(check-property e)
```

$e$  ist ein Ausdruck, der eine Eigenschaft liefern muss – in der Regel also ein `for-all`-Ausdruck. Die Form testet dann diese Eigenschaft. (Mehr dazu im nächsten Abschnitt.)

`check-property` funktioniert nur für Eigenschaften, bei denen aus den Signaturen sinnvoll Werte generiert werden können. Dies ist für die meisten Signaturen der Fall, aber nicht für `any` und Signaturvariablen definiert wurden.

Abbildung 13.2: `check-property`

```
(heat->temperature 206)
```

```
↪ 126
```

... und 126 ist größer als 100. Das Gegenbeispiel, das der `check-property`-Test gefunden hat, könnte, wenn wir den Fehler nicht absichtlich gemacht hätte, dabei helfen, ihn zu finden und zu beseitigen.

Abbildung 13.2 beschreibt den Aufbau von `check-property` genau.

### AUFGABE 13.2

Mache noch absichtlich ein paar weitere Fehler in `heat->temperature`. Welche davon werden von dem `check-property`-Test gefunden und welche nicht? □

Da `check-property` eine Eigenschaft testet, heißt diese Technik auch *property-based testing*. Die werden wir im Laufe dieses Kapitels noch auf andere Funktionen anwenden.

## 13.2 WIE CHECK-PROPERTY FUNKTIONIERT

Zunächst einmal machen wir aber einen kleinen Exkurs: Was passiert eigentlich bei so einem `check-property`-Test?

Toll wäre natürlich, wenn dieser mit Gewissheit sagen könnte, was die Eigenschaft besagt: Dass zum Beispiel die Temperatur wirklich für *alle* Eingaben höchstens 100 ist. Das ist im allgemeinen leider unmöglich.<sup>2</sup> In vielen Fällen ist es trotzdem möglich, Eigenschaften von Funktionen formal

<sup>2</sup> Dass das unmöglich ist, wurde mathematisch bewiesen und als Satz von Rice festgehalten. Der ist Thema in der theoretischen Informatik.

zu beweisen. Das ist allerdings (noch) meist recht aufwendig. (Dazu mehr in Abschnitt 13.10 auf Seite 445.)

**Check-property** kann also eine Eigenschaft nicht im allgemeinen beweisen. Stattdessen führt es Stichproben durch: Dafür wählt für die angegebenen Signaturen zufällig Werte aus, und wiederholt diesen Prozess, um so aus einem Testfall viele individuelle Tests zu machen – typischerweise mehr als 100.

Die Verwendung des Begriffs «zufällig» ist in diesem Zusammenhang in der Informatik so üblich, ein besseres Wort wäre aber «chaotisch». Tatsächlich produziert **check-property** bei jedem Durchlauf des Programms die gleichen Tests.

Die Technik des *property-based testing*, also zunächst allgemeine Eigenschaften zu formulieren und für diese dann Testfälle zu erzeugen, ist ursprünglich unter dem Namen *QuickCheck* veröffentlicht worden [CLAESSEN and HUGHES, 2000] und war ein großer Durchbruch bei der Effektivität von automatischen Tests.

### 13.3 MEHR EIGENSCHAFTEN UND INEXAKTE ZAHLEN

Es geht's weiter mit **heat-water**: **Heat->temperature** ist nur eine Hilfsfunktion dafür, zusammen mit **temperature->heat**. Auch hier könnten wir eine Eigenschaft aufschreiben, die etwas über den Zahlenbereich aussagt, der aus der Funktion herauskommt.

Allerdings gibt es noch eine weitaus ergiebigere Eigenschaft: **temperature->heat** soll ist ja gerade **heat->temperature** «umdrehen». Daraus können wir folgende Eigenschaft beziehungsweise folgenden Test machen:

```
(check-property
 (for-all ((temp real))
  (= (heat->temperature (temperature->heat temp))
     temp)))
```

Der besagt also dass, wenn eine Temperatur in Wärme gewandelt wird und wieder zurück in eine Temperatur, dass dann das gleiche herauskommen soll. Leider schlägt der Test fehl:

Eigenschaft falsifizierbar mit `temp = 0`

Wir können in der REPL für **temp** mal 0 einsetzen:

```
(heat->temperature (temperature->heat 0))
↪ cond: alle Bedingungen ergaben #f
```

In einer Eigenschaft steht die Form

$(\Rightarrow p\ e)$

dafür, dass die Eigenschaft  $e$  nur dann gelten muss, wenn der Ausdruck  $p$  den Wert  $\#t$  ergibt. Bei einer  $\Rightarrow$ -Form generiert `check-property` nur solche Tests, bei denen  $p\ \#t$  ergibt.

**Abbildung 13.3:** Voraussetzung bei Eigenschaften

Da war doch was? Vielleicht Erinnerst Du Dich: Eine Temperatur von  $0^\circ\text{C}$  kann nicht eindeutig einer Wärmezahl zugeordnet werden, die kann zwischen 0 und 80 liegen. Deshalb weigert sich auch `heat->temperature`, für die Eingabe 0 ein Ergebnis zu produzieren. Wir müssen also unseren Test anpassen, so dass da steht «für alle reellen Zahlen *außer* 0».

Das geht folgendermaßen:

```
(check-property
  (for-all ((temp real))
    ( $\Rightarrow$  (not (= temp 0))
      (= (heat->temperature (temperature->heat temp))
         temp))))
```

Der Pfeil  $\Rightarrow$  ist neu und funktioniert nur im Kontext einer Eigenschaft: Er bedeutet «gilt unter der Voraussetzung». Abbildung 13.3 erklärt im Detail, wie die Form funktioniert. Hier steht also, dass die Gleichung nur gelten muss unter der Voraussetzung, dass `temp` nicht 0 ist.

*Anmerkung:* Du könntest die Eigenschaft oben auch mit `if` statt  $\Rightarrow$  hinschreiben:

```
(check-property
  (for-all ((temp real))
    (if (= temp 0)
      #t
      (= (heat->temperature (temperature->heat temp))
         temp))))
```

Bedeutend würde die Eigenschaft so das gleiche. Allerdings behandelt `check-property` diese Schreibweise anders, nämlich schlechter. Wenn `check-property` 100 Tests generiert, dann werden alle, bei denen `temp` 0 ist, als Erfolg gewertet, obwohl da eigentlich nichts getestet wird. Wenn für (hypothetisch) drei von den Tests `temp` 0 ist, dann werden also nur 97 richtige Tests durchgeführt.

In der Variante mit `==>` allerdings stellt `check-property` sicher, dass tatsächlich 100 Tests durchgeführt werden, bei denen `temp` nicht 0 ist. Das ist besser, vor allem bei Tests, bei denen die Bedingung für viele mögliche Werte gilt, nicht nur für den einen wie hier.

Wenn wir den Testfall laufen lassen, gibt es allerdings eine merkwürdige Überraschung:

Eigenschaft falsifizierbar mit `temp = #i24.571428571428573`

Wie schon gesagt, die konkrete Zahl kann anders aussehen, aber es geht um das merkwürdige `#i`. Das steht für «inexakt», weil es sich um eine sogenannte *inexakte Zahl* handelt. Solche Zahlen werden von DrRacket (und so gut wie allen anderen Programmiersprachen) verwendet, um die Ergebnisse von Berechnungen darzustellen, die (zumindest mit vertretbarem Aufwand) nicht ganz genau durchgeführt werden können.

Bisher ging in diesem Buch alles noch ganz genau, weil unsere Programme bisher intern exakte Bruchrechnung verwendet haben. Um so eine inexakte Zahl zu berechnen, kannst Du zum Beispiel das hier in der REPL ausprobieren, um die Quadratwurzel («square root») von 2 auszurechnen:

`(sqrt 2)`

↪ `#i1.4142135623730951`

Die Wurzel von 2 hat unendlich viele Nachkommestellen, weswegen Racket davon nur einige ausrechnet und rundet. Und damit wir und Du wissen, dass gerundet wurde, steht das `#i` davor.

Das mit dem Runden ist sogar noch komplizierter, als es scheint: Es wird nämlich *binär* gerechnet. Wie genau abläuft, ist ziemlich kompliziert und würde ein weiteres Buch füllen. Mehr zu dem Thema findet sich zum Beispiel im Standardwerk von David Goldberg [GOLDBERG, 1991].

Aber was ist denn nun genau bei unserer Eigenschaft passiert? Wir können die `#i`-Zahl von Hand in die Eigenschaft einsetzen und in der REPL auswerten:

`(heat->temperature (temperature->heat #i24.571428571428573))`

↪ `#i24.57142857142857`

Du kannst sehen, dass offensichtlich beim Rechnen gerundet wurde, und zwar bei der letzten Nachkommastelle. Das erscheint Dir vielleicht merkwürdig, weil in `heat->temperature temperature->heat` doch ausschließlich addiert und subtrahiert wird – da ist keine Spur von «Runden». Wenn wir das `#i` weglassen, wird auch exakt gerechnet:<sup>3</sup>

<sup>3</sup> Falls Du es mal mit einer der anderen Sprachen zu tun hast, die beim Racket-System dabei sind: Bei den meisten von ihnen wird, anders als hier, jede Zahl mit Dezimalpunkt als inexakt behandelt.

```
(heat->temperature (temperature->heat 24.571428571428573))
↪ 24.571428571428573
```

Das liegt daran, dass «Inexaktheit» ansteckend ist: Wenn beim Aufruf einer Rechenfunktion wie `+` oder `*` auch nur eine Eingabe inexakt ist, wird gerundet. Bei unserer Eigenschaft haben für `temp` die Signatur `real` angegeben: In dieser Signatur sind auch inexakte Zahlen enthalten, deshalb nimmt da das Problem seinen Anfang. Wir können es auf zwei Arten angehen:

- Wir ersetzen in der Eigenschaft die Signatur `real` durch `rational`. In `rational` sind keine inexakten Zahlen drin. Das hat allerdings den Nachteil, dass auch nur auf exakten Zahlen getestet wird, obwohl die Funktionen auch auf inexakten Zahlen funktionieren.
- Wir berücksichtigen den Effekt der Rundung, indem wir die Bedingung in der Eigenschaft etwas aufweichen.

Wir machen letzteres und fordern nur, dass der Abstand zwischen echtem und erwartetem Ergebnis einen bestimmten Betrag nicht überschreitet:

```
(check-property
  (for-all ((temp real))
    (==> (not (= temp 0))
      (<= (abs
            (- (heat->temperature (temperature->heat temp))
              temp))
          0.0000001))))
```

Zur Erinnerung: Die eingebaute Funktion `abs` berechnet den Absolutbetrag, dreht also bei negativen Zahlen das Vorzeichen um, siehe Abschnitt 11.8 auf Seite 341.

Leider schlägt der Test immer noch fehl, zum Beispiel mit folgender Ausgabe:

```
Eigenschaft falsifizierbar mit temp = 105
```

Das können wir in der REPL ausprobieren:

```
(heat->temperature (temperature->heat 105))
cond: alle Bedingungen ergaben #f
```

Das liegt daran, dass `temperature->heat` nur für Temperaturen bis 100°C funktioniert: Wasser kann ja nicht heißer werden. Wir müssen also unsere Vorbedingung erweitern:

```
(check-property
  (for-all ((temp real))
    (==> (and (not (= temp 0))
```

```

    (< temp 100))
  (<= (abs
    (- (heat->temperature (temperature->heat temp))
      temp))
    0.0000001))))

```

Dieser Text drückt ganz gut aus, wie `heat->temperature` und `temperature->heat` zueinander stehen: Die eine dreht die andere um, zumindest ungefähr. In der Mathematik heißt es, dass die eine Funktion die *Inverse* andere anderen Funktion ist.

### AUFGABE 13.3

Versuchen, den letzten `check-property`-Test auszutricksen. Anstatt kleine Fehler einzuführen, versuche es mal mit ganz anderen Funktionen, die gar nichts mit Wassertemperaturen zu tun haben, aber trotzdem die obige Eigenschaft haben. □

Die Aufgabe zeigt, dass Eigenschaften kein Garant für Korrektheit sind: Genau wie bei Unit-Tests auch braucht es oft mehrere davon oder zusätzliche `check-expect`-Tests, um für genug Sicherheit zu sorgen.

Die Technik aus der Aufgabe ist dabei hilfreich: Überlege, wie Du einen Testfall durch falsche Funktionen austricksen kannst. Füge dann Testfälle hinzu, die diese Fehler aufspüren.

## 13.4 RELATIONALE PROBLEME

Die Eigenschaften für `heat-water` hätten wir auch durch eine (lange) Reihe von `check-expect`-Tests ersetzen können. Die Eigenschaften sind da hilfreich, aber nicht unverzichtbar. Aber es gibt Funktionen, bei denen Unit-Tests grundsätzlich nicht das richtige sind, nämlich solche, die sogenannte *relationale Probleme* lösen. Um die geht es in diesem Abschnitt.

Wir schreiben zunächst eine solche Funktion, um das Konzept zu erklären: Sie soll die Mitglieder einer Band nach Alter sortieren. Hier sind Daten- und Record-Definition für ein Bandmitglied:

```

(define-record band-member
  make-band-member
  (band-member-name string)
  (band-member-born natural))

```

Und hier die konkrete Band dazu (Stand 2021):

```
(define axl (make-band-member "Axl Rose" 1962))
(define duff (make-band-member "Duff McKagan" 1964))
(define slash (make-band-member "Slash" 1965))
(define dizzy (make-band-member "Dizzy Reed" 1963))
(define richard (make-band-member "Richard Fortus" 1966))
(define frank (make-band-member "Frank Ferrer" 1966))
(define melissa (make-band-member "Melissa Reese" 1990))

(define guns-n-roses
  (list axl duff slash dizzy richard frank melissa))
```

Für das Sortieren gibt viele Algorithmen, wir machen das mit einem besonders einfachen, wenn auch ineffizienten Verfahren namens *Insertionsort*. Die Idee ist folgende: Die Sortierfunktion arbeitet mit einer sortierte Liste als Akkumulator. Diese ist anfänglich leer, und die Funktion fügt jeweils ein Element aus der Eingabeliste hinzu, indem sie dies an der richtigen Stelle einfügt.

Hier Kurzbeschreibung und Signatur für die Hilfsfunktion zum Einfügen:

```
; Bandmitglied in eine sortierte Liste einfügen
(: insert (band-member (list-of band-member) -> (list-of band-
  member)))
```

Da dies eine Hilfsfunktion für die eigentliche Sortierfunktion ist, heben wir uns das Testen ausnahmsweise bis zu dieser auf. Konstruieren tun wir streng nach Anleitung. Gerüst und Schablone:

```
(define insert
  (lambda (band-member list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ...
       (first list)
       (insert band-member (rest list))
       ...))))
```

Im ersten Fall fügt die Funktion eine leere Liste ein: Das Ergebnis sollte dann die Liste mit **element** als einzigem Element sein. Im zweiten Fall ist noch unklar, wo **band-member** eingefügt wird, vor oder nach **(first list)**. Die Funktion muss die beiden Geburtsjahre miteinander vergleichen:



```

(define insert
  (lambda (band-member list)
    (cond
      ((empty? list) (cons band-member empty))
      ((cons? list)
       (if (<= (band-member-born band-member)
              (band-member-born (first list)))
           (cons band-member list)
           (cons (first list)
                  (insert band-member (rest list)))))))

```

Mit Hilfe von `insert` bauen wir nun die Funktion `sort-band`. Kurzbeschreibung, Signatur und Unit-Test:

```

; Band nach Alter sortieren
(: sort-band ((list-of band-member) -> (list-of band-member)))

(check-expect (sort-band guns-n-roses)
  (list axl dizzy duff slash richard frank melissa))

```

Hier die Schablone für die Funktion – mit Akkumulator:

```

(define sort-band
  (lambda (list0)
    ; Invariante: ...
    (define accumulate
      (lambda (list acc)
        (cond
          ((empty? list) ...)
          ((cons? list)
           (accumulate (rest list)
                        ... (first list) ... acc ...))))
      (accumulate list0 ...)))

```

Beim Akkumulieren enthält `list` die schon gesehenen Elemente aus `list0`, und zwar sortiert. Daraus können wir eine Invariante formulieren:

```

; Invariante: list enthält die Bandmitglieder
; zwischen list0 und list, sortiert.

```

Damit können wir die Lücken füllen: `acc` ist beim ersten Aufruf leer. Wenn `list` leer ist, dann ist `acc` das gewünschte Endergebnis. Das neue Zwischenergebnis berechnet die Funktion mit Hilfe von `insert`:

```
(define sort-band
  (lambda (list0)
    ; Invariante: list enthält die Bandmitglieder
    ; zwischen list0 und list, sortiert.
    (define accumulate
      (lambda (list acc)
        (cond
          ((empty? list) acc)
          ((cons? list)
           (accumulate (rest list)
                        (insert (first list) acc))))))
    (accumulate list0 empty)))
```

Fertig! Halt, da ist noch ein kleines Problem:

Der tatsächliche Wert

```
#<list
#<record:band-member "Axl Rose" 1962>
#<record:band-member "Dizzy Reed" 1963>
#<record:band-member "Duff McKagan" 1964>
#<record:band-member "Slash" 1965>
#<record:band-member "Frank Ferrer" 1966>
#<record:band-member "Richard Fortus" 1966>
#<record:band-member "Melissa Reese" 1990>>
```

ist nicht der erwartete Wert

```
#<list
#<record:band-member "Axl Rose" 1962>
#<record:band-member "Dizzy Reed" 1963>
#<record:band-member "Duff McKagan" 1964>
#<record:band-member "Slash" 1965>
#<record:band-member "Richard Fortus" 1966>
#<record:band-member "Frank Ferrer" 1966>
#<record:band-member "Melissa Reese" 1990>>.
```

Woran liegt's? Richard Fortus und Frank Ferrer sind im selben Jahr geboren. Der Unit-Test haben nimmt an, dass Fortus vor Ferrer einsortiert wird, aber `sort-band` macht es aber genau umgekehrt. Deswegen ist `sort-band` nicht verkehrt: Es gibt einfach mehrere korrekte Antworten.

Der Unit-Test ist also ungünstig, selbst wenn wir die Reihenfolge so wählen, dass er nicht fehlschlägt. Wenn wir die Suchfunktion irgendwann mal ändern, zum Beispiel um sie schneller zu machen, kann sich die Reihenfolge ändern und der Test schlägt wieder fehl, auch wenn die Funktion korrekt ist.

Da es für `sort-band` für eine gegebene Eingabe mehr als ein korrektes Ergebnis geben kann, sprechen wir von einem «relationalen Problem»: Es steht nicht das präzise Ergebnis fest, nur die Beziehung («Relation») zwischen Ein- und Ausgabe. Und um die zu beschreiben, ist eine Eigenschaft das richtige Mittel. Wie könnte eine sinnvolle Eigenschaft einer Funktion aussehen, die sortiert?

Nun, dass die Ausgabe sortiert ist. Um das festzustellen, schreiben wir eine Funktion:

```
; Band sortiert?
(: band-sorted? ((list-of band-member) -> boolean))
```

Die Tests lassen unterschiedliche Reihenfolgen zu, solange sie sortiert sind:

```
(check-expect (band-sorted? (list axl dizzy duff slash frank richard
                               melissa))
               #t)
(check-expect (band-sorted? (list axl dizzy duff slash richard frank
                               melissa))
               #t)
(check-expect (band-sorted? (list dizzy axl duff slash richard frank
                               melissa))
               #f)
(check-expect (band-sorted? (list axl dizzy duff richard slash frank
                               melissa))
               #f)
```

Hier sind Gerüst und Schablone:

```
(define band-sorted?
  (lambda (list)
    (cond
      ((empty? list) ...)
      ((cons? list)
       ... (first list) ...
```

```
... (band-sorted? (rest list) ...))))
```

Der **empty**-Fall ist einfach: Eine leere Liste ist sortiert. Im **cons**-Fall ist es etwas schwieriger. Um die Reihenfolge zu überprüfen, müssen wir zwei Elemente der Liste miteinander vergleichen, da ist aber nur **(first list)**. Wir brauchen also noch das zweite Element. Das gibt es nur bei Listen mit mehr als einem Element, weswegen wir eine zweite Verzweigung brauchen:

```
(define band-sorted?
  (lambda (list)
    (cond
      ((empty? list) #t)
      ((cons? list)
       (cond
         ((empty? (rest list)) ...)
         ((cons? (rest list))
          ...
          (first list)
          (first (rest list))))
       (band-sorted? (rest list))
      ...))))
```

Der innere **empty**-Fall ist die Liste mit einem Element: Die ist auch immer sortiert. Im **cons**-Fall schließlich können wir die beiden ersten Elemente vergleichen:

```
(define band-sorted?
  (lambda (list)
    (cond
      ((empty? list) #t)
      ((cons? list)
       (cond
         ((empty? (rest list)) #t)
         ((cons? (rest list))
          (and (<= (band-member-born (first list))
                  (band-member-born (first (rest list))))
               (band-sorted? (rest list)))))))))
```

Diese Funktion können wir nun benutzen, um aufzuschreiben, dass **sort-band** immer sortierte Listen produziert:

```
(check-property
  (for-all ((list (list-of band-member)))
    (band-sorted? (sort-band list))))
```

Diese Eigenschaft bringt auf den Punkt, was `sort-band` ausmacht, nämlich dass sie sortierte Listen produziert. Sie ist also schonmal gute Dokumentation.

Ist sie auch ein guter Testfall? Vielleicht hast Du ein mulmiges Gefühl, dass wir für den Test von `sort-band` eine weitere selbstgeschriebene Funktion benutzt haben, die fast ebenso kompliziert ist wie `sort-band` selbst. Was, wenn wir in `band-sorted?` einen Fehler gemacht haben, und zwar so, dass der Eigenschafts-Testfall dann einen Fehler in `sort-band` nicht mehr findet. Das ist natürlich theoretisch möglich, ist aber unwahrscheinlich und wird umso unwahrscheinlicher, je mehr Testfälle mit aussagekräftigen Eigenschaften dazukommen.

Diese Eigenschaften sind eine Form von Redundanz, analog dazu, bei Gebäuden lieber die tragenden Wände etwas stärker zu machen als unbedingt notwendig. Ob diese Redundanz die Arbeit wert ist, eine Funktion wie `band-sorted?` zu schreiben, die nur für das Testen gut sind, hängt vom Einzelfall ab: Je wichtiger die Korrektheit der Funktion und je komplizierter sie ist, desto größer ist der Wert solcher Testfälle.

Trotzdem kann man die obige Eigenschaft austricksen, ziemlich einfach sogar:

```
(define sort-band
  (lambda (list0)
    empty))
```

Das ist natürlich ein bisschen gemein. Aber die Funktion, die den Testfall austrickst, ist einfacher, als die richtige Funktion. Es ist also einfacher, es falsch zu machen als richtig. Deshalb sollten wir nach weiteren Eigenschaften suchen, die solche einfachen aber falschen Lösungen finden. Zum Beispiel könnten wir fordern, dass die Ausgabeliste genauso lang ist wie die Eingabe:

```
(check-property
  (for-all ((list (list-of band-member)))
    (= (length (sort-band list))
       (length list))))
```

#### AUFGABE 13.4

Die folgende Version von `sort-band` besteht die bisherigen Testfälle:

```
(define sort-band
  (lambda (list)
```

```
(cond
  ((empty? list) empty)
  ((cons? list)
   (map (lambda (band-member)
         (first list))
        list))))
```

Kannst Du einen Testfall schreiben, der den Fehler in dieser Funktion findet?

□

## 13.5 KONSTRUKTIONSANLEITUNGEN FÜR TESTFÄLLE?

Die bisherigen Beispiele für Testfälle mit Eigenschaften haben Dich hoffentlich überzeugt, dass es eine gute Idee ist, solche Testfälle zu schreiben. Aber *wie* geht das eigentlich bei der nächsten Funktion, die getestet werden soll? Toll wären Konstruktionsanleitungen analog zu denen für Funktionen, die zeigen, wie wir Eigenschaften aus der Signatur der zu testenden Funktion herleiten können. Das hätte bei zumindest bei den Funktionen der bisherigen Abschnitte dieses Kapitels nicht funktioniert.

Trotzdem gibt es ein paar Dinge, die Du probieren kannst, wenn Du Eigenschaften für eine neue Funktion formulieren willst:

1. Benutze das Wissen über die Größen Deines Problems.

Beispiele:

- Wenn es um die Temperatur von Wasser geht, weißt Du, dass die Temperatur nicht größer als 100°C sein kann.
- Wenn es um das Sortieren von Listen geht, dann weißt Du, dass die Ausgabe einer Sortierfunktion sortiert sein sollte.

2. Oft gehören zu einem Problem mehrere Funktionen, die auf bestimmte Art und Weise zusammenarbeiten. Schreibe auf, wie diese Zusammenarbeit aussieht.

Beispiel: `Temperature->heat` dreht `heat->temperature` um.

3. Versuche, die bestehenden Eigenschaften durch einfache aber fehlerhafte Varianten Deiner Funktion auszutricksen. Suche dann nach Eigenschaften, die das Austricksen verhindern.

Beispiel: Die leere Liste als Resultat `sort-band` ist immer sortiert, aber trotzdem falsch.

Abschnitt 13.7 auf Seite 423 wird diese Liste noch ergänzen um Eigenschaften, die tatsächlich mit den Signaturen der zu testenden Funktionen zu tun haben.

## 13.6 SUCHBÄUME TESTEN

In Abschnitt 12.7.2 auf Seite 383 haben wir die Funktion `balanced-search-tree-insert`, und die war ziemlich kompliziert. Vielleicht ging es Dir wie uns – wir hatten ein etwas mulmiges Gefühl, ob die Funktion auch wirklich korrekt ist. Sie hat ziemlich viele Verzweigungen, nicht nur in der Funktion selbst sondern auch in der Hilfsfunktion `make-balanced-node`. Dass die Testfälle wirklich jede mögliche Form von Suchbaum testet, ist ziemlich unwahrscheinlich.

Eine weitere grundsätzliche Schwierigkeit beim Testen von `balanced-search-tree-insert` ist, dass die Funktion ein relationales Problem löst: Es gibt mehr als ein korrektes Ergebnis. Wenn wir einen `check-expect`-Testfall schreiben, der ein bestimmtes Ergebnis von `balanced-search-tree-insert` erwartet, kann es sein, dass die Funktion ein anders, aber trotzdem korrektes Ergebnis liefert. Der Testfall schlägt dann fehl, und die Suche nach der Ursache ist mühsam.

Aber mit ein paar Eigenschaften können wir (hoffentlich) das mulmige Gefühl beseitigen und unser Vertrauen in die Funktion erhöhen. Der vorige Abschnitt hatte einige Vorschläge, wie wir vorgehen können. Der erste Vorschlag war:

Benutze das Wissen über die Größen Deines Problems.

Die Größe unseres Problems ist hier der balancierte Suchbaum. In dem Begriff steckt schon ziemlich viel Wissen:

1. Das Ergebnis von `balanced-search-tree-insert` sollte ein sortierter Baum sein, bei dem alle Markierungen im linken Teilbaum eines Knotens kleiner sein sollte als die Markierung des Knotens.
2. Die Bäume sind größenannotiert: Die Größenannotation sollte auch stimmen, also bei jedem Knoten die tatsächliche Größe des Baums darunter wiedergeben.
3. Schließlich und endlich sollte der balancierte Suchbaum natürlich auch *balanciert* sein.

Wir fangen mal mit dem zweiten Punkt an, und überlassen Dir danach den ersten als Übungsaufgabe. Um zu überprüfen, ob die Größenannotation stimmt, müssen wir die Annotation jeden Knoten des Baums betrachten. Dafür brauchen wir eine Hilfsfunktion. Sie hat folgende Kurzbeschreibung und Signatur:

```
; Stimmt die Größenannotation am Suchbaum?
(: proper-sized-search-tree? ((sized-search-tree-of %a) -> boolean))
```

Hier sind zwei einfache Testfälle dafür:

```
(check-expect (proper-sized-search-tree? sized-search-tree1) #t)
(check-expect (proper-sized-search-tree?
```

```

(make-sized-search-tree
  = <
  (make-sized-node
    5
    (make-node (make-sized-label 5 3) #f #f)
    (make-node (make-sized-label 2 7) #f #f)))
  #f)

```

Nur zwei popelige Testfälle, magst Du einwenden, mehr wären sicher besser. Allerdings werden wir sie noch im Rahmen von `check-property`-Tests zusammen mit `balanced-search-tree-insert` aufrufen und wir hoffen, dass dies die benötigte Redundanz schafft, was die Korrektheit von `balanced-search-tree-insert` betrifft. Aber ist diese Hoffnung auch berechtigt? Wir können uns zwei Szenarien vorstellen, in denen das nicht der Fall ist, also der Testfall keine Fehler findet, obwohl welche vorhanden sind:

- `Proper-sized-search-tree?` liefert *immer* `#t`, dann ist jeder Testfall damit nutzlos. Das haben wir aber durch den einen `check-expect`-Testfall ausgeschlossen.
- Fehler in `proper-sized-search-tree?` «passen» genau zu den Fehlern in `balanced-search-tree-insert`. Um dem vorzubeugen, werden wir `proper-sized-search-tree?` systematisch mit Hilfe der Konstruktionsanleitungen entwickeln. Wir werden sehen, dass dies bei `proper-sized-search-tree?` deutlich einfacher ist als `balanced-search-tree-insert`. Damit ist es wahrscheinlich, dass `proper-sized-search-tree?` korrekt gelingt.

Hier das Gerüst für die Funktion:

```

(define proper-sized-search-tree?
  (lambda (search-tree)
    ...))

```

Wir gehen beim Rumpf vor wie bei der Funktion `sized-search-tree-member?` auf Seite 384. Der eigentliche Baum steckt ja in dem `search-tree`-Record drin. Den extrahieren wir und lassen die eigentliche Arbeit von einer Hilfsfunktion erledigen.

```

(define proper-sized-search-tree?
  (lambda (search-tree)
    (define proper?
      (lambda (tree)
        ...))
    (proper? (sized-search-tree-tree search-tree))))

```

Hier die Schablone:



```

(define proper-sized-search-tree?
  (lambda (search-tree)
    (define proper?
      (lambda (tree)
        (cond
          ((node? tree)
           ...
           (proper? (node-left-branch tree))
           (proper? (node-right-branch tree))
           (node-label tree)
           ...))
        (else ...))))
    (proper? (sized-search-tree-tree search-tree))))

```

Ein paar Lücken können wir schon füllen: Bei einem Blatt ist die Größenannotation immer korrekt, weil nicht vorhanden. Bei Knoten ist eine Größenannotation korrekt, wenn sie die Summe der Größen der Teilbäume plus 1 (für den Knoten selbst ist). Außerdem muss die Funktion auch die Größenannotationen der Teilbäume überprüfen, das erledigen die rekursiven Aufrufe aus der Schablone. Fertig sieht das ganze so aus:

```

(define proper-sized-search-tree?
  (lambda (search-tree)
    (define proper?
      (lambda (tree)
        (cond
          ((node? tree)
           (define left (node-left-branch tree))
           (define right (node-right-branch tree))
           (and (proper? left)
                (proper? right)
                (= (sized-label-size (node-label tree))
                   (+ 1
                      (sized-tree-size left)
                      (sized-tree-size right)))))
          (else #t))))
    (proper? (sized-search-tree-tree search-tree))))

```

Aber geschrieben hatten wir die Funktion ja, um `balanced-search-tree-insert` zu testen. Wenn die korrekt ist, dann produziert sie nur solche Suchbäume, bei denen `proper-sized-search-tree?` als Ergebnis `#t` liefert. Man könnte versucht sein, so etwas hier zu schreiben:

```
(check-property
  (for-all ((search-tree (sized-search-tree-of natural))
            (element natural))
    (proper-sized-search-tree?
      (balanced-search-tree-insert element search-tree))))
```

Das funktioniert aber nicht, weil die Signatur `(sized-search-tree-of natural)` nicht dafür geeignet ist, Suchbäume zu generieren. Insbesondere steht in der Signatur nur, dass die Größenannotation eine natürliche Zahl ist, aber nicht welche: Bei den meisten so generierten Suchbäumen wäre sie falsch. Und wenn die Eingabe für `balanced-search-tree-insert` schon falsch ist, können wir nicht erwarten, dass die Ausgabe korrekt ist: Garbage in, Garbage out.

#### AUFGABE 13.5

Noch andere Aspekte der Signatur `(sized-search-tree-of natural)` machen sie ungeeignet, korrekte Suchbäume zu generieren. Finde einen! □

Um korrekte größenannotierte Suchbäume zu erzeugen, rufen wir stattdessen `balanced-search-tree-insert` wiederholt auf. Die Elemente dafür holen wir aus einer Liste und schreiben entsprechend eine Funktion, die aus einer Liste einen Suchbaum macht. Kurzbeschreibung und Signatur:

```
; Aus allen Elementen einer Liste einen Suchbaum machen
(: list->balanced-search-tree ((%a %a -> boolean) (%a %a -> boolean)
                                (list-of %a)
                                -> (sized-search-tree-of %a)))
```

Das ist eine typische Aufgabe für `fold`: Mit einem leeren Suchbaum anfangen (für dessen Konstruktion wir die Funktionen für den Vergleich der Markierungen brauchen) und dann für jedes Listenelement `balanced-search-tree-insert` aufrufen:

```
(define list->balanced-search-tree
  (lambda (label=? label<? elements)
    (fold (make-empty-sized-search-tree label=? label<?)
          balanced-search-tree-insert
          elements)))
```

Und mit Hilfe dieser Funktion können wir endlich den Test schreiben, der sicherstellt, dass `balanced-search-tree-insert` die Größenannotation richtig macht:

```
(check-property
  (for-all ((elements (list-of natural)))
    (proper-sized-search-tree?
      (list->balanced-search-tree = < elements)))))
```

### AUFGABE 13.6

Schreibe eine Funktion analog zu `proper-sized-search-tree?`, die feststellt, ob ein Suchbaum auch wirklich sortiert ist und mache daraus eine Eigenschaft und einen `check-property`-Testfall für `balanced-search-tree-insert`! □

Aus ist der Liste vom Anfang dieses Abschnitts haben wir zwei Punkte abgehakt: Einer fehlt noch, nämlich zu testen, ob `balanced-search-tree-insert` auch wirklich balanciert sind, wie es der Name suggeriert.

Nehmen wir an, `balanced-search-tree-insert` würde *nicht* korrekt balancieren. Das wäre ziemlich tückisch, weil die Funktion in allen Tests scheinbar korrekt funktioniert. Erst, wenn sie es mit größeren Datenmengen zu tun bekommt, gibt es Probleme, weil die Laufzeit beim Suchen im Baum größer ist als erwartet.

Was heißt eigentlich *genau* balanciert? Die Funktion `balanced-search-tree-insert` benutzt dafür ja ein «weiches» Kriterium, um nicht jedesmal mit viel Aufwand alles aufzuräumen. Wie oft die Funktion neu balanciert, wird durch die Definition von `ratio` auf Seite 387 kontrolliert. Stephen Adams, der Autor des zugrundeliegenden Papers, behauptet, dass bei die Suchbäume immer so ausbalanciert sind, dass für einen Knoten, dessen linker Teilbaum die Größe  $l$  und dessen rechter Teilbaum die Größe  $r$  gilt, folgende Formel gilt:

$$\frac{l}{\text{ratio}} \leq r \leq l \times \text{ratio}$$

Aus dieser Formel können wir eine Funktion machen, die einen Suchbaum darauf überprüft, dass sie durch den gesamten Baum hindurch eingehalten ist. Das geht nach dem gleichen Muster wie bei `proper-sized-search-tree?`:

```
; Ist ein Suchbaum balanciert?
(: balanced-search-tree? ((sized-search-tree-of %a) -> boolean))
```

```

(define balanced-search-tree?
  (lambda (search-tree)
    (define balanced?
      (lambda (tree)
        (cond
          ((node? tree)
           (define left (node-left-branch tree))
           (define right (node-right-branch tree))
           (define left-size (sized-tree-size left))
           (define right-size (sized-tree-size right))
           (and (balanced? left)
                (balanced? right)
                (<= (/ left-size ratio)
                    right-size
                    (* left-size ratio)))))
          (else
           #t))))
    (balanced? (sized-search-tree-tree search-tree))))

```

Ebenfalls nach dem gleichen Muster wie bei `proper-sized-search-tree?` können wir damit eine Eigenschaft und einen `check-property`-Testfall formulieren:

```

(check-property
  (for-all ((elements (list-of natural)))
    (balanced-search-tree?
     (list->balanced-search-tree = < elements)))))

```

(Inzwischen kommen soviele Tests zusammen, dass es schon eine Weile dauert, bis sie alle durchlaufen.)

Aber hoppla, der Testfall schlägt fehlt. Bei uns hat er folgendes Gegenbeispiel produziert:

Eigenschaft falsifizierbar mit `elements =` `#<list 1 1 3>`

Wir probieren mal in der REPL aus, was für ein Suchbaum bei dieser Liste herauskommt:

```

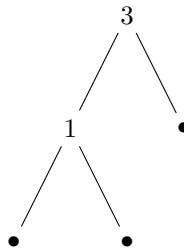
(list->balanced-search-tree = < (list 1 1 3))
↪ #<record:sized-search-tree-of #<function:=> #<function:<>
    #<record:node-of #<record:sized-label-of 2 3>
        #<record:node-of #<record:sized-label-of 1 1>

```

```
#f
#f>
#f>>
```

(Die zweite 1 hätte sich `check-property` auch sparen können. Offenbar ist es nicht ganz optimal programmiert.)

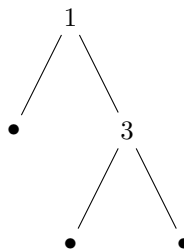
Was ist passiert? Grafisch dargestellt sieht Baum aus der REPL so aus:



Und in der Tat: Dieser Baum verletzt die Gleichung

$$\frac{l}{\text{ratio}} \leq r \leq l \times \text{ratio}.$$

In diesem Fall ist  $l = 1$  und  $r = 0$ . So richtig schlimm unbalanciert ist das allerdings gar nicht. Außerdem ist es gar nicht möglich, die Gleichung bei einem Suchbaum mit den Elementen 1 und 3 zu erfüllen. Die einzige andere Möglichkeit ist diese hier:



Hier gilt  $l = 0$  und  $r = 1$  und die Gleichung gilt ebenfalls nicht. Dass die Gleichung nicht erfüllbar ist, liegt daran, dass der Baum sehr klein ist: Nur zwei Elemente.

Dieser Umstand ist tatsächlich in der Programmierung von `balanced-search-tree-insert` berücksichtigt. Schau Dir nochmal `make-balanced-node` auf Seite 391 an. Falls Du nicht blättern magst – so geht es los:

```
(define make-balanced-node
  (lambda (label left-branch right-branch)
    (define left-size (sized-tree-size left-branch))
    (define right-size (sized-tree-size right-branch))
    (cond
      ((< (+ left-size right-size) 2)
        (make-sized-node label left-branch right-branch))
      ...)))
```

Der erste Zweig des `cond` schlägt genau dann zu, wenn der Baum sehr klein ist: Dann wird nicht rotiert oder anderweitig balanciert. Das ist bei uns gerade der Fall. Bei allen größeren Bäumen greift die Gleichung. Das ist also etwas schlampig formuliert in dem Paper, das der Funktion zugrunde liegt, und wir müssen es im Testfall berücksichtigen, indem wir

```
(<= (/ left-size ratio)
     right-size
     (* left-size ratio))))
```

ersetzen durch:

```
(or (< (+ left-size right-size) 2)
    (<= (/ left-size ratio)
         right-size
         (* left-size ratio))))
```

Und siehe da: Jetzt läuft der Testfall durch.

Es kommt immer mal wieder vor, dass ein `check-property`-Testfall Gegenbeispiele findet, die eigentlich keine sind, weil wir die Bedingung für den Test nicht ganz richtig formuliert haben. Den eigentlichen Fehler zu finden, nervt oft, aber manchmal erfahren wir auch etwas bei der Gelegenheit, was wir noch nicht gewusst haben. (In diesem Fall: Dass kleine Bäume gesondert behandelt werden müssen.)

#### AUFGABE 13.7

Was passiert, wenn Du den Zweig für kleine Bäume aus `make-balanced-node` entfernst? Findet das einer der Testfälle heraus? Was genau ist die Ursache des angezeigten Fehlers? □

Die Liste vom Anfang des Abschnitts haben wir damit abgearbeitet. Haben wir mit den bisherigen Tests ausreichend sichergestellt, dass `balanced-search-tree-insert` korrekt funktioniert?

Nun, wir haben überprüft, dass die entstehenden Suchbäume die korrekte Form haben. Allerdings haben wir noch gar nicht überprüft, dass auch der Inhalt stimmt. Dazu benutzen wir den zweiten Hinweis aus Abschnitt 13.5 auf Seite 412:

Oft gehören zu einem Problem mehrere Funktionen, die auf bestimmte Art und Weise zusammenarbeiten. Schreibe auf, wie diese Zusammenarbeit aussieht.

Die Funktion `balanced-search-tree-insert` ist ja nur insofern sinnvoll, dass die damit eingefügten Elemente mit `sized-search-tree-member?` auch gefunden werden können. Das ist gerade das gefragte Zusammenspiel mehrerer Funktionen. Zu diesem Zweck benutzen wir wieder `list->balanced-search-tree`, um aus einer Liste einen Suchbaum herzustellen und überprüfen, dass `sized-search-tree-member?` auch für jedes Element der Liste `#t` liefert:

```
(check-property
  (for-all ((elements (list-of natural)))
    (define search-tree (list->balanced-search-tree = < elements))
    (every? (lambda (element)
              (sized-search-tree-member? element search-tree))
            elements)))
```

Die Funktion `every?` haben wir extra für diesen Testfall programmiert – strikt nach Konstruktionsanleitung für Listen:

```
; Liefert eine Funktion für alle Elemente einer Liste #t?
(: every? ((%a -> boolean) (list-of %a) -> boolean))
```

```
(check-expect (every? even? (list 2 4 6)) #t)
(check-expect (every? positive? (list 1 2 3)) #t)
(check-expect (every? positive? (list 1 0 3)) #f)
```

```
(define every?
  (lambda (p? list)
    (cond
      ((empty? list) #t)
      ((cons? list)
       (and (p? (first list))
            (every? p? (rest list)))))))
```

Das ist schonmal gut, aber Du erinnerst Dich an den dritten Hinweis in Abschnitt 13.5 auf Seite 412?

Versuche, die bestehenden Eigenschaften durch einfache aber fehlerhafte Varianten  
Deiner Funktion auszutricksen.

#### AUFGABE 13.8

Finde eine einfache, aber fehlerhafte Variante der Funktion `balanced-search-tree-insert`,  
die alle bisherigen Testfälle austrickst! □

Wir benötigen also auf jeden Fall noch eine Funktion, die überprüft, dass Werte, die *nicht* in einem  
Suchbaum stehen, auch tatsächlich *nicht* gefunden werden. Wir fangen so an wie bisher:

```
(check-property
  (for-all ((elements (list-of natural)))
    ...
    (list->balanced-search-tree = < elements)
    ...))
```

Wir brauchen jetzt noch irgendeine Zahl, die wir darauf überprüfen können, dass sie nicht im Such-  
baum ist. Die holt sich der Testfall auch mit `for-all`:

```
(check-property
  (for-all ((elements (list-of natural))
            (number natural))
    ...
    (not (sized-search-tree-member?
          number
          (list->balanced-search-tree = < elements))))
  ...))
```

Ganz fertig ist das noch nicht. Wir müssen noch sicherstellen, dass `number` nicht zufällig doch im  
Suchbaum ist. Dazu benutzen wir `==>`:

```
(check-property
  (for-all ((elements (list-of natural))
            (number natural))
    (==> (not (member? = number elements))
          (not (sized-search-tree-member?
                number
                (list->balanced-search-tree = < elements))))))
```



Da steht also umgangssprachlich formuliert: Für alle Listen von Elementen `elements` und jede Zahl `number` gilt: Wenn `number` nicht zu `elements` gehört, dann liefert `sized-search-tree-member?` für `number` und den Suchbaum aus den Zahlen in `elements` das Ergebnis `#f`.

Damit haben wir das Zusammenspiel zwischen `balanced-search-tree-insert` und `sized-search-tree-member?` ausreichend beschrieben und können ziemlich ruhig schlafen in der Gewissheit, dass die Funktionen korrekt arbeiten.

## 13.7 ALGEBRAISCHE EIGENSCHAFTEN

Hast Du das hier schonmal gesehen?

$$(a + b) + c = a + (b + c)$$

Oder das hier?

$$(a \times b) \times c = a \times (b \times c)$$

Diese Gleichungen sind unter dem Namen *Assoziativität* bekannt, manchmal auch als «Assoziativgesetz». Kurz kam das Wort schonmal auf Seite 224 vor, und da steht folgende Gleichung:

$$\text{cat}(u, \text{cat}(v, w)) = \text{cat}(\text{cat}(u, v), w).$$

Die Assoziativität ist eine äußerst nützliche Gleichung, weil sie aussagt, dass in einer Aneinanderreihung von  $+$  oder  $\times$  jeweils die Klammern vollkommen egal sind: Und wenn sie schon egal sind, kann man sie auch weglassen.

Die Assoziativität ist nicht einfach eine Gleichung. (Hier sind es ja schon drei.) Tatsächlich ist die Assoziativität eine Eigenschaft – aber wovon? Zunächst einmal ist sie eine Eigenschaft von  $+$  und  $\times$  und `cat`. Entsprechend kann man auch sagen « $+$  ist assoziativ» oder «`cat` ist assoziativ». Die Funktionen  $+$ ,  $\times$  und `cat` haben allesamt etwas gemeinsam, nämlich die Struktur ihrer Signaturen. Hier sind diese Signaturen, so als wenn sie in einem Programm stünden:

```
(: + (number number -> number))
(: * (number number -> number))
(: concatenate ((list-of %a) (list-of %a) -> (list-of %a)))
```

(Statt der mathematischen Funktion `cat` haben wir die Funktion `concatenate` aus Abschnitt 6.6.3 auf Seite 183 aufgeführt.)

## AUFGABE 13.9

Sind `and` und `or` assoziativ? Wie passen ihre Signaturen zu denen von den Funktionen oben? ☐

Diese Form von Signaturen haben wir schon gesehen, als es um Kombinatoren ging, in Kapitel 5 auf Seite 131. Eine assoziative Funktion ist also eine spezielle Sorte Kombinator.

Assoziative Operationen gibt es so viele, dass es sich lohnt, auf Verdacht danach zu suchen, wenn Du eine neue Datendefinition erstellst. Und wenn Du einen assoziativen Kombinator gefunden hast, lohnt es sich, diese Erkenntnis aufzuschreiben. Im Gespräch mit anderen kannst Du dann einfach sagen «cat ist assoziativ», und im besten Fall wissen alle, was gemeint ist.

Wir können die Assoziativität aber auch als Eigenschaft mit `check-property` aufschreiben. Wir machen das erstmal mit `concatenate`, `+` und `*` kommen danach.

```
(check-property
  (for-all ((a (list integer))
            (b (list integer))
            (c (list integer)))
    (expect (concatenate (concatenate a b) c)
            (concatenate a (concatenate b c)))))
```

Du fragst Dich vielleicht, warum `(list integer)` als Signatur und nicht `(list %a)`, wie es in der Signaturdeklaration von `concatenate` steht. Das liegt daran, dass bei Signaturen in `for-all` keine Signaturvariablen stehen dürfen, damit DrRacket dafür vernünftig zufällige Werte generieren kann. (Siehe dazu auch Abbildung 13.2 auf Seite 400.)

## AUFGABE 13.10

Schreibe mit Hilfe des Teachpacks `image.rkt` Eigenschaften für die Assoziativität der Kombinatoren aus Abschnitt 5.2 auf Seite 5,2 –`overlay`, `beside` und `above`. ☐

Eine assoziative Funktion macht aus zwei «Dingen» ein «Dings». `+` macht aus zwei Zahlen eine, `concatenate` aus zwei Listen eine, `and` aus zwei booleschen Werten einen.

Über die Signaturdeklarationen von `+`, `*`, `concatenate` undsoweiter können wir abstrahieren; sie haben alle die folgende Form:

```
(: op (s s -> s))
```

Zur Assoziativität gehört also neben der Operation `op` auch die Signatur `s`, so dass `op` die Signatur `(s s -> s)` hat. (Und natürlich die Assoziativitätsgleichung selbst.)

Solche «Dreigestirne» aus «Grundsignatur», Operation (oder mehreren Operationen) und Gleichungen werden in der mathematischen Algebra studiert und klassifiziert. (Dort wird statt mit Signaturen mit Mengen hantiert.) Die besonders nützlichen, interessanten oder häufig vorkommenden dieser Dreigestirne – algebraischer Strukturen – bekommen eigene Namen.

Eine Kombination aus Signatur, Koinikator mit Signatur ( $s \ s \rightarrow s$ ) und Assoziativgesetz heißt zum Beispiel *Halbgruppe*.<sup>4</sup>

Viele Halbgruppen erfüllen noch eine weitere Eigenschaft: Sie besitzen einen Wert der Signatur, der «nichts macht», das sogenannte «neutrale Element». Das hat in Abschnitt 6.1 auf Seite 6.1 schonmal eine Rolle gespielt. Jetzt können wir es endlich richtig einordnen.

Nennen wir das neutrale Element  $n$ . Dann gilt folgende Eigenschaft:

```
(for-all ((a s))
  (and (expect (op a n) a)
        (expect (op n a) a)))
```

Das neutrale Element kann von Halbgruppe zu Halbgruppe variieren, selbst wenn die Signatur  $s$  immer die Gleiche ist. Zum Beispiel ist 0 das neutrale Element von  $+$  und 1 das neutrale Element von  $\times$ , auch wenn es beide Male um Zahlen geht.

#### AUFGABE 13.11

Was ist das neutrale Element von `or` beziehungsweise `and`? □

Bei `concatenate` ist das neutrale Element die leere Liste. Das können wir mit `check-property` festhalten:

```
(check-property
  (for-all ((a (list integer)))
    (and (expect (concatenate a empty) a)
          (expect (concatenate empty a) a))))
```

Auch für `overlay`, `beside` und `above` gibt es ein neutrales Element. Es heißt `empty-image`.

#### AUFGABE 13.12

Formuliere `check-property`-Testfälle für die Neutrales-Element-Eigenschaft von `empty-image`. □

<sup>4</sup> Wenn es «Halbgruppen» gibt, gibt es dann auch «Gruppen», magst Du fragen. Gibt es, und in der Algebra gibt es einen großen Teilbereich namens «Gruppentheorie». Gruppen gibt es in der Programmierung aber viel seltener als Halbgruppen, darum konzentrieren wir uns auf letztere.

Wie Du siehst, haben alle Halbgruppen, die wir bisher betrachtet haben, ein neutrales Element. In der Tat ist das meistens so. Sogar so oft, dass, wenn bei Deiner Datenanalyse eine Halbgruppe herausgekommen ist, Du immer auch nach einem neutralen Element Ausschau halten solltest. Eine Halbgruppe mit einem neutralen Element hat entsprechend auch einen Namen, nämlich *Monoid*.

### 13.7.1 BINÄRE OPERATIONEN

Eine allgemein bekannte Eigenschaft der Addition ist die *Kommutativität*:

$$a + b = b + a$$

Auch wenn intuitiv die Bedeutung klar ist, ist die Eigenschaft genau genommen so noch nicht präzise schriftlich festgehalten, da nicht notiert ist, was  $a$  und  $b$  sind: Die Idee ist natürlich, dass  $a$  und  $b$  beliebige *Zahlen* sind. Im allgemeinen also:

$$\forall a \in \mathbb{C}, b \in \mathbb{C} : a + b = b + a$$

(Wer sich mit der Vorstellung komplexer Zahlen nicht wohlfühlt, kann das  $\mathbb{C}$  auch durch  $\mathbb{R}$  oder  $\mathbb{Q}$  ersetzen.)

FIXME

In den Lehrsprachen lässt sich diese Eigenschaft für die eingebaute Funktion  $+$  aufschreiben – das  $\forall$  ist auf Tastaturen nicht vertreten und wird darum ausgeschrieben (siehe Abbildung 13.1):

```
(for-all ((a number)
          (b number))
  (= (+ a b) (+ b a)))
```

Interessanter wird es erst bei Eigenschaften, die nicht stimmen. Zum Beispiel ist die Subtraktion – nicht kommutativ:

```
(check-property
  (for-all ((a number)
            (b number))
    (= (- a b)
       (- b a))))
```

Hierfür liefert DrRacket folgende Meldung:

Eigenschaft falsifizierbar mit  $a = \boxed{\#i1.25}$   $b = \boxed{\#i-4.0}$

*Falsifizierbar* bedeutet, dass es ein *Gegenbeispiel* für die Eigenschaft gibt, also Werte für die Variablen **a** und **b**, welche die Eigenschaft falsch werden lassen. DrRacket hat in diesem Fall ein Gegenbeispiel gefunden, bei dem **a** den Wert `#i1.25` und **b** den Wert `#i-4.0` hat:

```
(- #i1.25 #i-4.0)
↪ #i5.25
(- #i-4.0 #i1.25)
↪ #i-5.25
```

Dieses Beispiel widerlegt also tatsächlich die Behauptung der Eigenschaft.

Hinter den Kulissen hat DrRacket verschiedene Werte für **a** und **b** durchprobiert und in die Eigenschaft eingesetzt, also effektiv nach einem Gegenbeispiel gesucht. Für die Kommutativität von **+** gibt es kein Gegenbeispiel – DrRacket konnte also auch keins finden. Dass ausgerechnet das merkwürdige Beispiel `0.0` und `-1.5` herauskam, liegt an der relativ komplexen Suchstrategie von DrRacket.

Auf diese Art und Weise lassen sich eine Reihe von interessanten Eigenschaften formulieren, so zum Beispiel die Assoziativität von **+**:

```
(check-property
  (for-all ((a number)
            (b number)
            (c number))
    (= (+ a (+ b c))
       (+ (+ a b) c))))
```

Hierbei gibt es allerdings eine böse Überraschung – DrRacket produziert ein Gegenbeispiel:

FIXME: inexakte Zahlen

Eigenschaft falsifizierbar mit

**a** = `2.6666666666666665` **b** = `6.857142857142857` **c** = `-6.857142857142857`

Es ist kein Zufall, dass es sich um Zahlen mit vielen Nachkommastellen handelt. Wenn dieses Gegenbeispiel in die Eigenschaft eingesetzt wird, liefert die REPL folgende Ergebnisse:

```
(+ 2.6666666666666665 (+ 6.857142857142857 -6.857142857142857))
↪ 2.6666666666666665
(+ (+ 2.6666666666666665 6.857142857142857) -6.857142857142857)
↪ 2.6666666666666667
```

Hier wird sichtbar dass, wie bereits in Abschnitt ?? erwähnt, bei Berechnungen mit sogenannten *inexakten Zahlen*, das sind Zahlen mit einem Dezimalpunkt, die mathematischen Operationen nur mit einer begrenzten Anzahl von Stellen durchgeführt werden und dann runden – da auch noch binär und nicht dezimal gerundet wird, sieht das Ergebnis dieser Rundung oft unintuitiv aus. Dieses Beispiel zeigt nun, dass Addition plus binäre Rundung nicht assoziativ ist. Die Assoziativität gilt nur für das Rechnen mit *exakten Zahlen*. Immerhin sind alle Zahlen mit der Signatur **rational** exakt, die Eigenschaft lässt sich also reformulieren:

```
(check-property
  (for-all ((a rational)
            (b rational)
            (c rational))
    (= (+ a (+ b c))
       (+ (+ a b) c))))
```

Und tatsächlich, in dieser Form wird die Eigenschaft nicht beanstandet.

Kommutativität und Assoziativität sind jeweils Eigenschaften einer einzelnen Operation, in diesem Fall  $+$ . Manche Eigenschaften beschreiben auch das Zusammenspiel mehrerer Operationen, wie zum Beispiel die Distributivität, die für Addition und Multiplikation gilt:

$$\forall a \in \mathbb{C}, b \in \mathbb{C}, c \in \mathbb{C} : a \times (b + c) = a \times b + b \times c$$

Auch dies lässt sich direkt in Code übersetzen, diesmal gleich mit **rational** statt **number**:

```
(check-property
  (for-all ((a rational)
            (b rational)
            (c rational))
    (= (* a (+ b c))
       (+ (* a b) (* a c)))))
```

Auch hier hat DrRacket nichts zu meckern.

Neben der Addition ist auch die Multiplikation kommutativ:

```
(check-property
  (for-all ((a rational)
            (b rational))
    (= (* a b)
       (* b a))))
```

Wenn Du diese Eigenschaft neben die Kommutativität für  $+$  legst, siehst Du, dass diese fast identisch sind und damit natürliche Kandidaten für Abstraktion: Nur die Operation  $-$   $*$  im einen und  $+$  im anderen Fall  $-$  ist unterschiedlich. Wenn wir über die Operation abstrahieren, bekommen wir so etwas wie eine allgemeine Definition der Kommutativität, und das sieht so aus:

```
(define commutativity
  (lambda (op)
    (for-all ((a rational)
              (b rational))
      (= (op a b)
         (op b a)))))
```

Mit Hilfe dieser Definition können wir die Kommutativität von  $+$  und  $*$  deutlich kompakter formulieren:

```
(check-property (commutativity *))
(check-property (commutativity +))
```

Über dem `check-property` können wir nicht abstrahieren – es muss ganz außen stehen, damit DrRacket Fehlermeldungen den dazu passenden Programmstellen zuordnen kann.

Der Vollständigkeit halber braucht `commutativity` noch eine Signatur:  $+$  und  $*$  sind jeweils Funktionen, die zwei Zahlen akzeptieren und wieder eine Zahl zurückliefern. Der Rückgabewert von `commutativity` ist eine Eigenschaft, für die in DrRacket die Signatur `property` fest eingebaut ist. Die fertige Signatur ist also diese hier:

```
(: commutativity ((rational rational -> rational) -> property))
```

Diese drei Eigenschaften – Kommutativität, Assoziativität und Distributivität – tauchen immer wieder auf, da sie nicht nur für arithmetische Operationen gelten (auch die Multiplikation ist kommutativ und assoziativ) sondern auch anderswo.

Zum Beispiel gelten Kommutativität und Assoziativität auch für das logische `and`:

```
(check-property
  (for-all ((a boolean)
            (b boolean))
    (boolean=? (and a b)
               (and b a))))
```

```
(check-property
  (for-all ((a boolean)
```

```

      (b boolean)
      (c boolean))
  (boolean=? (and a (and b c))
             (and (and a b) c))))

```

Hier muss die eingebaute Funktion `boolean=?` verwendet werden, die boolesche Werte vergleicht, analog zu `=`, die nur Zahlen vergleichen kann.

Schön wäre natürlich, wenn wir auch für die Kommutativität von `and` die obige Funktion `commutativity` verwenden könnten: Das Problem ist aber, dass sich die Kommutativität von `and` an zwei weiteren Stellen von der Kommutativität für `*` und `+` unterscheidet, nämlich bei der Signatur (`boolean` statt `rational`) und auch bei der Vergleichsoperation (`boolean=?` statt `=`). Um auch `and` in den Einzugsbereich von `commutativity` zu holen, müssen wir also auch noch über diese beiden Werte abstrahieren:

```

(define commutativity
  (lambda (op sig =?)
    (for-all ((a sig)
              (b sig))
      (=? (op a b)
          (op b a)))))

```

Für `*` und `+` müssen wir `commutativity` nun wie folgt aufrufen:

```

(check-property (commutativity * (signature rational) =))
(check-property (commutativity + (signature rational) =))

```

Denk an das `signature`, das immer notwendig ist, wenn eine Signatur außerhalb einer Signatur-Deklaration mit `:` sowie einem `for-all` vorkommt.

Um `commutativity` auch auf `and` und `or` loszulassen, gibt es allerdings noch ein weiteres Hindernis: Das Argument zu `op` muss eine Funktion sein – `and` und `or` sind aber Spezialformen. Wir können sie aber zu Funktionen machen, indem wir `lambdas` darumwickeln:

```

(check-property (commutativity (lambda (a b) (and a b))
                              (signature boolean) boolean=?))
(check-property (commutativity (lambda (a b) (or a b))
                              (signature boolean)
                              boolean=?))

```

Bei der neuen Version von `commutativity` fehlt noch die Signatur. Wir müssen dazu die ursprüngliche Signatur



```
(: commutativity ((rational rational -> rational) -> property))
```

ziemlich radikal renovieren: Das erste Argument ist zwar immer noch eine zweistellige Funktion, aber nicht mehr notwendigerweise auf rationalen Zahlen. Wir skizzieren erstmal, was wir wissen:

```
(: commutativity ((? ? -> ?) signature (? ? -> boolean) -> property))
```

Die eingebaute Signatur `signature` ist für Signaturen zuständig – das zweite Argument ist ja eine Signatur. Von der Vergleichsfunktion an dritter Stelle ist klar, dass sie ein `boolean` liefert. Für die restlichen Fragezeichen ist die genaue Signatur abhängig vom konkreten Operator und dieser (ebenfalls variablen) Signatur, wir müssen also Signaturvariablen verwenden.

Was ist noch bekannt? Die beiden Argumente der Funktion `op` müssen auf dieselbe Signatur passen, da sie ja vertauschbar sind:

```
(: commutativity ((%a %a -> ?) signature (? ? -> boolean) ->
  property))
```

Außerdem wird der Rückgabewert von `op` in die Vergleichsfunktion gefüttert, für die restliche drei Fragezeichen müssen wir also dieselbe Signatur einsetzen. Ist erforderlich, dass der Rückgabewert von `op` auf die gleiche Signatur passt wie die Argumente? Der Rückgabewert wird nicht wieder in `op` hineingefüttert, die Antwort ist also nein. Wir können also eine von `%a` verschiedene Signaturvariable benutzen:

```
(: commutativity ((%a %a -> %b) signature (%b %b -> boolean) ->
  property))
```

Genauso wie bei der Kommutativität können wir auch bei der Assoziativität abstrahieren. Hier die Abstraktion, die dabei herauskommt:

```
(define associativity
  (lambda (op sig =?)
    (for-all ((a sig)
              (b sig)
              (c sig))
      (=? (op a (op b c))
          (op (op a b) c)))))
```

Benutzen können wir sie ähnlich wie bei der Kommutativität:

```
(check-property (associativity + (signature rational) =))
(check-property (associativity * (signature rational) =))
(check-property (associativity (lambda (a b) (and a b))
```

```

(signature boolean)
boolean=?))
(check-property (associativity (lambda (a b) (or a b))
(signature boolean)
boolean=?))

```

Auch hier die Formulierung der Signatur nicht so einfach. Die erste Skizze könnte so aussehen:

```
(: associativity ((? ? -> ?) signature (? ? -> boolean) -> property))
```

Wie bei `commutativity` wird der Rückgabewert von `op` als Argument für die Vergleichsfunktion verwendet: Die letzten drei Fragezeichen müssen also wieder gleich sein. Anders als bei der Kommutativität wird der Rückgabewert von `op` auch wieder als Argument in `op` hereingefüttert. Damit müssen auch die ersten beiden Fragezeichen den anderen entsprechen. Die beste Signatur ist also wie folgt:

```
(: associativity ((%a %a -> %a) signature (%a %a -> boolean) ->
property))
```

And und or erfüllen auch zwei Distributivgesetze. Damit beschäftigt sich Aufgabe 13.15.

Auch das DeMorgan'sche Gesetz (siehe Abschnitt A.1) können wir als Code formulieren:

```
(check-property
(for-all ((a boolean)
(b boolean))
(boolean=? (not (and a b))
(or (not a) (not b))))))

```

Bei vielen Operationen ist außerdem interessant, ob sie ein *neutrales Element* besitzen, also ein Argument, das dafür sorgt, dass die Operation ein anderes Argument unverändert zurückgibt. Die Addition hat z.B. die 0 als neutrales Element:

```
(check-property
(for-all ((a rational))
(= (+ a 0) a)))

```

Streng genommen ist damit nur gesichert, dass 0 *rechtsneutrales Element* ist, also von rechts addiert das andere Argument unverändert herauskommt. Aus der Kommutativität folgt aber, dass jedes rechtsneutrale Element auch ein linksneutrales Element ist.

Bei manchen Operationen gibt es neben dem neutralen Element zu jedem Element auch ein *inverses Element*: Wenn eine binäre Operation auf ein Element und sein Inverses angewendet wird,

so muss das neutrale Element herauskommen. Bei der Addition entsteht das Inverse zu einer Zahl durch Umdrehen des Vorzeichens:

```
(check-property
  (for-all ((a rational))
    (= (+ a (- a)) 0)))
```

```
(check-property
  (for-all ((a rational))
    (= (+ (- a) a) 0)))
```

Hier noch einmal eine Zusammenfassung der in diesem Abschnitt behandelten Eigenschaften, mit Kurzfassungen der mathematischen Formulierungen:

#### MANTRA 9 (EIGENSCHAFTEN VON BINÄREN OPERATIONEN)

Folgende Eigenschaften sind prinzipiell auf alle *binären* Operationen denkbar, die zwei Elemente einer Menge  $M$  akzeptieren und wiederum ein Element von  $M$  zurückgeben.

- Kommutativität  $a \star b = b \star a$
- Assoziativität  $(a \star b) \star c = a \star (b \star c)$
- Distributivität  $a \otimes (b \star c) = (a \otimes b) \star (a \otimes c)$ ;  $(b \star c) \otimes a = (b \otimes a) \star (c \otimes a)$
- neutrales Element  $(a \star \nu = a; \nu \star a = a)$
- inverses Element  $a \star a^{-1} = \nu; a^{-1} \star a = \nu$

### 13.7.2 EIGENSCHAFTEN VON BINÄREN PRÄDIKATEN

#### FIXME: Begriff Prädikat

Die Funktion `=` passt nicht in das Schema der Eigenschaften des folgenden Abschnitts. Sie hat folgende Signatur:

```
(: = (number number -> boolean))
```

Damit akzeptiert sie zwar zwei Argumente aus derselben Menge, liefert aber einen booleschen Wert zurück. Stattdessen handelt es sich um ein *binäres Prädikat* beziehungsweise eine *binäre Relation*. Für binäre Relationen kommt ein anderer Satz von Eigenschaften in Frage. (Die mathematische Seite ist in Anhang A.5 beschrieben.) Insbesondere ist `=` eine *Äquivalenzrelation* und damit *reflexiv*, *symmetrisch* und *transitiv*.

Die Reflexivität besagt, dass jedes Element der Grundmenge (in diesem Fall die Menge der Zahlen) zu sich selbst in Beziehung steht:

Eine *Implikation* in einer Eigenschaft wird folgendermaßen geschrieben:

`(==> e ep)`

Dabei muss  $e$  ein Ausdruck mit booleschem Wert sein (die *Voraussetzung*) und  $e_p$  eine Eigenschaft oder ein boolescher Ausdruck. Die Implikation liefert ihrerseits wieder eine Eigenschaft, die gilt, wenn  $e_p$  immer dann gilt, wenn die Voraussetzung erfüllt ist, also `#t` liefert.

**Abbildung 13.4:** `==>`

```
(check-property
 (for-all ((a number)
           (= a a)))
```

Die Symmetrie bedeutet für `=`, dass aus `(= a b) |evalsto| #t` das «Spiegelbild» `(= b a) |evalsto| #t` folgt. Mathematisch geschrieben sähe das so aus:

$$\forall a \in \mathbb{C}, b \in \mathbb{C} : a = b \Rightarrow b = a$$

Der Implikationspfeil  $\Rightarrow$  wird in den Lehrsprachen `==>` geschrieben. (Siehe Abbildung 13.4.)

Der Test der Symmetrie sieht also folgendermaßen aus:

```
(check-property
 (for-all ((a number)
           (b number))
  (==> (= a b)
        (= b a))))
```

Ähnlich läuft es mit der Transitivität: Wenn zwei Zahlen  $a$  und  $b$  gleich sind sowie  $b$  und eine dritte Zahl  $c$ , dann müssen auch  $a$  und  $c$  gleich sein:

```
(check-property
 (for-all ((a number)
           (b number)
           (c number))
  (==> (and (= a b) (= b c))
        (= a c))))
```

Neben den drei Eigenschaften von Äquivalenzrelationen tritt auch gelegentlich die Eigenschaft *Antisymmetrie* auf (die mathematische Definition steht in Anhang A.5).

Folgende Eigenschaften sind für binäre Prädikate denkbar:

- Reflexivität  $a \rightsquigarrow a$
- Symmetrie  $a \rightsquigarrow b \Rightarrow b \rightsquigarrow a$
- Transitivität  $a \rightsquigarrow b \wedge b \rightsquigarrow c \Rightarrow a \rightsquigarrow c$
- Antisymmetrie  $a \rightsquigarrow b \wedge b \rightsquigarrow a \Rightarrow a = b$

## 13.8 EIGENSCHAFTEN VON FUNKTIONEN AUF LISTEN

Es wird Zeit, Eigenschaften von selbstgeschriebenen Funktionen zu überprüfen. In diesem Abschnitt geht es um einige der Funktionen, die auf Listen operieren: `concatenate`, `invert`, und `list-sum`.

### 13.8.1 EIGENSCHAFTEN VON `CONCATENATE`

Die Funktion `concatenate` aus Abschnitt 6.6.3 hängt zwei Listen aneinander. Auch `concatenate` ist assoziativ: Wenn drei Listen mit Hilfe von `concatenate` aneinandergehängt werden, spielt es keine Rolle, ob zuerst die ersten beiden oder zuerst die letzten beiden Listen aneinandergehängt werden. Nach dem Muster der Assoziativität von `+` und `and` sieht der Test dieser Eigenschaft folgendermaßen aus:

```
(check-property
  (associativity concatenate (signature (list-of number)) ...))
```

Beim Test ist die Signatur von `lis-1`, `lis-2` und `lis-3` jeweils `(list-of number)`. Die Signatur von `concatenate`

```
(: concatenate ((list-of %a) (list-of %a) -> (list-of %a)))
```

suggeriert allerdings, dass die Signatur von `lis-1`, `lis-2` und `lis-3` jeweils `(list-of \%a)` lauten sollte, also allgemeiner als `(list-of number)`. Signaturen mit Signaturvariablen funktionieren allerdings nicht im Zusammenhang mit Eigenschaften, wie folgendes Beispiel zeigt:

```
(check-property
  (for-all ((x %a))
    ...))
```

Dieser Code liefert die Fehlermeldung «Signatur hat keinen Generator»: Das liegt daran, dass die Signaturvariable `%a` zuwenig Information über die zugrundeliegenden Werte liefert, als dass DrRacket sinnvoll Werte für die Tests generieren könnte. Aus diesem Grund müssen in `for-all` immer «konkrete» Signaturen ohne Signaturvariablen angegeben werden. (Aus ähnlichen Gründen funktionieren auch einige andere Arten von Signaturen nicht bei `for-all`, insbesondere Record-Signaturen. Funktionssignaturen sind allerdings zulässig und werden in Abschnitt 13.9 behandelt.)

Für `concatenate` wäre es zwar gründlicher, die Tests auch noch für andere Sorten von Listenelementen als `number` durchzuführen – da aber `concatenate` mit den Listenelementen nichts anfängt, außer sie in weitere Liste zu stecken, reicht die Formulierung der Eigenschaft mit `(list-of number)` aus.

Es bleibt noch ein weiteres Problem bei der Formulierung der Assoziativität für `concatenate`: Es steht noch keine Funktion für den Vergleich der beiden Listen zur Verfügung, die muss erst noch geschrieben werden. Kurzbeschreibung und Signatur: `FIXME: über = abstrahieren`

`; Zwei Listen aus Zahlen vergleichen`

`(: number-list=? ((list-of number) (list-of number) -> boolean))`

Die Testfälle sollten insbesondere Listen unterschiedlicher Länge berücksichtigen:

```
(check-expect (number-list=? empty empty) #t)
(check-expect (number-list=? (list 1.0 2.0 3.0) (list 1.0 2.0 3.0)) #t)
(check-expect (number-list=? (list 1.0 2.0 3.0) (list 1.0 2.0)) #f)
(check-expect (number-list=? (list 1.0 2.0) (list 1.0 2.0 3.0)) #f)
(check-expect (number-list=? (list 1.0 2.0 3.0) (list 1.0 2.1 3.0)) #f)
```

Die Schablone für Listen als Eingabe, ausgewählt nach dem ersten Listenparameter `lis-1`, sieht so aus:

```
(define number-list=?
  (lambda (lis-1 lis-2)
    (cond
      ((empty? lis-1) ...)
      ((cons? lis-1) ... (first lis-1) ...
        ... (number-list=? (rest lis-1) ...) ...))))
```

Die Schablone für den zweiten Listenparameter `lis-2` wird in beide Zweige des `cond` eingesetzt:

```

(define number-list=?
  (lambda (lis-1 lis-2)
    (cond
      ((empty? lis-1)
       (cond
         ((empty? lis-2) ...)
         ((cons? lis-2)
          ... (first lis-2) ...
          ... (number-list=? ... (rest lis-2))))))
      ((cons? lis-1)
       ... (first lis-1) ...
       ... (number-list=? (rest lis-1) ...) ...
       (cond
         ((empty? lis-2) ...)
         ((cons? lis-2)
          ... (first lis-2) ...
          ... (number-list=? ... (rest lis-2)))))))

```

Es gibt also insgesamt vier Fälle bei den Verzweigungen:

- Im ersten Fall sind beide Listen leer, das Ergebnis ist also `#t`.
- Im zweiten Fall ist die erste Liste leer und die zweite nichtleer. Das Ergebnis ist also `#f` und die Schablonelemente sind überflüssig.
- Im dritten Fall ist die erste Liste nichtleer und die zweite leer. Das Ergebnis ist also wiederum `#f`.
- Im vierten Fall sind beide Listen nichtleer und in der Schablone stehen die jeweils ersten Elemente von `lis-1` und `lis-2`. Die beiden Listen sind nur gleich, wenn die beiden ersten Elemente gleich sind. Außerdem müssen natürlich die beiden Reste der Listen ebenfalls gleich sind – die beiden rekursiven Aufrufe aus den Schablonen können also kombiniert werden:

```

(define number-list=?
  (lambda (lis-1 lis-2)
    (cond
      ((empty? lis-1)
       (cond
         ((empty? lis-2) #t)
         ((cons? lis-2) #f)))
      ((cons? lis-1)

```

```
(cond
  ((empty? lis-2) #f)
  ((cons? lis-2)
   (and (= (first lis-1) (first lis-2))
        (number-list=? (rest lis-1) (rest lis-2))))))
```

Damit kann jetzt die Assoziativität von `concatenate` getestet werden:

```
(check-property
  (associativity concatenate (signature (list-of number)) number-list=?
  ))
```

`Concatenate` hat außerdem ein neutrales Element, und zwar sowohl im linken als auch im rechten Argument:

```
(check-property
  (for-all ((lis (list-of number)))
    (number-list=? lis (concatenate empty lis))))

(check-property
  (for-all ((lis (list-of number)))
    (number-list=? lis (concatenate lis empty))))
```

`Concatenate` ist allerdings demonstrierbar nicht kommutativ. Der entsprechende Test sieht so aus:

```
(check-property
  (commutativity concatenate (signature (list-of number)) number-list=?
  ))
```

DrRacket liefert hierfür ein Gegenbeispiel:

Eigenschaft falsifizierbar mit

```
lis-1 = #<list -3.75> lis-2 = #<list 1.5 1.5>
```

### 13.8.2 EIGENSCHAFTEN VON `NUMBER-LIST=?`

Wie der Zufall so will, hat auch die Hilfsfunktion `number-list=?` interessante Eigenschaften: Wie `=` muss auch `number-list=?` eine Äquivalenzrelation sein – schließlich testet sie wie `=` auf



Gleichheit. Die dazugehörigen Eigenschaften – Reflexivität, Symmetrie und Transitivität – können ebenso wie bei = formuliert werden:

Reflexivität:

```
(check-property
  (for-all ((lis (list-of number)))
    (number-list=? lis lis)))
```

Symmetrie:

```
(check-property
  (for-all ((lis-1 (list-of number))
            (lis-2 (list-of number)))
    (==> (number-list=? lis-1 lis-2)
          (number-list=? lis-2 lis-1))))
```

Transitivität

```
(check-property
  (for-all ((lis-1 (list-of number))
            (lis-2 (list-of number))
            (lis-3 (list-of number)))
    (==> (and (number-list=? lis-1 lis-2)
              (number-list=? lis-2 lis-3))
          (number-list=? lis-1 lis-3))))
```

### 13.8.3 EIGENSCHAFTEN VON INVERT

Die Funktion `invert` aus Abschnitt 10.1 dreht die Reihenfolge der Elemente einer Liste um. Eine naheliegende Eigenschaft von `invert` ist, dass zweimaliges Umdrehen wieder die Ursprungsliste liefern sollte:

```
(check-property
  (for-all ((lis (list-of number)))
    (number-list=? lis (invert (invert lis)))))
```

Auch bei `invert` enthält die Signatur eine Signaturvariable:

```
(: invert ((list-of %a) -> (list-of %a)))
```

**Expect** liefert eine Eigenschaft analog zur Funktionsweise von **check-expect**. Ein **expect**-Ausdruck hat folgende Form:

```
(expect e1 e2)
```

$e_1$  und  $e_2$  sind Ausdrücke. Die resultierende Eigenschaft ist erfüllt, wenn  $e_1$  und  $e_2$  den gleichen Wert liefern – der Vergleich wird dabei wie bei **check-expect** angestellt.

**Abbildung 13.5:** **expect**

Genau wie bei **concatenate** macht **invert** mit den Listenelementen nichts spezielles, es können also auch zum Beispiel Zeichenketten benutzt werden. Diese Änderung allein funktioniert allerdings nicht:

```
(check-property
  (for-all ((lis (list-of string)))
    (\textbf{number-list=?} lis (invert (invert lis)))))
```

Die Funktion **number-list=?** funktioniert nur auf Listen von Zahlen. Es wäre möglich, **number-list=?** über der Vergleichsfunktion auf den Elementen zu abstrahieren, aber es wäre trotzdem umständliche Arbeit nur für den Zweck des Testens. Deshalb gibt es eine Vereinfachung analog zu **check-expect**. Die eingebaute Form **expect** akzeptiert zwei beliebige Werte und ist dann erfüllt, wenn diese Werte gleich sind. (Siehe Abbildung 13.5.) Die Eigenschaft von **invert** sieht damit so aus:

```
(check-property
  (for-all ((lis (list-of string)))
    (expect lis (invert (invert lis)))))
```

Viele Funktionen auf Listen haben Eigenschaften, welche die Funktion jeweils im Zusammenspiel mit einer oder mehreren anderen Funktionen zeigen. Bei Funktionen mit Listen ist es häufig interessant, das Zusammenspiel mit **concatenate** zu betrachten. Damit **concatenate** etwas sinnvolles tun kann, sind zwei Listen notwendig:

```
(check-property
  (for-all ((lis-1 (list-of number))
            (lis-2 (list-of number))
            ...))
```

Auf diese zwei Listen kann **concatenate** aber auch jeweils **invert** angewendet werden:

```
(check-property
  (for-all ((lis-1 (list-of number))
```

```

      (lis-2 (list-of number)))
... (invert lis-1) ...
... (invert lis-2) ...
... (invert (concatenate lis-1 lis-2)) ...)

```

Wie lässt sich die Liste `(invert (concatenate lis-1 lis-2))` noch beschreiben? Angenommen, `lis-1` ist die Liste `#<list 1 2 3>` und `lis-2` die Liste `#<list 4 5 6>`. Dann gilt:

```

(invert (concatenate lis-1 lis-2))
=
(invert (concatenate #<list 1 2 3> #<list 4 5 6>))
⇒ ... ⇒ (invert #<list 1 2 3 4 5 6>))
                        lis-1   lis-2

⇒ ... ⇒ #<list   6 5 4   3 2 1   >
                        (invert lis-2) (invert lis-1)

```

Dies lässt vermuten, dass die gesuchte Eigenschaft folgendermaßen aussieht:

```

(check-property
  (for-all ((lis-1 (list-of number))
            (lis-2 (list-of number)))
    (expect (invert (concatenate lis-1 lis-2))
            (concatenate (invert lis-2) (invert lis-1))))))

```

## MANTRA 11 (EIGENSCHAFTEN VON FUNKTIONEN AUF LISTEN)

Funktionen, die Listen akzeptieren, haben häufig interessante Eigenschaften im Zusammenspiel mit `concatenate`.

### 13.8.4 EIGENSCHAFTEN VON `LIST-SUM`

`List-sum` aus Abschnitt 6.1 ist, wie `invert`, eine Funktion, die eine Liste akzeptiert. Genau wie bei `invert` ist es eine gute Idee, die Interaktion zwischen `list-sum` und `concatenate` zu untersuchen. Es müssen also wieder zwei Listen her – die zu `invert` analoge Vorgehensweise liefert folgende Schablone:

```

(check-property
  (for-all ((lis-1 (list-of number))

```

`Expect-within` liefert eine Eigenschaft analog zur Funktionsweise von `check-within`. Ein `expect-within`-Ausdruck hat folgende Form:

```
(expect-within  $e_1$   $e_2$   $e_\delta$ )
```

$e_1, e_2$  und  $e_\delta$  sind Ausdrücke, wobei  $e_\delta$  eine reelle Zahl liefern muss. Die resultierende Eigenschaft ist erfüllt, wenn  $e_1$  und  $e_2$  den gleichen Wert liefern – der Vergleich wird dabei wie bei `check-within` angestellt, das heißt alle inexakten Zahlen in den Ergebnissen von  $e_1$  und  $e_2$  müssen nicht gleich sein, dürfen sich aber höchstens um  $e_\delta$  voneinander unterscheiden.

Abbildung 13.6: `expect-within`

```
(lis-2 (list-of number)))
... (list-sum lis-1) ...
... (list-sum lis-2) ...
... (list-sum (concatenate lis-1 lis-2)) ...
```

Da `list-sum` die Elemente der Liste addiert und die Addition assoziativ ist, müsste folgendes gelten:

```
(check-property
  (for-all ((lis-1 (list-of number))
            (lis-2 (list-of number)))
    (expect (+ (list-sum lis-1) (list-sum lis-2))
            (list-sum (concatenate lis-1 lis-2)))))
```

Hier allerdings schlägt das Rundungsproblem aus Abschnitt 13.7.1 wieder zu: Die Addition auf `number` ist eben *nicht* assoziativ, aber immerhin auf `rational`. Der fertige Test muss also so aussehen:

```
(check-property
  (for-all ((lis-1 (list-of rational))
            (lis-2 (list-of rational)))
    (expect (+ (list-sum lis-1) (list-sum lis-2))
            (list-sum (concatenate lis-1 lis-2)))))
```

Eine Alternative ist die Verwendung der Form `expect-within`, die eine Eigenschaft analog zu `check-within` erzeugt. (Siehe Abbildung 13.6.)

Mit `expect-within` sieht der Testfall folgendermaßen aus:

```
(check-property
  (for-all ((lis-1 (list-of number))
            (lis-2 (list-of number)))
```

```
(expect-within (+ (list-sum lis-1) (list-sum lis-2))
               (list-sum (concatenate lis-1 lis-2))
               0.1)))
```

Auch dieser Testfall läuft durch.

Auch der Test für die Assoziativität von `+` aus Abschnitt 13.7.1 kann mit `expect-within` formuliert werden:

```
(check-property
  (for-all ((a number)
            (b number)
            (c number))
    (expect-within (+ a (+ b c))
                  (+ (+ a b) c)
                  0.1)))
```

So wie sich die Assoziativität von `+` in einer Eigenschaft von `list-sum` niederschlägt, tut dies auch die Kommutativität: Sie besagt, dass die Reihenfolge der Elemente der Liste keine Rolle spielt. Eine einfache Möglichkeit, dies zu testen, ist wiederum mit zwei Listen zu arbeiten und diese einmal in der einen und dann in der anderen Richtung aneinanderzuhängen:

```
(check-property
  (for-all ((lis-1 (list-of rational))
            (lis-2 (list-of rational)))
    (expect (list-sum (concatenate lis-1 lis-2))
            (list-sum (concatenate lis-2 lis-1))))))
```

## 13.9 EIGENSCHAFTEN VON FUNKTIONEN HÖHERER ORDNUNG

In Abschnitt 9.5 wurde bereits eine Eigenschaft von `curry` und `uncurry` aufgeführt:

$$(\text{uncurry } (\text{curry } p)) \equiv p$$

Mit anderen Worten: `curry` und `uncurry` sind jeweils Inverse voneinander. Auch diese Eigenschaft lässt sich direkt mit `check-property` und `for-all` formulieren. Zu beachten ist wieder, obwohl `curry` und `uncurry` polymorphe Signaturen mit Signaturvariablen haben, dass für den Test mit `check-property` eine «konkrete» Signatur ohne Signaturvariablen für das Funktions-Argument benutzt werden muss, also zum Beispiel `string`:

```
(check-property
  (for-all ((proc (string string -> string)))
    (expect (curry (uncurry proc))
      proc)))
```

Leider schlägt dieser Test fehl, und zwar mit einer mysteriösen Meldung:

Eigenschaft falsifizierbar mit #<procedure:??>

Offenbar ist DrRacket also der Ansicht, es hat eine Funktion gefunden, welche die Eigenschaft nicht erfüllt, kann aber nicht genau sagen, welche Funktion: Das liegt daran, dass es prinzipiell unmöglich ist, zwei Funktionen auf Gleichheit zu überprüfen – Gleichheit zweier Funktionen heißt ja, dass die eine Funktion angewendet auf einen Satz Argumente *immer* das gleiche Ergebnis wie die andere Funktion liefert. Im obigen Beispiel akzeptiert `proc` zwei Zeichenketten, von denen es unendlich viele gibt; die Gleichheit zu überprüfen, würde also unendlich lange dauern. `Expect` versucht es darum gar nicht erst, sondern sieht es als notwendige (nicht als hinreichende) Bedingung für die Gleichheit zweier Funktionen an, dass sie Wert desselben `lambda`-Ausdrucks sind. `Expect` testet also bei Funktionen auf sogenannte *intensionale Gleichheit*, was soviel heißt, dass `expect` vergleicht, *auf welche Weise* die beiden Funktionen entstanden sind, nicht aber, ob sich die beiden Funktionen *gleich verhalten*. Die letztere Eigenschaft heißt *extensionale Gleichheit* – und ist, wie gesagt, nicht effektiv testbar.

Der `lambda`-Ausdruck der Funktion, die von `(curry (uncurry proc))` geliefert wird, ist aber der Rumpf von `curry`, während der `lambda`-Ausdruck von `proc` i.d.R. woanders steht; damit sind die beiden Funktionen intensional ungleich, und der obige Test muss fehlschlagen, auch wenn die beiden Operanden von `expect` äquivalent sind.

Damit ein `check-property`-Test die Äquivalenz testen kann, muss er selbst `(curry (uncurry proc))` anwenden:

```
(check-property
  (for-all ((a string)
            (b string)
            (proc (string string -> string)))
    (expect ((uncurry (curry proc)) a b)
      (proc a b))))
```

Dieser Test funktioniert.

## 13.10 PROGRAMME BEWEISEN

**Check-property** ist nützlich, um zu überprüfen, ob eine Eigenschaft gilt oder nicht. Da **check-property** allerdings nur eine endliche Menge zufälliger Tests durchführt, reicht es nicht aus, um sicherzugehen, dass eine bestimmte Eigenschaft für alle Werte der **for-all**-Variablen gilt: Dazu ist ein mathematischer Beweis notwendig.

An verschiedenen Stellen im Buch wurden Beweise für mathematische Funktionen durchgeführt – zuletzt in Kapitel 8 für eine rekursive Funktion. Beweise über mathematische Funktionen erlauben, bei jedem Schrittan beliebigen Stellen Gleichungen einzusetzen. Beweise über Funktionen in Programmen sind schwieriger, da sie das Substitutionsmodell berücksichtigen müssen: Bei jedem Reduktionsschritt kommt immer nur eine bestimmte Substitution in Frage.

### 13.10.1 ARITHMETISCHE GESETZE BEWEISEN

Als erstes Beispiel für den Beweis an einem Programm dient der Beweis der Kommutativität von  $+$ . Der Beweis ist nicht besonders tiefgreifend, demonstriert aber die wichtigsten Techniken, die beim Beweisen von Programmen zum Einsatz kommen. Zu beweisen ist:

```
(= (+ a b) (+ b a))  
⇒ ... ⇒ #t
```

... und zwar für beliebige Bindungen von  $a$  und  $b$  an Zahlen. Seien die Zahlen, die an  $a$  beziehungsweise  $b$  gebunden sind,  $x$  und  $y$  mit  $x, y \in \mathbb{C}$ . (Die «mathematischen» Namen könnten auch  $a$  und  $b$  sein, aber das birgt ein Verwirrungsrisiko mit  $a$  und  $b$ .) Wenn nun also der obige Term für bestimmte Werte von  $x$  und  $y$  im Substitutionsmodell reduziert wird, wird zuerst  $x$  für  $a$  und  $y$  für  $b$  eingesetzt. Für  $x = 5$  und  $y = 17$  also:

```
(= (+ 5 17) (+ 17 5))
```

FIXME:  $\lceil \_ \rceil$  schwierig

Der Beweis soll aber für beliebige Werte für  $x$  und  $y$  funktionieren:  $x$  und  $y$  müssen also im Beweis auftauchen. Um den Unterschied zwischen Variablen des Programms  $a$  und  $b$  und den Zahlen zu machen, die für  $x$  und  $y$  eingesetzt werden, werden diese noch mit  $\lceil \_ \rceil$  umgeben:  $\lceil x \rceil$  in einem Reduktionsschritt des Substitutionsmodell ist also ein Platzhalter für «die Zahl, für die  $x$  steht» – entsprechend für  $y$ . Es gilt also:

```
(= (+ a b) (+ b a))  
= (= (+  $\lceil x \rceil$   $\lceil y \rceil$ ) (+  $\lceil y \rceil$   $\lceil x \rceil$ ))
```

Dort ist der erste Teilausdruck unterstrichen, der beim ersten Substitutionsschritt ersetzt wird. Wenn die Programm-Funktion `+` tatsächlich die mathematische Operation `+` realisiert,<sup>5</sup> wird der Teilausdruck `(+ [x] [y])` durch  $x + y$  ersetzt – beziehungsweise durch die Zahl, für die der mathematische Ausdruck  $x + y$  steht. Es kommt also wieder `[_]` zum Einsatz:

```
(= (+ [x] [y]) (+ [y] [x]))
⇒ (= [x + y] (+ [y] [x]))
```

Entsprechend geht es weiter mit der zweiten Summe und schließlich der Vergleichsoperation `=`, die dem mathematischen `=` entspricht:

```
⇒ (= [x + y] [y + x])
⇒ [x + y = y + x]
= #t
```

Die Kommutativität der Programm-Funktion `+` folgt also aus der Kommutativität des mathematischen `+`, durch das sie definiert ist.

## 13.11 REKURSIVE PROGRAMME BEWEISEN

Beweise über rekursive Programme sind anspruchsvoller als der Beweis der Kommutativität von `+`, benutzen aber die gleichen Techniken sowie – genau wie bei Beweisen mathematischer rekursiver Funktionen – Induktion als Beweisprinzip.

### 13.11.1 REKURSIVE PROGRAMME AUF LISTEN

Als erstes Beispiel dient die Reflexivität. Es gilt für die Bindung von `lis` an eine beliebige Liste von Zahlen folgendes zu beweisen:

```
(number-list=? lis lis)
⇒ ... ⇒ #t
```

Wieder wird für den Wert der Bindung eine mathematische Variable eingeführt – *l*. Dann läuft der Beweis auf folgendes hinaus:

<sup>5</sup> Die Komplikationen durch inexacte Zahlen und Rundungen bleiben hier unberücksichtigt.



```

(number-list=? lis lis)
= (number-list=? [l] [l])
⇒ ... ⇒ #t

```

... und dies ist zu beweisen für *alle* Listen von Zahlen  $l$ . Für diesen Beweis kommt uns die induktive Struktur der Listen zur Hilfe, die der Struktur der endlichen Folgen – bekannt aus Abschnitt 8.4 auf Seite 221 – entspricht. Dazu müssen wir zunächst einmal zwischen allen Fällen der gemischten Datendefinition unterscheiden, also zwischen den leeren und den nichtleeren Listen.

**Leere Liste** | Angenommen,  $l$  ist die leere Liste. Dann beginnt die Reduktion folgendermaßen:

```

(number-list=? [l] [l])
⇒ ((lambda (lis-1 lis-2) ...) [l] [l])
⇒ (cond ((empty? [l]) ...) ((cons? [l]) ...))
⇒ (cond ((empty? [l]) ...) ((cons? [l]) ...))
⇒ (cond (#t (cond ...)) ((cons? [l]) #f))
⇒ (cond ((empty? [l]) #t) ((cons? [l]) #f))
⇒ (cond (#t #t) ((cons? [l]) #f))
⇒ #t

```

**Nichtleere Liste** | Für diesen Fall stimmt die Behauptung also. Angenommen,  $l$  ist *nicht* die leere Liste, hat also erstes Element  $f$  und Rest  $r$ . In diesem Fall können wir strukturelle Induktion benutzen. Die *Induktionsvoraussetzung* bezieht sich dann auf den Rest  $r$ :

```

(number-list=? [r] [r])
⇒ ... ⇒ #t

```

Diese können wir benutzen. Zunächst einmal müssen wir aber soweit reduzieren wie möglich:

```

(number-list=? [l] [l])
⇒ ((lambda (lis-1 lis-2) ...) [l] [l])
⇒ (cond ((empty? [l]) ...) ((cons? [l]) ...))
⇒ (cond (#f ...) ((cons? [l]) ...))
⇒ (cond (#t (cond ...)))
⇒ (cond ((empty? [l]) ...) ((cons? [l]) ...))
⇒ (cond (#f ...) ((cons? [l]) ...))
⇒ (cond (#t (and ...)))
⇒ (and (= (first [l]) (first [l])) (number-list=? ...))

```

Da `(first [l])` das erste Element  $f$  liefert, geht es so weiter:

```
⇒ (and (= [f] (first [l])) (number-list=? ...))
⇒ (and (= [f] [f]) (number-list=? ...))
⇒ (and #t (number-list=? ...))
⇒ (number-list=? (rest [l]) (rest [l]))
```

Da `(rest [l])` den Rest  $r$  liefert, geht es dann so weiter:

```
⇒ (number-list=? [r] (rest [l]))
⇒ (number-list=? [r] [r])
```

Nach der Induktionsvoraussetzung wissen wir aber, dass der letzte Ausdruck zu `#t` reduziert wird. Die Behauptung ist damit bewiesen.

### 13.11.2 REKURSIVE PROGRAMME AUF ZAHLEN

Die Definition von `factorial` am Anfang von Abschnitt ?? folgt der induktiven Definition der zugrundeliegenden Daten, der natürlichen Zahlen. Dementsprechend ist der Induktionsbeweis für dessen Korrektheit einfach. Es ist aber entscheidend, die zu beweisende Eigenschaft, welche die Korrektheit von `factorial` begründet, sorgfältig aufzuschreiben:

Die Funktion `factorial` soll die Fakultät berechnen, es soll also für alle natürlichen Zahlen  $k$  gelten:

```
(factorial [k])
⇒ ... ⇒ [k!]
```

(Diese Eigenschaft lässt sich nicht sinnvoll mit `for-all` hinschreiben, da die mathematische Fakultät nicht fest eingebaut ist.)

Da es um natürliche Zahlen geht, ist vollständige Induktion anwendbar. Wir verwenden das Schema aus Abschnitt 8.2 auf Seite 217:

1. Die Behauptung ist bereits in der geforderten Form.
2.  $k = 0$ :

```
(factorial 0)
⇒ [0!]
```

3. Beweis für  $k = 0$ :

```
(factorial 0)
⇒ ((lambda (n) ...) 0)
```

```

⇒ (if (= 0 0) ...)
⇒ (if #t 1 ...)
⇒ 1
= [0!]

```

4. Induktionsvoraussetzung:

```

(factorial [k])
⇒ ... ⇒ [k!]

```

5. Induktionsschluss (zu zeigen):

```

(factorial [k + 1])
⇒ ... ⇒ [(k + 1)!]

```

Der Beweis sieht so aus:

```

(factorial [k + 1])
⇒ ((lambda (n) ...) [k + 1])
⇒ (if (= [k + 1] 0) ...)
⇒ (if #f 1 (* ...))
⇒ (* [k + 1] (factorial (- [k + 1] 1)))
⇒ (* [k + 1] (factorial [k]))
= (* [k + 1] (factorial [k]))

```

Mit der Induktionsannahme kann `(factorial [k])` ersetzt werden:

```

(* [k] (factorial [l]))
⇒ ... ⇒ (* [k + 1] [k!])
⇒ [(k + 1) × k!]
= [(k + 1)!]

```

Damit ist der Beweis fertig.

Die Technik funktioniert auch mit Beispielen, bei denen die zu beweisende Eigenschaft nicht so einfach zu sehen ist wie bei `factorial`.

Die folgende Funktion verrät nicht auf den ersten Blick, was sie berechnet:

```

(: f (natural -> rational))

```

```

(define f
  (lambda (n)
    (if (= n 0)

```

```

0
(+ (f (- n 1))
  (/ 1 (* n (+ n 1))))))

```

Tatsächlich berechnet der Funktionsaufruf  $(f \lceil k \rceil)$  für eine natürliche Zahl  $k$  die Zahl  $\frac{k}{k+1}$ .

Die Eigenschaft ist plausibel, wie sich mit `check-property` feststellen lässt:

```

(check-property
 (for-all ((k natural))
  (= (f k) (/ k (+ k 1)))))

```

Ein Beweis schafft Sicherheit. Wieder gehen wir nach dem bekannten Schema vor:

6. Behauptung:

Für alle  $k \in \mathbb{N}$  gilt:

```

(f  $\lceil k \rceil$ )
 $\implies \dots \implies \lceil \frac{k}{k+1} \rceil$ 

```

2.  $k = 0$ :

```

(f  $\lceil 0 \rceil$ )
 $\implies \dots \implies \lceil \frac{0}{0+1} \rceil$ 

```

3. Beweis:

```

(f  $\lceil 0 \rceil$ )
 $\implies ((\text{lambda } (n) \dots) 0)$ 
 $\implies (\text{if } (= 0 0) 0 \dots)$ 
 $\implies (\text{if } \#t 0 \dots)$ 
 $\implies 0$ 
 $= \lceil \frac{0}{0+1} \rceil$ 

```

4. Induktionsvoraussetzung:

```

(f  $\lceil k \rceil$ )
 $\implies \dots \implies \lceil \frac{k}{k+1} \rceil$ 

```

5. Induktionsschluss (zu zeigen):

```

(f  $\lceil k+1 \rceil$ )
 $\implies \dots \implies \lceil \frac{k+1}{(k+1)+1} \rceil$ 

```

6. Beweis:

```

(f ⌈k + 1⌉)
⇒ ((lambda (n) (if ...)) ⌈k + 1⌉)
⇒ (if (= ⌈k + 1⌉ 0) ...)
⇒ (if #f ... (+ ...))
⇒ (+ (f (- ⌈k + 1⌉ 1)) (/ 1 (* ⌈k + 1⌉ (+ ⌈k + 1⌉ 1)))))
⇒ (+ (f ⌈k⌉) (/ 1 (* ⌈k + 1⌉ (+ ⌈k + 1⌉ 1)))))
⇒ ... ⇒ (+ ⌈ $\frac{k}{k+1}$ ⌉ (/ 1 (* ⌈k + 1⌉ (+ ⌈k + 1⌉ 1)))))
⇒ (+ ⌈ $\frac{k}{k+1}$ ⌉ (/ 1 (* ⌈k + 1⌉ ⌈(k + 1) + 1⌉))))
⇒ (+ ⌈ $\frac{k}{k+1}$ ⌉ (/ 1 ⌈(k + 1) × ((k + 1) + 1)⌉)))
⇒ (+ ⌈ $\frac{k}{k+1}$ ⌉ ⌈ $\frac{1}{(k+1) \times ((k+1)+1)}$ ⌉))
⇒ ⌈ $\frac{k}{k+1} + \frac{1}{(k+1) \times ((k+1)+1)}$ ⌉
= ⌈ $\frac{k}{k+1} + \frac{1}{(k+1) \times (k+2)}$ ⌉
= ⌈ $\frac{k \times (k+2) + 1}{(k+1) \times (k+2)}$ ⌉
= ⌈ $\frac{k^2 + 2k + 1}{(k+1) \times (k+2)}$ ⌉
= ⌈ $\frac{(k+1)^2}{(k+1) \times (k+2)}$ ⌉
= ⌈ $\frac{k+1}{k+2}$ ⌉
= ⌈ $\frac{k+1}{(k+1)+1}$ ⌉

```

Damit ist die Behauptung bewiesen.

## 13.12 INVARIANTEN

Die bisher angewendete Technik für den Beweis rekursiver Funktionen mit Induktion funktioniert bei Funktionen mit Akkumulator nicht mehr direkt: Angenommen, die Korrektheit der endrekursiven Fakultät `!` aus Abschnitt ?? soll ähnlich die die Korrektheit von `factorial` bewiesen werden. Wieder sei `n` an eine natürliche Zahl `k` gebunden:

```

(! n)
= (! ⌈k⌉)
⇒ ((lambda (n) (!-helper n 1)) ⌈k⌉)
⇒ (!-helper ⌈k⌉ 1)
⇒ ((lambda (n acc) ...) ⌈k⌉ 1)
⇒ (if (= ⌈k⌉ 0) 1 (!-helper (- ⌈k⌉ 1) (* 1 ⌈k⌉)))

```

Wie bei `factorial` muss zwischen  $k = 0$  und  $k > 0$  unterschieden werden. Für  $k = 0$  geht es folgendermaßen weiter:

```
⇒ (if (= [k] 0) 1 (!-helper (- [k] 1) (* 1 [k])))
⇒ (if #t 1 (!-helper (- [k] 1) (* 1 [k])))
⇒ 1
```

Für  $k = 0$  funktioniert also der Beweis. Für  $k > 0$  allerdings verläuft die Reduktion folgendermaßen:

```
⇒ (if (= [k] 0) 1 (!-helper (- [k] 1) (* 1 [k])))
⇒ (if #f 1 (!-helper (- [k] 1) (* 1 [k])))
⇒ (!-helper (- [k] 1) (* 1 [k]))
```

Hier gibt es zwar einen rekursiven Aufruf mit Argument  $(- [k] 1)$ , aber *der Akkumulator hat sich auch verändert*. Damit ist die naheliegende Induktionsannahme für  $(!-helper (- [k] 1) [a])$  (falls der Wert des Akkumulators `acc a` ist) wertlos. Funktionen mit Akkumulator sind also nicht nur schwieriger zu schreiben als «regulär» rekursive Funktionen – sie sind auch schwerer zu beweisen.

Stattdessen ist es bei Funktionen mit Akkumulator nützlich, eine *Invariante* aufzustellen, also eine Eigenschaft, welche Zwischenergebnis und noch zu leistende Arbeit in Beziehung setzt. Wie in Abschnitt ?? beschrieben, geht die Fakultätsfunktion mit Akkumulator folgendermaßen vor, um  $(! 4)$  auszurechnen:

$$(((1 \times 4) \times 3) \times 2) \times 1$$

Bei jedem rekursiven Aufruf lässt sich dieser Aufruf in «geleistete Arbeit» (die durch den Akkumulator repräsentiert ist) und «noch zu leistende Arbeit» unterteilen. Zum Beispiel entsteht ein rekursiver Aufruf  $(!-helper 2 12)$ , bei der Akkumulator der Wert des unterstrichenen Teilaufrufs ist:

$$(((1 \times 4) \times 3) \times 2) \times 1$$

Es ist zu sehen, dass die noch zu leistende Arbeit gerade darin besteht, den Akkumulator noch mit der Fakultät von 2 zu multiplizieren. Wenn bei einem rekursiven Aufruf von `!-helper` der Wert von  $n$   $k$  ist und der Wert des Akkumulators  $a$ , und am Ende die Fakultät von  $N$  berechnet werden soll, dann gilt bei jedem rekursiven Aufruf  $(!-helper [n] [a])$  immer:

$$a \times k! = N!$$

Dies ist die Invariante von `!-helper` und heißt so, weil sie beim rekursiven Aufruf von `!-helper` unverändert bleibt. Dies ist zunächst nur eine Behauptung, aber wenn sie gelten sollte, dann folgt

daraus automatisch die Korrektheit der Funktion, da bei  $k = 0$  gilt:

$$a \times 0! = a \times 1 = a = N!$$

Dass  $a \times k! = N!$  tatsächlich die Invariante ist, lässt sich folgendermaßen folgern:

- Sie gilt für den ersten Aufruf von `!-helper` von `!`, da dort  $k = N$  und  $a = 1$  gilt, also:

$$a \times k! = 1 \times N! = N!$$

- Jeder rekursive Aufruf erhält die Invariante. Angenommen, sie gilt für  $k$  und  $a$ , dann sind die neuen Werte für  $k$  und  $a$  im rekursiven Aufruf `(!-helper (- [n] 1) (* [a] [k]))` gerade  $k \mapsto k - 1$  und  $a \mapsto a \times n$ , die Invariante wäre also:

$$\begin{aligned} (a \times k) \times (k - 1)! &= a \times (k \times (k - 1)! \\ &= a \times k! \\ &= N! \end{aligned}$$

Diese Technik funktioniert auch bei weniger offensichtlichen Funktionen. Hier eine Funktion, die äquivalent ist zu der Funktion `f` aus Abschnitt 13.11:

```
; n/(n+1) berechnen
(: f (natural -> natural))

(define f
  (lambda (n)
    (f-helper n 0)))

(define f-helper
  (lambda (n acc)
    (if (= n 0)
        acc
        (f-helper (- n 1)
                    (+ (/ 1 (* n (+ n 1)))
                      acc)))))
```

Die Funktion geht folgendermaßen vor, um das Ergebnis für eine Eingabe  $n$  zu berechnen:

$$(\dots((\frac{1}{n \times (n+1)} + \frac{1}{(n-1) \times ((n-1)+1)}) + \frac{1}{(n-2) \times ((n-2)+1)}) + \dots + \frac{1}{1 \times (1+1)})$$

Diese Summe ist bei jedem rekursiven Aufruf aufgeteilt als Summe von zwei Teilen, z.B. wie folgt:

$$(\dots(\underbrace{(\frac{1}{n \times (n+1)} + \frac{1}{(n-1) \times ((n-1)+1)})}_{\text{Akkumulator}} + \frac{1}{(n-2) \times ((n-2)+1)}) + \dots + \frac{1}{1 \times (1+1)})$$

Der unterstrichene Teil ist gerade der Wert des Akkumulators, die Summe rechts davon der noch zu berechnende Summand. Wenn die Funktion tatsächlich  $n/(n+1)$  berechnen sollte, ist dieser rechte Teil im Beispiel  $(n-3)/((n-3)+1)$ . Damit ergibt sich die Invariante als  $a + n/(n+1)$ , wobei  $a$  der Wert von `acc` ist. Um die Annahme zu beweisen, dass dies die Invariante ist, muss im wesentlichen folgende Gleichung bewiesen werden:

$$a + \frac{n}{n+1} = a + \frac{n-1}{n} + \frac{1}{n \times (n+1)}$$

Dies ist eine lohnende Fingerübung.

## 13.13 KORREKTHEIT VON SUCHBÄUMEN BEWEISEN

### FIXME

Die beiden `check-property`-Tests stärken also das Vertrauen in die korrekte Funktionsweise von `search-tree-insert` und `search-tree-member?`. Aber auch hier ist Kontrolle über einen Beweis der Korrektheit noch besser: Es lohnt sich, etwas formaler über die Korrektheit von `search-tree-insert` nachzudenken. Zunächst einmal ist es wichtig, zu formulieren, was der Begriff «Korrektheit» im Zusammenhang mit `search-tree-insert` überhaupt bedeutet:

**SATZ 13.1** `Search-tree-insert` erhält die Suchbaumeigenschaft. Oder mit anderen Worten: Wenn das `search-tree`-Argument von `search-tree-insert` die Suchbaumeigenschaft erfüllt, so erfüllt auch der zurückgegebene Baum die Suchbaumeigenschaft.

**Beweis** Die Korrektheit ist an der Hilfsfunktion `insert` festgemacht: Wenn das Argument von `insert` die Suchbaumeigenschaft erfüllt, so muß auch der Rückgabewert sie erfüllen. Der Beweis funktioniert über strukturelle Induktion über den Wert `t`, der an den Baum `τ` gebunden sei. Im Beweis gibt es vier Fälle, die den Zweigen der `cond`-Formen entsprechen:



- $\tau$  ist der leere Baum. Der dann zurückgegebene Baum der Form  
TBD  
erfüllt offensichtlich die Suchbaumeigenschaft.
- $\tau$  ist ein Knoten, dessen Markierung mit  $l$  übereinstimmt. Dann gibt `insert`  $\tau$  zurück. Da  $\tau$  nach Voraussetzung die Suchbaumeigenschaft erfüllt, ist auch hier die Suchbaumeigenschaft erhalten.
- $\tau$  ist ein Knoten, dessen Markierung *größer* ist als  $l$ , sieht also so aus:  
TBD  
wobei sowohl  $a$  als auch  $b$  selbst die Suchbaumeigenschaft erfüllen. In diesem Fall sieht der entstehende Baum folgendermaßen aus:  
TBD  
Per Induktionsannahme erfüllt `(insert [a])` die Suchbaumeigenschaft. Da  $b$  auch die Suchbaumeigenschaft erfüllt, muß nur noch gezeigt werden, daß alle Markierungen in `(insert [a])` kleiner sind als  $m$ . Es gibt in `insert` drei Aufrufe von `make-node`, die neue Knoten erzeugen können. Alle fügen höchstens 1 zu der Menge der Markierungen des Baumes hinzu. Alle anderen Markierungen sind nach Voraussetzung kleiner als  $m$ , ebenso wie 1. Das Resultat erfüllt also ebenfalls die Suchbaumeigenschaft.
- Im vierten Fall ist  $\tau$  ein Knoten, dessen Markierung *kleiner* ist als  $l$ . Dieser Fall geht analog zum dritten Fall.

□

## AUFGABEN

### AUFGABE 13.13

Welche interessanten Eigenschaften hat die Division? Schreibe diese als Eigenschaften von `/` auf.

### AUFGABE 13.14

Schreibe eine möglichst vollständige Liste interessanter Eigenschaften sowohl der Ihnen bekannten arithmetischen Operationen als auch der logischen Operationen auf. Beziehe dazu auch die Vergleichsoperationen  $<$ ,  $\leq$  etc. ein. Finde außerdem für jede Operation eine interessante Eigenschaft, die *nicht* gilt und überprüfen, ob DrRacket jeweils ein Gegenbeispiel findet.

### AUFGABE 13.15

Für **and** und **or** gelten zwei Distributivgesetze analog dem Distributivgesetz für **\*** und **+**: Formuliere diese als Eigenschaften und lasse sie DrRacket sie überprüfen.

Abstrahiere dann über die nun insgesamt drei Distributivgesetze analog zu **commutativity** und **associativity** und formuliere die drei Distributivgesetze mit Hilfe der Abstraktion neu. Schreibe eine möglichst aussagekräftige Signatur für Deine Abstraktion!

#### AUFGABE 13.16

Schreibe Abstraktionen analog zu **commutativity** und **associativity** für folgende Eigenschaften:

1. DeMorgan
2. Reflexivität
3. Symmetrie
4. Antisymmetrie
5. Transitivität
6. linksneutrales Element
7. rechtsneutrales Element
8. inverses Element

#### AUFGABE 13.17

Versuche, die Eigenschafts-Tests für **number-list=?** auszutricksen, also eine fehlerhafte Version von **number-list=?** zu schreiben, die alle drei **check-property**-Tests besteht. Die **check-expect**-Tests sind für diese Aufgabe nicht relevant.

#### AUFGABE 13.18

Formuliere Eigenschaften von **filter** und **map** im Zusammenhang mit **concatenate** und teste diese.

#### AUFGABE 13.19

Finde eine präzisere Formulierung der Kommutativität von **list-sum** als die in Abschnitt 13.8.4, also eine, an der sich die Eigenschaft, dass die «Reihenfolge der Elemente der Liste keine Rolle spielt» klarer zu sehen ist.

Schreibe dazu eine Funktion, welche die Reihenfolge der Elemente einer Liste abhängig von einer natürlichen Zahl  $n$  verändert, zum Beispiel indem die  $n$ te Permutation der Elemente ausgewählt wird.

#### AUFGABE 13.20

Schreibe einen `check-property`-Test für folgende Eigenschaft:

$$(\text{uncurry } (\text{curry } p_2)) \equiv p_2$$

#### AUFGABE 13.21

Formuliere sinnvolle Eigenschaften von `compose` und `repeat` von Seite ?? und überprüfe diese mit `check-property`!

#### AUFGABE 13.22

Beweise, dass für Funktionen  $p_1$  mit einem Parameter, die einparametrische Funktionen zurückgeben, und Funktionen  $p_2$  mit zwei Parametern gilt:

$$\begin{aligned} (\text{curry } (\text{uncurry } p_1)) &\equiv p_1 \\ (\text{uncurry } (\text{curry } p_2)) &\equiv p_2 \end{aligned}$$

#### AUFGABE 13.23

Beweise die entsprechend dem Beweis der Kommutativität von  $+$  in Abschnitt 13.10.1 die Assoziativität von  $+$  sowie die Distributivität von  $+$  und  $*$  aus Abschnitt 13.7.1.

#### AUFGABE 13.24

Beweise, dass die folgende Funktion natürliche Zahlen quadriert:

```
; Quadrat einer Zahl berechnen
(: square (natural -> natural))
(define square
  (lambda (n)
    (if (= n 0)
```

```

0
(+ (square (- n 1))
  (- (+ n n) 1))))

```

Formuliere dazu auch eine Eigenschaft und überprüfe diese mit `check-property`.

#### AUFGABE 13.25

Beweise, dass auch die folgende Funktion `square` natürliche Zahlen quadriert. Gib die Invariante von `square-helper` an!

```

; Quadrat einer Zahl berechnen
(: square (natural -> natural))
(define square
  (lambda (n)
    (square-helper n 0)))

(define square-helper
  (lambda (n acc)
    (if (= n 0)
        acc
        (square-helper (- n 1)
                        (+ acc
                          (- (+ n n) 1))))))

```

Formuliere dazu auch eine Eigenschaft und überprüfe diese mit `check-property`.

#### AUFGABE 13.26

Beweise mit Hilfe des Substitutionsmodells, dass die `concatenate`-Funktion aus Abschnitt ?? assoziativ ist, dass also für Listen  $l_1$ ,  $l_2$  und  $l_3$  gilt:

$$(\text{concatenate } l_1 (\text{concatenate } l_2 l_3)) = (\text{concatenate } (\text{concatenate } l_1 l_2) l_3)$$

#### AUFGABE 13.27

Beweise die Korrektheit von `search-tree-member?` in Abschnitt 12.5 auf Seite 372. Formuliere zunächst eine geeignete Korrektheitseigenschaft und beweisen diese mit Hilfe von Induktion!

## AUFGABE 13.28

Falls Du Aufgabe 12.20 auf Seite 395 gelöst hast ([search-tree-delete](#)): Beweisen, dass die Funktion die Suchbaumeigenschaft erhält.



## 14 FORTGESCHRITTENES PROGRAMMIEREN MIT REKURSION

FIXME: Vielleicht sowas wie Mergesort, parallele Listenverarbeitung

### 14.1 LASTWAGEN OPTIMAL BELADEN

FIXME: Backtracking sagen

Hier ist ein weiteres Problem, zu dessen Lösung Listen hervorragend taugen: Die Aufgabe ist, einen Lastwagen optimal auszulasten: Aus einem Lager mit zu transportierenden Artikeln, jeder mit einem bestimmten Gewicht, sind solche Artikel auszuwählen, dass die Tragfähigkeit des Lastwagens möglichst gut ausgeschöpft wird.

Für die Lösung muß erst einmal festgelegt werden, wie Ein- und Ausgabe des Programms aussehen sollen. Die elementare Größe im Problem ist ein *Artikel*, der aus einer Artikelnummer und seinem Gewicht besteht. Folgende Daten- und Record-Definition passen dazu:

```
; Ein Artikel ist ein Wert
; (make-article n w)
; wobei n die Nummer des Artikels ist
; und w das Gewicht des Artikels in Kilo
(define-record-procedures article
  make-article article?
  (article-number article-weight))
```

Ein Beispiel-Lager wird durch die folgende Liste beschrieben:

```
; Ein Beispiel-Lager
(: stock (list(article))
(define stock
  (list (make-article 1 274)
        (make-article 2 203)
        (make-article 3 268)
        (make-article 4 264)
        (make-article 5 229)
        (make-article 6 406))
```

```
(make-article 7 220)
(make-article 8 232)
(make-article 9 356)
(make-article 10 197)
(make-article 11 207)
(make-article 12 373)))
```

Die Lösung des Problems soll eine Funktion `load-list` sein, welche eine Liste der Artikel zurückliefert, die in den Lastwagen geladen werden sollen. Neben der Liste der Artikel braucht `load-list` auch noch die Tragfähigkeit eines Lastwagens. Die Funktion soll Kurzbeschreibung und Signatur wie folgt haben:

```
; Maximale Liste von Artikeln berechnen,
; die auf einen Lastwagen passen
(: load-list (list(article) number -> list(article)))
```

Im Fall des Beispiel-Lagers und eines Lastwagens mit 1800 kg Tragfähigkeit soll folgendes passieren, wenn das Programm fertig ist:

```
(load-list stock 1800)
↪ #<list #<record:article 1 274> #<record:article 2 203>
    #<record:article 3 268> #<record:article 6 406>
    #<record:article 7 220> #<record:article 8 232>
    #<record:article 10 197>>
```

Die Funktion arbeitet auf Listen, was folgende Schablone nahelegt:

```
(define load-list
  (lambda (articles capacity)
    (cond
      ((empty? articles) ...)
      ((cons? articles)
       ... (first articles) ...
       ... (load-list (rest articles) capacity) ...))))
```

Wenn keine Artikel da sind, kommen auch keine in den Lastwagen. Die Liste im ersten Zweig ist also leer. Der zweite Fall ist etwas komplizierter. Das liegt daran, daß es dort eine weitere Fallunterscheidung gibt, je nach dem ersten Artikel (`first articles`): die Funktion muß entscheiden, ob dieser erste Artikel schließlich in den Lastwagen kommen soll oder nicht. Ein erstes Ausschlußkriterium ist, wenn der Artikel schwerer ist als die Tragfähigkeit erlaubt:



```
(define load-list
  (lambda (articles capacity)
    (cond
      ((empty? articles) empty)
      ((cons? articles)
       (if (> (article-weight (first articles)) capacity)
           (load-list (rest articles) capacity)
           ... (load-list (rest articles) capacity) ...))))
```

Damit ist die Frage, ob der erste Artikel im Lastwagen landet, aber immer noch nicht abschließend beantwortet. Schließlich muß `load-list` noch entscheiden, ob unter Verwendung dieses Artikels der Lastwagen optimal vollgepackt werden kann. Dazu muß die Funktion vergleichen, wie ein Lastwagen *mit* dem ersten Artikel und wie er *ohne* diesen Artikel am besten vollgepackt werden würde. Die Variante «ohne» wird mit folgendem Ausdruck ausgerechnet:

```
(load-list (rest articles) capacity)
```

Die Variante «mit» ist etwas trickreicher – sie entsteht, wenn im Lastwagen der Platz für den ersten Artikel reserviert wird und der Rest der Tragfähigkeit mit den restlichen Artikeln optimal gefüllt wird. Die optimale Füllung für den Rest wird mit folgendem Ausdruck berechnet, der die Induktionsannahme für `load-list` benutzt:

```
(load-list (rest articles)
  (- capacity (article-weight (first articles))))
```

Die vollständige Artikelliste entsteht dann durch nachträgliches Wieder-Anhängen des ersten Artikels:

```
(cons (first articles)
  (load-list (rest articles)
    (- capacity (article-weight (first articles)))))
```

Diese beiden Listen müssen jetzt nach ihrem Gesamtgewicht verglichen werden. Die Liste mit dem größeren Gewicht gewinnt. Als erster Schritt werden die beiden obigen Ausdrücke in die Schablone eingefügt:

```
(define load-list
  (lambda (articles capacity)
    (cond
```

```

(empty? articles) empty)
(cons? articles)
  (if (> (article-weight (first articles)) capacity)
      (load-list (rest articles) capacity)
      ... (load-list (rest articles) capacity) ...
      ... (cons (first articles)
                 (load-list
                  (rest articles)
                  (- capacity
                     (article-weight (first articles)))))
          ...))))))

```

Die Ausdrücke für die beiden Alternativen sind in dieser Form unhandlich groß, was die Funktion schon unübersichtlich macht, bevor sie überhaupt fertig ist. Es lohnt sich also, ihnen Namen zu geben:

```

(define load-list
  (lambda (articles capacity)
    (cond
      ((empty? articles) empty)
      ((cons? articles)
       (if (> (article-weight (first articles)) capacity)
           (load-list (rest articles) capacity)
           (let ((articles-1 (load-list (rest articles) capacity))
                 (articles-2
                  (cons (first articles)
                        (load-list
                         (rest articles)
                         (- capacity
                            (article-weight (first articles)))))
                  ...)))))))

```

Der Ausdruck `(article-weight (first articles))` kommt zweimal vor. Die Einführung einer weiteren lokalen Variable macht die Funktion noch übersichtlicher:

```

(define load-list
  (lambda (articles capacity)
    (cond

```



```

articles-1
articles-2)))))))))

```

Es fehlt noch `articles-weight`, die wieder streng nach Anleitung geht und für welche die Schablone folgendermaßen aussieht:

```

(define articles-weight
  (lambda (articles)
    (cond
      ((empty? articles) ...)
      ((cons? articles)
       ... (first articles) ...
       ... (articles-weight (rest articles)) ...))))

```

Das Gesamtgewicht der leeren Liste ist 0 – der erste Fall ist also wieder einmal einfach. Im zweiten Fall interessiert vom ersten Artikel nur das Gewicht:

```

(define articles-weight
  (lambda (articles)
    (cond
      ((empty? articles) 0)
      ((cons? articles)
       ... (article-weight (first articles)) ...
       ... (articles-weight (rest articles)) ...))))

```

Nach Induktionsannahme liefert `(articles-weight (rest articles))` das Gewicht der restlichen Artikel. Das Gewicht des ersten Artikels muß also nur addiert werden:

```

(define articles-weight
  (lambda (articles)
    (cond
      ((empty? articles) 0)
      ((cons? articles)
       (+ (article-weight (first articles))
          (articles-weight (rest articles)))))))

```

Mit `articles-weight` läßt sich bestimmen, wie voll der Lastwagen beladen ist. Im Falle von `stock` ist das Ergebnis sehr erfreulich; kein Platz wird verschenkt:

```
(articles-weight (load-list stock 1800))
```

↪ 1800

## AUFGABEN

### AUFGABE 14.1

Schreiben Sie eine Funktion `drink-machine`, die das Finanzmanagement eines Getränkeautomaten durchführt.

`Drink-machine` soll drei Argumente akzeptieren: den Preis eines Getränks (in Euro-Cents), eine Liste der Centbeträge der Wechselgeldmünzen, die noch im Automaten sind, und den Geldbetrag, den der Kunde eingeworfen hat. (Also gibt es ein Listenelement pro Münze. Falls z.B. mehrere Groschen im Automaten sind, finden sich in der Liste mehrmals die Zahl 10.) Herauskommen soll eine Liste der Centbeträge der Münzen, welche die Maschine herausgibt oder `#f`, falls die Maschine nicht herausgeben kann.

```
(drink-machine 140 (list 50 100 500 10 10) 200)
```

```
↪ #<list 50 10>
```

```
(drink-machine 140 (list 50 20 20 20) 200)
```

```
↪ #<list 20 20 20>
```

### AUFGABE 14.2

Schreiben Sie eine Funktion `backlog-list`, welche die gleichen Eingaben wie `load-list` akzeptiert, aber die nach dem Beladen eines Lastwagens im Lager zurückbleibenden Artikel zurückgibt. Die Funktion soll `load-list` als Hilfsfunktion verwenden.



## 15 SCHRITTWEISE VERFEINERUNG

In der Praxis gelingt es nur selten, Programme erst vollständig zu planen und dann umzusetzen: Manchmal ist die Aufgabenstellung bei Entwicklungsbeginn noch nicht vollständig bekannt. Manchmal ist die Aufgabenstellung zwar bekannt aber unübersichtlich. Manchmal ändert sich die Aufgabenstellung nachträglich. Deshalb muß die Entwicklung oft beginnen, bevor die gesamte Planung steht. Das Programm wird dann bis zu einem bestimmten Grad entwickelt und dann *schrittweise verfeinert*, um die noch vorher unbekannten Aspekte der Aufgabenstellung abzudecken. Dabei ist es wichtig, zwischenzeitlich gemachte Annahmen klar herauszustellen, fehlerhafte Programmteile auch wieder bereitwillig zu löschen, und Schnittstellen gegen Änderungen zu isolieren. In diesem Kapitel werden verschiedene Beispiele und Techniken für diese schrittweise Verfeinerung vorgestellt.

### 15.1 LÖSCHEN IN SUCHBÄUMEN

Das Löschen eines Elements in einem Suchbaum ist deutlich schwieriger als das Einfügen. Beim Suchbaum s3 erscheint die Aufgabe noch einfach. Hier das Bild dazu:

TBD

Angenommen, die 3 soll gelöscht werden. In diesem Fall kann der Suchbaum durch dessen rechten Teilbaum ersetzt werden:

TBD

Angenommen, die 17 soll gelöscht werden: Dann kann der Teilbaum mit 17 an der Spitze durch seinen eigenen linken Teilbaum ersetzt werden: TBD

Ähnlich kann der Teilbaum mit der 5 an der Spitze durch den leeren Baum ersetzt werden:

TBD

Wird der Suchbaum nur minimal erweitert, wird die Sache schon schwieriger:

TBD

Wie könnte hier die 17 gelöscht werden? Es ist nicht möglich, den Teilbaum mit 17 an der Spitze einfach durch seinen linken oder rechten Teilbaum zu ersetzen, weil dann der jeweils andere Teilbaum unter den Tisch fiel. Es ist also notwendig, die Baumstruktur zu reorganisieren. Mit dieser Beobachtung ist klar, daß die Aufgabe nicht einfach dadurch zu lösen ist, daß die Konstruktionsanleitung einmal befolgt wird: Stattdessen empfiehlt es sich, etwas nachzudenken und zumindest eine grobe Lösungsstrategie zu entwickeln, bevor mit dem Programmieren begonnen wird.

Eine Möglichkeit für eine Lösung wäre, den kompletten Suchbaum neu zu bilden: Dazu könnte die Lösch-Funktion alle Elemente in einer Liste aufsammeln, und dann mit `search-tree-insert` sukzessive aus allen Elementen bis auf das zu löschende einen neuen Suchbaum konstruieren. (Siehe Aufgabe 15.1.) Dies wäre allerdings eine zeitaufwendige Methode, die überproportional länger dauert, je größer der Suchbaum ist.

Besser wäre also, wenn beim Löschen soviel wie möglich des restlichen Baums intakt bleibt: das soll für die Überlegungen erst einmal die Maxime sein. Hypothetisch wäre es möglich, daß dieser Ansatz nicht zum Ziel führt und diese Annahme wieder rückgängig gemacht werden muß. Darum ist es sinnvoll, die Annahme deutlich sichtbar zu dokumentieren:

*Annahme:* Es ist möglich, ein Element zu löschen, indem nur der Teilbaum verändert wird, an dessen Wurzel das Element sitzt; der Rest soll gleichbleiben.

Etwas Nachdenken ergibt, daß der oben beschriebene Ansatz, nur den Teilbaum zu ersetzen, an dessen Spitze das zu löschende Element steht und den restlichen Baum unverändert zu lassen, auch hier funktioniert. Es gibt zwei Möglichkeiten, dies zu tun:

TBD

Mit anderen Worten, es wird ein Element von weiter unten «hochgezogen». Es bleibt die Frage, wie die Funktion das Element auswählen könnte, das hochgezogen werden muß. Offensichtlich kann entweder ein Element aus dem linken oder aus dem rechten Teilbaum ausgewählt werden. Ideal wäre, wenn der jeweils andere Teilbaum intakt bliebe. Es muß also ein Element so hochgezogen werden, daß die Suchbaumeigenschaft nicht verletzt wird. Angenommen, das Element soll aus dem linken Teilbaum genommen werden: Dann muß es größer sein als alle anderen Elemente des linken Teilbaums. (Es muß auch kleiner als alle Elemente des rechten Teilbaums sein, aber dies ist automatisch gegeben, weil die Suchbaumeigenschaft im ursprünglichen Baum bereits gilt.) Das hochziehende Element des linken Teilbaums wäre also das *Maximum* aller Elemente des linken Teilbaums. (Umgekehrt wäre es genauso möglich, das Minimum der Elemente des rechten Teilbaums auszuwählen: reine Geschmacksfrage.)

Zu diesem Zeitpunkt stellen sich möglicherweise bereits Kopfschmerzen beim Programmierer ein: Wie wird das hochziehende Element gefunden? Wie funktioniert der Hochziehprozeß selbst? Daß diese Aufgaben lösbar sind, ist klar, aber zusammen mit den anderen zu lösenden Teilaufgaben vom Anfang dieses Abschnitts ist es vielleicht etwas viel auf einmal.

Darum ist es sinnvoll, erst einmal die schon klaren Teile der Aufgabenlösung umzusetzen und die noch unklaren Teile per Wunschdenken auf später zu verschieben. Ein Teil der Strategie ist



bereits vollständig erkennbar: Das Löschen funktioniert, indem der Teilbaum, an dessen Wurzel das zu löschende Element klebt, durch einen anderen Baum ersetzt wird; der Rest des Baums ist intakt. Es gibt also bereits zwei erkennbare Teilaufgaben:

1. Den zu ersetzenden Teilbaum finden.
2. Aus einem Baum mit einem zu löschenden Element an der Spitze einen anderen Baum machen, der das Element nicht mehr enthält.

Insbesondere müsste es leicht möglich sein, die zweite Teilaufgabe soweit zu lösen, daß die Löscho-Funktion zu zumindest für `s3` funktioniert:

```
; Element aus Suchbaum löschen
(: search-tree-delete (%a (search-tree %a) -> (search-tree %a)))

(check-expect (search-tree-tree (search-tree-delete 3 s3))
  (make-node 17
    (make-node 5 the-empty-tree the-empty-tree)
    the-empty-tree))
(check-expect (search-tree-tree (search-tree-delete 5 s3))
  (make-node 3
    the-empty-tree
    (make-node 17 the-empty-tree the-empty-tree)))
(check-expect (search-tree-tree (search-tree-delete 17 s3))
  (make-node 3 the-empty-tree (make-node 5 the-empty-tree the-empty-tree)))

(check-expect (search-tree-tree (search-tree-delete 28 s3))
  (search-tree-tree s3))
```

Das Gerüst ist das gleiche, das schon für `search-tree-member?` und `search-tree-insert` entwickelt wurde. Die zwei Hilfsfunktionen werden mit `letrec` gebunden:

```
(define search-tree-delete
  (lambda (e s)
    (let ((=proc (search-tree-label-equal-proc s))
          (<proc (search-tree-label-less-than-proc s)))
      (letrec
        ; Element finden und löschen
        ; (: delete ((tree %a) -> (tree %a))
        ((delete
```

```

        (lambda (t)
          ...))
; Element an Wurzel löschen
; (: delete-top ((node %a (tree %a) (tree %a)) -> (tree %a))
(delete-top
  (lambda (n)
    ...)))
(make-search-tree
  =proc <proc
    (delete (search-tree-tree s))))))

```

Die Hilfsfunktion `delete` folgt dabei dem gleichen Muster wie die Hilfsfunktionen in `search-tree-member?` und `search-tree-insert`: Die rekursiven Aufrufe folgen der Struktur des Baums gemäß der Suchbaumeigenschaft. Wenn das gesuchte Element gefunden ist, kann die Hilfsfunktion `delete-top` den Rest der Arbeit übernehmen. Bei `delete-top` sind bereits zwei Fälle klar: Wenn ein Teilbaum leer ist, so kann der Baum durch den jeweils anderen Teilbaum ersetzt werden:

```

(define search-tree-delete
  (lambda (e s)
    (let ((=proc (search-tree-label-equal-proc s))
          (<proc (search-tree-label-less-than-proc s)))
      (letrec
        ; Element finden und löschen
        ; (: delete ((tree %a) -> (tree %a))
        ((delete
          (lambda (t)
            (cond
              ((empty-tree? t) t)
              ((node? t)
               (cond
                 ((=proc e (node-label t))
                  (delete-top t))
                 ((<proc e (node-label t))
                  (make-node (node-label t)
                             (delete (node-left-branch t))
                             (node-right-branch t)))
                 (else

```

```

                (make-node (node-label t)
                           (node-left-branch t)
                           (delete (node-right-branch t)))))))))
; Element an Wurzel löschen
; (: delete-top ((node %a (tree %a) (tree %a)) -> (tree %a))
(delete-top
  (lambda (n)
    (cond
      ((empty-tree? (node-left-branch n))
       (node-right-branch n))
      ((empty-tree? (node-right-branch n))
       (node-left-branch n))
      (else ...))))))
(make-search-tree
  =proc <proc
  (delete (search-tree-tree s))))))

```

Die Testfälle mit `s3` funktionieren damit schon einmal. Es ist durchaus sinnvoll, bewußt die einfachen Fälle zuerst zu programmieren und auch zu testen. Gibt es später Probleme bei den komplizierteren Fällen, so sind diese einfacher zu lokalisieren.

Der folgende Testfall allerdings beschwert sich noch über den fehlenden Code im `else`-Zweig:

```

(define s4
  (make-search-tree
    = <
    (make-node 17
      (make-node 5 the-empty-tree the-empty-tree)
      (make-node 20 the-empty-tree the-empty-tree))))

(check-expect (search-tree-tree (search-tree-delete 17 s4))
  (make-node 5 the-empty-tree
    (make-node 20 the-empty-tree the-empty-tree)))

```

Für die Ellipse muß nun also das maximale Element des linken Teilbaums hochgezogen werden; der rechte Teilbaum bleibt unangetastet. Das könnte etwa so aussehen:

```

(make-node (find-maximum (node-left-branch n))
  ...

```

```
(node-right-branch n))
```

Per Wunschdenken wird hier eine Funktion `find-maximum` vorausgesetzt:

```
; Maximum von Teilbaum finden
(: find-maximum ((tree %a) -> %a))
```

Aber auch im obigen Ausdruck fehlt noch etwas, nämlich der linke Teilbaum des neu konstruierten Baums. Der soll alle Elemente des ursprünglichen Teilbaums finden, *außer dem Maximum*. Wie kann `delete-top` das Maximum loswerden? Am einfachsten durch die Verwendung der Lösch-Funktion, die gerade geschrieben wird. Erster Anlauf:

```
(make-node (find-maximum (node-left-branch n))
  (search-tree-delete
    (find-maximum (node-left-branch n))
    (node-left-branch n))
  (node-right-branch n))
```

Hier taucht `(find-maximum (node-left-branch n))` mehrfach auf, es sollte also durch ein `let` gebunden werden:

```
(let ((el (find-maximum (node-left-branch n))))
  (make-node el
    (search-tree-delete
      el
      (node-left-branch n))
    (node-right-branch n)))
```

Leider ist der Code so nicht richtig, da `(node-left-branch n)` kein `search-tree`-Wert ist, sondern nur ein `tree`-Wert: es gibt also eine Vertragsverletzung. Der Baum `(node-left-branch n)` muß also noch in ein `search-tree`-Record eingewickelt werden. Außerdem kommt bei `search-tree-delete` wieder ein `search-tree`-Wert heraus, während `make-node` einen `tree`-Wert erwartet, der hinterher wieder ausgewickelt werden muß:

```
(let ((el (find-maximum (node-left-branch n))))
  (make-node el
    (search-tree-tree
      (search-tree-delete
```

```

el
  (make-search-tree =proc <proc
                    (node-left-branch n)))
(node-right-branch n)))

```

Alternativ zum umständlichen Ein- und Auswickeln kann auch die `delete`-Hilfsfunktion um einen Parameter für das zu löschende Element erweitert werden:

```

(letrec
  ; Element finden und löschen
  ; (: delete (%a (tree %a) -> (tree %a)))
  ((delete
    (lambda (e t)
      (cond
        ((empty-tree? t) t)
        ((node? t)
         (cond
           ((=proc e (node-label t))
            (delete-top t))
           (<proc e (node-label t))
            (make-node (node-label t)
                       (delete e (node-left-branch t))
                       (node-right-branch t)))
          (else
           (make-node (node-label t)
                       (node-left-branch t)
                       (delete e (node-right-branch t))))))))))

```

Dann lautet der Code im `else`-Zweig:

```

(let ((el (find-maximum (node-left-branch n))))
  (make-node el
              (delete el (node-left-branch n))
              (node-right-branch n)))

```

Fehlt nur noch die Hilfsfunktion `find-maximum`. Beim Ermitteln des Maximums hilft die Suchbaumeigenschaft: Es ist einfach das Element, das am weitesten rechts im Baum steht. Die Funktion ist also ganz einfach zu schreiben. Komplett sieht `search-tree-delete` dann so aus:

---

```

(define search-tree-delete
  (lambda (e s)
    (let ((=proc (search-tree-label-equal-proc s))
          (<proc (search-tree-label-less-than-proc s)))
      (letrec
        ; Element finden und löschen
        ; (: delete (%a (tree %a) -> (tree %a))
        ((delete
          (lambda (e t)
            (cond
              ((empty-tree? t) t)
              ((node? t)
               (cond
                 ((=proc e (node-label t))
                  (delete-top t))
                 ((<proc e (node-label t))
                  (make-node (node-label t)
                             (delete e (node-left-branch t))
                             (node-right-branch t)))
                 (else
                  (make-node (node-label t)
                             (node-left-branch t)
                             (delete e (node-right-branch t))))))))))
          ; Element an der Wurzel löschen
          ; (: delete-top ((node %a (tree %a) (tree %a)) -> (tree %a))
          (delete-top
           (lambda (n)
             (cond
               ((empty-tree? (node-left-branch n))
                (node-right-branch n))
               ((empty-tree? (node-right-branch n))
                (node-left-branch n))
               (else
                (let ((el (find-maximum (node-left-branch n))))
                  (make-node el

```

```

                                (delete el (node-left-branch n))
                                (node-right-branch n))))))
; Maximum von Teilbaum finden
; (: find-maximum ((tree %a) -> %a))
(find-maximum
  (lambda (t)
    (cond
      ((empty-tree? t)
        (violation "unerwarteter leerer Baum"))
      ((node? t)
        (let ((right (node-right-branch t)))
          (if (empty-tree? right)
              (node-label t)
              (find-maximum (node-right-branch t))))))))))

(make-search-tree
  =proc <proc
  (delete e (search-tree-tree s))))))

```

Beim Testen fällt auf, daß der rekursive Aufruf von `find-maximum` noch nicht abgedeckt ist. (Daß der Aufruf von `violation` nicht abgedeckt ist, ist erwünscht – es soll ja nicht passieren.) Es muß also noch ein Testfall konstruiert werden:

```

(define s4
  (make-search-tree
    = <
    (make-node 17
      (make-node 5 the-empty-tree the-empty-tree)
      (make-node 20 the-empty-tree the-empty-tree))))

(check-expect (search-tree-tree (search-tree-delete 17 s4))
  (make-node 5 the-empty-tree
    (make-node 20 the-empty-tree the-empty-tree)))

```

## 15.2 DATENVERFEINERUNG

Bei «realen» Programmieraufgaben sind selten schon präzise Datendefinitionen vorgegeben: Stattdessen liegen oft nur ein paar vage Anforderungen an die Software vor, die sich unter Umständen auch noch mit der Zeit ändern. In solchen Fällen ist es nicht möglich, wie bei den meisten Übungsaufgaben dieses Buchs, die Datendefinition direkt an der Aufgabenstellung abzulesen und alle zu anfallenden Probleme mit einem Wurf zu lösen.

Das Beispiel dieses Abschnitts ist, die Verwaltungsdaten von Wohnungen zu modellieren. Über eine Wohnung läßt sich viel sagen, von dem nur manches je nach Kontext und Aufgabe wichtig ist. Darum ist es schwierig, im ersten Anlauf eine vollständige und präzise Datendefinition zu entwerfen. In diesem Abschnitt geht es darum, was passiert, wenn sich die Anforderungen und damit die Datendefinitionen ändern. Angenommen, es geht zunächst einmal vage um die Adresse, die Bewohner, die Zimmer und die Wohnfläche. Immerhin sind daran schon eine Daten- und eine Record-Definition ablesbar:

```
; Eine Wohnung besteht aus:
; - Adresse
; - Bewohner
; - Zimmern
(define-record-procedures apartment
  make-apartment apartment?
  (apartment-address apartment-resident apartment-rooms))

(: make-apartment (string string (list room) -> apartment))
```

Hier ist allerdings bereits eine willkürliche Entscheidung gefallen: Die Datendefinition enthält nicht direkt die Grundfläche der Wohnung. Diese soll pro Zimmer angegeben werden. Der Vertrag `room` für Zimmer ist noch offen; da über die Zimmer außer der Grundfläche nichts weiter bekannt ist, wäre eine Möglichkeit die folgende Definition:

```
(define-contract room rational)
```

Diese naive Definition hat zwei Nachteile:

- Es gibt unmittelbar kein eindeutiges Prädikat für Zimmer.
- Es ist klar, daß sich über ein Zimmer noch mehr Dinge sagen lassen als die Grundfläche: Früher oder später werden Zimmer zu zusammengesetzten Daten.

Aus diesen Gründen empfiehlt es sich, von vornherein von zusammengesetzten Daten auszugehen:



```
; Ein Zimmer besteht aus:
; - Fläche
(define-record-procedures room
  make-room room?
  (room-area))

(: make-room (rational -> room))
```

Hier sind zwei Wohnungen für Tests:

```
(define a1
  (make-apartment
    "Entenhausener Straße 15"
    "Sperber"
    (list (make-room 30) (make-room 50) (make-room 60))))

(define a2
  (make-apartment
    "Froschgasse 17"
    "Crestani"
    (list (make-room 20) (make-room 40))))
```

Die Gesamtfläche einer Wohnung läßt sich einfach durch Addieren der Flächen der Zimmer berechnen:

```
; Fläche einer Wohnung berechnen
(: apartment-area (apartment -> number))
(check-expect (apartment-area a1) 140)
(check-expect (apartment-area a2) 60)

(define apartment-area
  (lambda (a)
    (fold 0 +
      (map room-area (apartment-rooms a))))))
```

An dieser Stelle wird eine positive Eigenschaft der bisherigen Daten- und Funktionsdefinitionen

deutlich. Angenommen, die Record-Definition hätte die Grundfläche direkt einer Wohnung zugeordnet:

```
; Eine Wohnung besteht aus:
; - Adresse
; - Bewohner
, - Grundfläche
; - Zimmern
(define-record-procedures apartment
  make-apartment apartment?
  (apartment-address apartment-resident apartment-rooms apartment-area))

(: make-apartment (string string (list room) rational -> apartment))
```

In diesem Fall würde der Selektor `apartment-area` sich nach außen genauso verhalten wie die Funktion `apartment-area` – gleicher Vertrag, gleiche Funktion. Die Tatsache, daß der Record-Selektor eine reguläre Funktion ist, die sich von «selbstgeschriebenen» Funktionen nach außen nicht unterscheidet, isoliert Benutzer der Wohnungs-Funktionalität von den Unterschieden zwischen den zwei Varianten.

Aus Sicht eines Wohnungsbesitzers interessieren möglicherweise noch weitere Daten, zum Beispiel, ob ein Bewohner noch Untermieter hat. Es ist sogar möglich, daß Untermieter ihrerseits selbst Untermieter haben. Dieser Idee ließe sich Rechnung tragen, indem eine Wohnung statt aus Zimmern aus «Abschnitten» besteht, und jeder Abschnitt entweder ein Zimmer oder ein untervermieteter Wohnungsteil ist. Die Definition für Wohnungen müsste also folgendermaßen geändert werden:

```
; Eine Wohnung besteht aus:
; - Adresse
; - Bewohner
; - Abschnitten
(define-record-procedures apartment
  make-apartment apartment?
  (apartment-address apartment-resident apartment-sections))

(: make-apartment (string string (list section) -> apartment))
```

Bei Abschnitten (`section`) handelt es sich klar um gemischte Daten:

```
; Ein Abschnitt ist eins der folgenden:
```

```
; - ein Zimmer
; - ein untervermieter Teil der Wohnung
(define-contract section (mixed room sublet))
```

Zimmer sind wie gehabt. Untervermietete Wohnungsteile sind wie Wohnungen, haben aber keine Adresse:

```
; Ein unvermieteter Teil des Wohnung besteht aus:
; - Bewohner
; - Abschnitten
(define-record-procedures sublet
  make-sublet sublet?
  (sublet-resident sublet-sections))
```

```
(: make-sublet (string (list section) -> sublet))
```

Hier ist ein Beispiel für solche einen untervermieteten Teil sowie ein Apartment, das es enthält:

```
(define s1
  (make-sublet "Taschenbier"
    (list (make-room 20)
      (make-sublet "Sams"
        (list (make-room 2)))
      (make-room 10))))

(define a3
  (make-apartment
    "Flugentenschneise 17"
    "Rotkohl"
    (list (make-room 100)
      s1
      (make-room 30))))
```

Es kann nun sein, daß die «Wohnungs-Funktionalität» schon in ein größeres Programm eingebaut wurde, daß dessen Funktionen benutzt, insbesondere den Selektor `apartment-rooms`. Der ist im Zuge der Erweiterung um Untermiete unter den Tisch gefallen. Das Konzept der Zimmer einer Wohnung ist aber immer noch sinnvoll; entsprechend ist es möglich, `apartment-rooms` wieder zur Verfügung zu stellen, diesmal als «normale» Funktion mit einer Hilfsfunktion, um die Zimmer eines Abschnitts aufzuzählen:

```

; Zimmer einer Wohnung aufzählen
(: apartment-rooms (apartment -> (list room)))

(check-expect (apartment-rooms a1)
  (list (make-room 30) (make-room 50) (make-room 60)))
(check-expect (apartment-rooms a3)
  (list (make-room 100)
        (make-room 20)
        (make-room 2)
        (make-room 10)
        (make-room 30)))

(define apartment-rooms
  (lambda (a)
    (fold empty append
      (map section-rooms (apartment-sections a))))))

; Zimmer eines Wohnungsabschnitts aufzählen
(: section-rooms (section -> (list room)))

(check-expect (section-rooms (make-room 30))
  (list (make-room 30)))
(check-expect (section-rooms s1)
  (list (make-room 20)
        (make-room 2)
        (make-room 10)))

(define section-rooms
  (lambda (s)
    (cond
      ((room? s) (list s))
      ((sublet? s)
       (fold empty append
         (map section-rooms (sublet-sections s)))))))

```

Bei der Abspaltung von Hilfsfunktionen gibt es Ermessensspielraum: Die Hilfsfunktion `section-rooms`

ist notwendig, weil sie rekursiv ist.

«Von außen» ist kein Unterschied zwischen dem alten Selektor `apartment-rooms` und der neuen Funktion zur erkennen – die Schnittstelle ist gleich geblieben: Dies ist ein wichtiger Beitrag zur *Modularität*, die es erlaubt, Änderungen an der Wohnungs-Funktionalität vorzunehmen, ohne daß andere Programmteile, welche die Funktionalität benutzen, verändert werden müssen.

Zu den Programmteilen, die `apartment-rooms` verwenden, gehört auch `apartment-area`, und tatsächlich funktioniert `apartment-area` ohne Anpassungen wie vorher. Außerdem funktioniert es auch bei Wohnungen mit Untermietern, wie ein zusätzlicher Testfall ermittelt:

```
(check-expect (apartment-area a3) 162)
```

Mit den Erweiterungen an den Datendefinitionen sind allerdings auch neue Funktionalitäten möglich. Zum Beispiel könnte eine Funktion für die Bewohner einer Wohnung präzise Adressen ausrechnen, etwa so:

```
; Adresse eines Bewohners berechnen
(: resident-address (string apartment -> string))

(check-expect (resident-address "Sperber" a1) "Sperber")

(check-expect (resident-address "Taschenbier" a3)
               "Taschenbier bei Rotkohl")
(check-expect (resident-address "Sams" a3)
               "Sams bei Taschenbier bei Rotkohl")
```

Die Funktion muß also gleichzeitig nach dem Bewohner suchen und die Adreßangabe konstruieren. Wenn der gesuchte Bewohner gerade der «Hauptbewohner» der Wohnung ist, ist die Aufgabe einfach:

```
(define resident-address
  (lambda (r a)
    (if (string=? r (apartment-resident a))
        r
        ...)))
```

Damit funktioniert bereits der erste Testfall.

Die anderen Adreßangaben haben jetzt alle die Form *x bei y* wobei *y* der Bewohner der Wohnung ist:

```
(define resident-address
  (lambda (r a)
    (if (string=? r (apartment-resident a))
        r
        (string-append ... " bei " (apartment-resident a)))))
```

Der Präfix  $x$  hängt vom Wohnabschnitt ab, in dem der gesuchte Bewohner wohnt. Per Wunschdenken sei eine Funktion vorausgesetzt, die in den Abschnitten einer Wohnung nach dem Bewohner sucht und den entsprechenden Präfix berechnet:

```
; Adreß-Präfix eines Abschnitts-Bewohners berechnen
(: sections-prefix (string (list section) -> string))
```

Mit Hilfe dieser Funktion kann `resident-address` erweitert werden:

```
(define resident-address
  (lambda (r a)
    (if (string=? r (apartment-resident a))
        r
        (string-append (sections-prefix r (apartment-sections a))
                        " bei "
                        (apartment-resident a)))))
```

Nun zur Funktion `sections-prefix`. Hier ein geeigneter Testfall:

```
(check-expect (sections-prefix "Taschenbier" (apartment-sections a3))
               "Taschenbier")
```

Gerüst und Schablone:

```
(define sections-prefix
  (lambda (r lis)
    (cond
      ((empty? lis) ...)
      ((cons? lis)
       ... (first lis) ...
       ... (sections-prefix r (rest lis)) ...))))
```

Beim Ausfüllen der Schablone wird sofort klar, daß ein geeigneter Rückgabewert für den `empty?`-Zweig noch fehlt: In diesem Fall ist der Bewohner gar nicht gefunden worden. Für diesen Umstand (der in den Überlegungen bisher noch gar nicht auftauchte) muß ein Extra-Wert eingeführt werden:

```
; Wert für nicht gefundenen Bewohner
(define-record-procedures not-found
  make-not-found not-found?
  ())
```

```
(define the-not-found (make-not-found))
```

Entsprechend muß der Vertrag von `sections-prefix` erweitert werden:

```
(: sections-prefix (string (list section) -> (mixed string not-found)))
```

Da es sich bei `(first lis)` um gemischte Daten handelt, wird eine Verzweigung fällig. Wieder wird per Wunschdenken eine Hilfsfunktion angenommen, diesmal `sublet-prefix`, die den Adreß-Präfix für eine Untermiete ausrechnet, möglich:

```
(define sections-prefix
  (lambda (r lis)
    (cond
      ((empty? lis) the-not-found)
      ((cons? lis)
       (let ((f (first lis)))
         (cond
           ((room? f) (sections-prefix r (rest lis)))
           ((sublet? f)
            (let ((prefix (sublet-prefix r f)))
              (cond
                ((not-found? prefix)
                 (sections-prefix r (rest lis)))
                ((string? prefix)
                 prefix))))))))))
```

Die Hilfsfunktion `sublet-prefix` soll sich folgendermaßen verhalten:

```
; Adreß-Präfix eines Untermieters bewohnen
(: sublet-prefix (string sublet -> (mixed string not-found)))

(check-expect (sublet-prefix "Taschenbier" s1) "Taschenbier")
(check-expect (sublet-prefix "Sams" s1) "Sams bei Taschenbier")
(check-expect (sublet-prefix "Merz" s1) the-not-found)
```

Bei Untervermietungen ist die Arbeit dann erledigt, wenn der Bewohner gerade der gesuchte ist:

```
(define sublet-prefix
  (lambda (r s)
    (if (string=? r (sublet-resident s))
        r
        ...)))
```

Dies sieht ähnlich aus wie bei `apartment-prefix`; auch hier funktioniert schon der erste Testfall. Für die Alternative kann `sections-prefix` benutzt werden:

```
(define sublet-prefix
  (lambda (r s)
    (if (string=? r (sublet-resident s))
        r
        (let ((prefix (sections-prefix r (sublet-sections s))))
          (cond
            ((not-found? prefix) the-not-found)
            ((string? prefix)
             (string-append prefix " bei " (sublet-resident s))))))))
```

Damit sieht das Programm fast fertig aus. Allerdings ist der Änderung im Vertrag von `sections-prefix` noch nicht überall Rechnung getragen worden, namentlich nicht in `resident-address`: Einerseits muß `resident-address` eine Fallunterscheidung für den Rückgabewert von `sections-prefix` einführen; dies heißt andererseits aber auch, daß `resident-address` selbst in die Lage kommen kann, daß der gesuchte Bewohner nicht gefunden wird. Damit werden neben der Erweiterung des Rumpfes der Funktion auch eine Änderung des Vertrags und ein neuer Testfall fällig:

```
(: resident-address (string apartment -> (mixed string not-found)))
(check-expect (resident-address "Müller" a1) the-not-found)
```

```
(define resident-address
  (lambda (r a)
    (if (string=? r (apartment-resident a))
        r
        (let ((prefix (sections-prefix r (apartment-sections a))))
          (cond
            ((not-found? prefix)
```



```

the-not-found)
((string? prefix)
 (string-append prefix
  " bei "
  (apartment-resident a))))))

```

Solche Änderungen an Verträgen, die im Rahmen der schrittweisen Verfeinerung auftreten, sind wie Wellen im Teppichboden: Sie müssen durch die gesamte Ausdehnung des Programms weitergegeben werden, bis sie schließlich am Rand angekommen sind.

Am Beispiel der Wohnungsverwaltung werden einige typische Einsichten über die Software-Entwicklung deutlich:

- Einmal begonnene Software wird fast immer später erweitert und für Aufgaben eingesetzt, für die sie ursprünglich nicht gedacht war. Darum ist es sinnvoll, von vornherein die Möglichkeit späterer Erweiterungen zu berücksichtigen
- Auch wenn scheinbar das gesamte Problem vor dem Beginn der Programm-Entwicklung bekannt ist, treten häufig noch beim Programmieren neue Einsichten zutage, die Anpassungen erfordern.
- Änderungen an Verträgen – also Schnittstellen – ziehen häufig andere Änderungen in Funktionen nach sich, die diese benutzen. Es ist wichtig, diese gewissenhaft durchzuführen.

In Mantra-Form sehen diese Einsichten so aus:

MANTRA 12

Programme wachsen.

MANTRA 13

Es ist nicht möglich, für alle Eventualitäten im voraus zu planen.

MANTRA 14

Änderungen an Schnittstellen müssen vollständig durch ein Programm propagiert werden.

## AUFGABEN

### AUFGABE 15.1

Eine alternative Implementierung von `search-tree-delete` funktioniert folgendermaßen:

1. Alle Elemente des Baums werden in einer Liste aufgezählt.
2. Das zu löschende Element wird aus der Liste entfernt.
3. Ein völlig neuer Suchbaum aus wird aus der bereinigten Liste konstruiert.

Realisieren Sie diese Alternative! Wie unterscheidet sie sich in der Laufzeit von der im Text präsentierten Implementierung?

#### AUFGABE 15.2

Zwischen den Record-Definitionen von `apartment` und `sublet` gibt es Gemeinsamkeiten. Ist es möglich, von diesen Gemeinsamkeiten zu profitieren und den Code zu vereinfachen? Ändern Sie den Code entsprechend und beurteilen Sie, ob es sich gelohnt hat.

## 16 DER $\lambda$ -KALKÜL

Nachdem die bisherigen Kapitel den Bogen von der praktischen Konstruktion einfacher Programme bis zur objektorientierten Programmierung geschlagen haben, beleuchtet dieses Kapitel eine wichtige theoretische Grundlage der Programmierung.

Für die Informatik ist der Begriff des *logischen Kalküls* von zentraler Bedeutung. Ein Kalkül dient dazu, auf formale Art und Weise wahre Aussagen abzuleiten, ohne daß es dabei nötig wird, über den Sinn der Aussagen nachzudenken. Der  $\lambda$ -Kalkül ist ein logischer Kalkül, der als die Basis für eine formale Beschreibung des Verhaltens von Computerprogrammen dient. Die Lehrsprache baut direkt auf dem  $\lambda$ -Kalkül auf: es ist kein Zufall, daß das Schlüsselwort für die Abstraktion `lambda` heißt. Es gibt noch viele weitere Einsatzgebiete für den  $\lambda$ -Kalkül, insbesondere bei der Konstruktion von besonders effizienten Übersetzern für Programmiersprachen, in der Logik und der Linguistik, und bei der Entwicklung und Sicherheitsüberprüfung von mobilem Code im Internet.

### 16.1 SPRACHE UND REDUKTIONSEMANTIK

DEFINITION 16.1 (SPRACHE DES  $\lambda$ -KALKÜLS  $\mathcal{L}_\lambda$ )

Sei  $V$  eine abzählbare Menge von Variablen. Die Sprache des  $\lambda$ -Kalküls, die Menge der  $\lambda$ -Terme,  $\mathcal{L}_\lambda$ , ist durch folgende Grammatik definiert:

$$\begin{aligned} \langle \mathcal{L}_\lambda \rangle &\rightarrow \langle V \rangle \\ &| (\mathcal{L}_\lambda \ \mathcal{L}_\lambda) \\ &| (\lambda \langle V \rangle . \mathcal{L}_\lambda) \end{aligned}$$

Ein  $\lambda$ -Term der Form  $(e_0 \ e_1)$  heißt *Applikation* mit *Operator*  $e_0$  und *Operand*  $e_1$ . Ein Term der Form  $(\lambda x . e)$  heißt *Abstraktion*, wobei  $x$  *Parameter* der Abstraktion heißt und  $e$  *Rumpf*. In diesem Kapitel steht  $e$  immer für einen  $\lambda$ -Term,  $v$  und  $x$  stehen für Variablen.

Es ist kein Zufall, daß die Lehrsprachen genau die gleichen Begriffe verwendet wie der  $\lambda$ -Kalkül. Ein Lambda-Ausdruck mit einem Parameter entspricht einer Abstraktion im  $\lambda$ -Kalkül, und die Applikationen in den Lehrsprachen entsprechen den Applikationen im  $\lambda$ -Kalkül. Die Lehrsprachen wurde bewußt auf dem  $\lambda$ -Kalkül aufgebaut.

Die Intuition für die Bedeutung der  $\lambda$ -Terme ist ähnlich wie in Scheme: Eine Abstraktion steht für eine mathematische Funktion, speziell für eine solche Funktion, die sich durch ein Computerprogramm berechnen läßt.<sup>1</sup> Eine Applikation steht gerade für die Applikation einer Funktion, und eine Variable bezieht sich auf den Parameter einer umschließenden Abstraktion und steht für den Operanden der Applikation.

Der einfachste  $\lambda$ -Term ist die Identität:

$$(\lambda x.x)$$

Der folgende  $\lambda$ -Term wendet eine Funktion  $f$  auf ein Argument  $x$  an:

$$(\lambda f.(\lambda x.(f\ x)))$$

An diesem Beispiel wird deutlich, daß sich im  $\lambda$ -Kalkül, wie in den Lehrsprachen auch, die Klammern schnell häufen, wenn die Terme größer werden. Darum werden redundante Klammern beim Aufschreiben von  $\lambda$ -Termen oft weggelassen. Damit wird aus dem obigen Term der folgende:

$$\lambda f.\lambda x.f\ x$$

Dieser Ausdruck läßt sich unterschiedlich klammern:  $(\lambda f.((\lambda x.f)\ x))$ ,  $(\lambda f.(\lambda x.(f\ x)))$  oder  $((\lambda f.\lambda x.f)\ x)$ . Bei solchen Mehrdeutigkeiten erstreckt sich der Rumpf einer Abstraktion so weit wie möglich nach rechts. Die richtige Variante ist also  $(\lambda f.(\lambda x.(f\ x)))$ .

Die Funktionen im  $\lambda$ -Kalkül sind auf einen Parameter beschränkt. Dies ist keine wirkliche Einschränkung: Funktionen mit mehreren Parametern werden geschönfinkelt, um aus ihnen mehrstufige Funktionen mit jeweils einem Parameter zu machen, vgl. Abschnitt 9.5.

Wegen der Verwandtschaft zwischen Funktionen mit mehreren Parametern und ihren geschönfinkelten Pendanten gibt es zwei weitere Abkürzungen in der Notation von  $\lambda$ -Termen:

- $\lambda x_1 \dots x_n.e$  steht für  $\lambda x_1.(\lambda x_2.(\dots \lambda x_n.e) \dots)$ .
- $e_0 \dots e_n$  steht für  $(\dots (e_0\ e_1)\ e_2) \dots e_n$ .

<sup>1</sup> Die Wahl des Buchstabens  $\lambda$  für die Notation von Abstraktionen war eher ein Unfall: Zur Zeit der Entstehung des Kalküls war der Ausdruck  $2\hat{x} + 1$  eine historische Notation für eine Funktion  $f$  mit  $f(x) \stackrel{\text{def}}{=} 2x + 1$ . ALONZO CHURCH, der Erfinder des  $\lambda$ -Kalküls, hatte ursprünglich die Notation  $\hat{x}.2x + 1$  in der ersten Publikation über den Kalkül vorgesehen. Der Satzsetzer konnte allerdings aus technischen Gründen das Hütchen nicht über dem  $x$  positionieren und setzte es deshalb davor, womit aus dem Ausdruck  $\hat{x}.2x + 1$  wurde. Ein weiterer Setzer machte aus dem einsamen Hütchen ein  $\lambda$  und der  $\lambda$ -Kalkül war geboren.

Dementsprechend ist  $\lambda fxy.f\ x\ y$  eine andere Schreibweise für den Term

$$(\lambda f.(\lambda x.(\lambda y.((f\ x)\ y)))) .$$

Bemerkenswert am  $\lambda$ -Kalkül ist, daß es dort *nur* Funktionen gibt, noch nicht einmal Zahlen, boolesche Werte oder zusammengesetzte Daten. Darum erscheint die Sprache des Kalküls auf den ersten Blick noch spartanisch und unintuitiv: So unmittelbar läßt sich noch nicht einmal eine Funktion hinschreiben, die zwei Zahlen addiert – schließlich gibt es keine Zahlen. Wie sich jedoch weiter unten in Abschnitt 16.3 herausstellen wird, lassen sich all diese Dinge durch Funktionen nachbilden.

Der  $\lambda$ -Kalkül selbst legt das Verhalten von  $\lambda$ -Termen fest; er ist ein Reduktionskalkül, der beschreibt, wie ein  $\lambda$ -Term in einen anderen, gleichbedeutenden, überführt werden kann. Die Konstruktion dieses Kalküls erfordert sehr sorgfältigen Umgang mit Variablen, was eine Hilfsdefinition notwendig macht:

#### DEFINITION 16.2 (FREIE UND GEBUNDENE VARIABLEN)

Die Funktionen

$$\text{free, bound} : \mathcal{L}_\lambda \rightarrow \mathcal{P}(V)$$

liefern die Mengen der *freien* beziehungsweise der *gebundenen* Variablen eines  $\lambda$ -Terms.

$$\begin{aligned} \text{free}(e) &\stackrel{\text{def}}{=} \begin{cases} \{v\} & \text{falls } e = v \\ \text{free}(e_0) \cup \text{free}(e_1) & \text{falls } e = e_0\ e_1 \\ \text{free}(e_0) \setminus \{v\} & \text{falls } e = \lambda v.e_0 \end{cases} \\ \text{bound}(e) &\stackrel{\text{def}}{=} \begin{cases} \emptyset & \text{falls } e = v \\ \text{bound}(e_0) \cup \text{bound}(e_1) & \text{falls } e = e_0\ e_1 \\ \text{bound}(e_0) \cup \{v\} & \text{falls } e = \lambda v.e_0 \end{cases} \end{aligned}$$

Außerdem ist  $\text{var}(e) \stackrel{\text{def}}{=} \text{free}(e) \cup \text{bound}(e)$  die *Menge der Variablen* von  $e$ . (Es läßt sich leicht zeigen, daß diese Menge alle vorkommenden Variablen eines  $\lambda$ -Terms enthält.) Ein  $\lambda$ -Term  $e$  heißt *abgeschlossen* beziehungsweise *Kombinator*, falls  $\text{free}(e) = \emptyset$ .

Einige Beispiele:

$$\begin{aligned}
 \text{free}(\lambda x.y) &= \{y\} \\
 \text{bound}(\lambda x.y) &= \{x\} \\
 \text{free}(\lambda y.y) &= \emptyset \\
 \text{bound}(\lambda y.y) &= \{y\} \\
 \text{free}(\lambda x.\lambda y.\lambda x.x (\lambda z.a y)) &= \{a\} \\
 \text{bound}(\lambda x.\lambda y.\lambda x.x (\lambda z.a y)) &= \{x, y, z\}
 \end{aligned}$$

In einem Term kann die gleiche Variable sowohl frei als auch gebunden vorkommen:

$$\begin{aligned}
 \text{free}(\lambda x.y (\lambda y.y)) &= \{y\} \\
 \text{bound}(\lambda x.y (\lambda y.y)) &= \{x, y\}
 \end{aligned}$$

Entscheidend ist dabei, daß das  $y$  einmal innerhalb und einmal außerhalb einer bindenden Abstraktion auftaucht. Das Frei- und Gebundensein bezieht sich also immer auf bestimmte *Vorkommen* einer Variablen in einem  $\lambda$ -Term.

Im  $\lambda$ -Kalkül gilt, genau wie in den Lehrsprachen, das Prinzip der lexikalischen Bindung (siehe Abschnitt ??): das Vorkommen einer Variable  $v$  als  $\lambda$ -Term gehört immer zur innersten umschließenden Abstraktion  $\lambda v.e$ , deren Parameter ebenfalls  $v$  ist. Bei  $\lambda x.y (\lambda y.y)$  aus dem Beispiel oben ist also das erste  $y$  das freie, während das zweite  $y$  durch die zweite Abstraktion gebunden wird.

Der  $\lambda$ -Reduktionskalkül ist darauf angewiesen, Variablen durch andere zu ersetzen, ohne dabei die Zugehörigkeit von Variablenvorkommen und den dazu passenden Abstraktionen zu verändern. Der Mechanismus dafür heißt auch hier *Substitution*:

#### DEFINITION 16.3 (SUBSTITUTION)

Für  $e, f \in \mathcal{L}_\lambda$  ist  $e[v \mapsto f]$  – in  $e$  wird  $v$  durch  $f$  *substituiert* – induktiv definiert:

$$e[v \mapsto f] \stackrel{\text{def}}{=} \begin{cases} f & \text{falls } e = v \\ x & \text{falls } e = x \text{ und } x \neq v \\ \lambda v. e_0 & \text{falls } e = \lambda v. e_0 \\ \lambda x. (e_0[v \mapsto f]) & \text{falls } e = \lambda x. e_0 \text{ und } x \neq v, x \notin \text{free}(f) \\ \lambda x'. (e_0[x \mapsto x'][v \mapsto f]) & \text{falls } e = \lambda x. e_0 \\ & \text{und } x \neq v, x \in \text{free}(f), x' \notin \text{free}(e_0) \cup \text{free}(f) \\ (e_0[v \mapsto f]) (e_1[v \mapsto f]) & \text{falls } e = e_0 e_1 \end{cases}$$

Die Definition der Substitution erscheint auf den ersten Blick kompliziert, folgt aber letztlich nur direkt dem Prinzip der lexikalischen Bindung. Die erste Regel besagt, daß das Vorkommen einer Variable durch eine Substitution genau dieser Variablen ersetzt wird:

$$v[v \mapsto f] = f$$

Die zweite Regel besagt, daß das Vorkommen einer *anderen* Variable durch die Substitution nicht betroffen wird:

$$x[v \mapsto f] = x \quad x \neq v$$

Die dritte Regel ist auf den ersten Blick etwas überraschend:

$$(\lambda v. e_0)[v \mapsto f] = \lambda v. e_0$$

Ein  $\lambda$ -Ausdruck, dessen Parameter gerade die Variable ist, die substituiert werden soll, bleibt unverändert. Das liegt daran, daß mit dem  $\lambda$ -Ausdruck die Zugehörigkeit aller Vorkommen von  $v$  in  $e_0$  bereits festgelegt ist: ein Vorkommen von  $v$  in  $e_0$  gehört entweder zu dieser Abstraktion oder einer anderen Abstraktion mit  $v$  als Parameter, die in  $e_0$  weiter innen steht –  $v$  ist in  $(\lambda v. e_0)$  gebunden und  $v \in \text{bound}(\lambda v. e_0)$ . Da die Substitution diese Zugehörigkeiten nicht verändern darf, läßt sie das  $v$  in Ruhe.

Anders sieht es aus, wenn die Variable der Abstraktion eine andere ist – die vierte Regel:

$$(\lambda x. e_0)[v \mapsto f] = \lambda x. (e_0[v \mapsto f]) \quad x \neq v, x \notin \text{free}(f)$$

In diesem Fall wird die Substitution auf den Rumpf der Abstraktion angewendet. Wichtig ist dabei, daß  $x$  nicht frei in  $f$  vorkommt – sonst könnte es vorkommen, daß beim Einsetzen von  $f$  ein freies

Vorkommen von  $x$  plötzlich durch die umschließende Abstraktion gebunden wird. Damit würde auch wieder die durch die lexikalische Bindung definierte Zugehörigkeitsregel verletzt.

Was passiert, wenn  $x$  eben doch frei in  $f$  vorkommt, beschreibt die fünfte Regel:

$$(\lambda x. e_0)[v \mapsto f] = \lambda x'. (e_0[x \mapsto x'] [v \mapsto f]) \quad \begin{array}{l} x \neq v, x \in \text{free}(f) \\ x' \notin \text{free}(e_0) \cup \text{free}(f) \end{array}$$

Hier kann es passieren, daß die freien  $x$  in  $f$  durch die Abstraktion «eingefangen» werden. Aus diesem Grund wird einfach das  $x$  in der Abstraktion aus dem Weg geschafft und durch ein «frisches»  $x'$  ersetzt, das noch nirgendwo frei vorkommt.

Die letzte Regel beschreibt schließlich, wie die Substitution auf Applikationen wirkt: sie taucht einfach in Operator und Operanden rekursiv ab:

$$(e_0 e_1)[v \mapsto f] = (e_0[v \mapsto f])(e_1[v \mapsto f])$$

Hier ist ein etwas umfangreicheres Beispiel für die Substitution:

$$\begin{aligned} (\lambda x. \lambda y. x (\lambda z. z) z)[z \mapsto x y] &= \lambda x'. ((\lambda y. x (\lambda z. z) z)[x \mapsto x'] [z \mapsto x y]) \\ &= \lambda x'. ((\lambda y. ((x (\lambda z. z) z)[x \mapsto x'])) [z \mapsto x y]) \\ &= \lambda x'. ((\lambda y. (x[x \mapsto x'] ((\lambda z. z)[x \mapsto x'])) z[x \mapsto x']) [z \mapsto x y]) \\ &= \lambda x'. ((\lambda y. (x' (\lambda z. z) z)) [z \mapsto x y]) \\ &= \lambda x'. \lambda y'. ((x' (\lambda z. z) z)[y \mapsto y'] [z \mapsto x y]) \\ &= \lambda x'. \lambda y'. ((x'[y \mapsto y'] ((\lambda z. z)[y \mapsto y']) z[y \mapsto y']) [z \mapsto x y]) \\ &= \lambda x'. \lambda y'. ((x' (\lambda z. z) z)[z \mapsto x y]) \\ &= \lambda x'. \lambda y'. x'[z \mapsto x y] ((\lambda z. z)[z \mapsto x y]) z[z \mapsto x y] \\ &= \lambda x'. \lambda y'. x' (\lambda z. z) (x y) \end{aligned}$$

Deutlich zu sehen ist, wie die freien Variablen  $x$  und  $y$  aus der Substitution  $z \mapsto x y$  auch im Ergebnis frei bleiben, während die gebundenen Variablen  $x$  und  $y$  aus dem ursprünglichen Term umbenannt werden, um eine irrtümliche Bindung ihrer hereinsubstituierten Namensvettern zu vermeiden.

Mit Hilfe der Definition der Substitution ist es möglich, die Reduktionsregeln des  $\lambda$ -Kalküls zu formulieren.

#### DEFINITION 16.4 (REDUKTIONSREGELN)



Die Reduktionsregeln im  $\lambda$ -Kalkül sind die  $\alpha$ -Reduktion  $\rightarrow_\alpha$  und die  $\beta$ -Reduktion  $\rightarrow_\beta$ :

$$\begin{aligned} \lambda x.e &\rightarrow_\alpha \lambda y.(e[x \mapsto y]) \quad y \notin \text{free}(e) \\ (\lambda v.e) f &\rightarrow_\beta e[v \mapsto f] \end{aligned}$$

Für  $x \in \{\alpha, \beta\}$  ist  $\rightarrow_x^*$  jeweils der reflexiv-transitive Abschluß der Relation. (Siehe dazu Definition A.9). Außerdem ist  $\leftrightarrow_x$  jeweils der symmetrische Abschluß, und  $\leftrightarrow_x^*$  der reflexiv-transitiv-symmetrische Abschluß.

Die  $\alpha$ -Reduktion (oft auch  $\alpha$ -Konversion genannt) benennt eine gebundene Variable in eine andere um.

Die  $\beta$ -Reduktion, die zentrale Regel des  $\lambda$ -Kalküls, steht für Funktionsapplikation: eine Abstraktion wird angewendet, indem die Vorkommen ihres Parameters durch den Operanden einer Applikation ersetzt werden.

Wie in anderen Reduktionskalkülen auch, also zum Beispiel wie in  $\text{RC}_1$ , werden die Regeln auf Subterme fortgesetzt. So gilt zum Beispiel:

$$\lambda x.(\lambda y.y) x \rightarrow_\beta \lambda x.x$$

Auch der Begriff des Redex ist im  $\lambda$ -Kalkül analog zu  $\text{RC}_1$  und bezeichnet einen reduzierbaren Subterm. Im obigen Beispiel ist der Redex gerade  $(\lambda y.y) x$ .

Als Reduktionskalkül ist die Hauptaufgabe des  $\lambda$ -Kalküls der Beweis von Gleichungen: Zwei Terme gelten als äquivalent wenn sie durch Reduktionen ineinander überführt werden können.

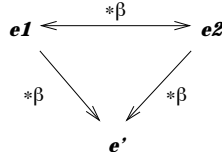
DEFINITION 16.5 (ÄQUIVALENZ IM  $\lambda$ -KALKÜL)

Zwei Terme  $e_1, e_2 \in \mathcal{L}_\lambda$  heißen  $\alpha\beta$ -äquivalent oder einfach nur äquivalent, wenn  $e_1 \leftrightarrow_{\alpha,\beta}^* e_2$  gilt, wobei  $\leftrightarrow_{\alpha,\beta} \stackrel{\text{def}}{=} \leftrightarrow_\alpha \cup \leftrightarrow_\beta$ .

Die Schreibweise dafür ist  $e_1 \equiv e_2$ .

## 16.2 NORMALFORMEN

Im  $\lambda$ -Kalkül ist es erlaubt, jederzeit beliebige Teilausdrücke zu reduzieren, solange sie nur  $\alpha$ - oder  $\beta$ -Redexe sind. Zum Beispiel gibt es für den folgenden  $\lambda$ -Term zwei verschiedene Möglichkeiten



**Abbildung 16.1:** Die Church/Rosser-Eigenschaft (FIXME)

zur  $\beta$ -Reduktion. Der gewählte Redex ist jeweils unterstrichen:

$$\begin{aligned}
 ((\lambda x. (\underline{\lambda y. y}) z) a) &\rightarrow_{\beta} (\lambda x. z) a \\
 (\underline{((\lambda x. (\lambda y. y) z) a)}) &\rightarrow_{\beta} (\lambda y. y) z
 \end{aligned}$$

In diesem Beispiel kommt eine weitere  $\beta$ -Reduktion sowohl von  $(\lambda x. z) a$  als auch von  $(\lambda y. y) z$  zum gleichen Ergebnis  $z$  – ein Indiz dafür, daß schließlich alle Abfolgen von  $\beta$ -Reduktionen zum gleichen Ergebnis kommen. Eine solche Eigenschaft eines Kalküls heißt *Normalformeigenschaft*. Hier ist die Definition des Normalformbegriffs für den  $\lambda$ -Kalkül:

DEFINITION 16.6 (NORMALFORM)

Sei  $e$  ein  $\lambda$ -Term. Ein  $\lambda$ -Term  $e'$  ist eine *Normalform* von  $e$ , wenn  $e \rightarrow_{\beta}^* e'$  gilt und kein  $\lambda$ -Term  $e''$  existiert mit  $e' \rightarrow_{\beta} e''$ .

Nun wäre es schön, wenn Normalformen dazu benutzt werden könnten, um den Beweis von Gleichungen im Kalkül zu erleichtern: Der Beweis von  $e_1 \equiv e_2$  erfordert dann lediglich den Vergleich der Normalformen von  $e_1$  und  $e_2$  – wenn diese  $\alpha$ -äquivalent sind, dann gilt  $e_1 \equiv e_2$ , sonst nicht.

Leider haben manche  $\lambda$ -Terme überhaupt keine Normalform. Hier ein Beispiel:

$$(\lambda x. x x)(\lambda x. x x) \rightarrow_{\beta} (\lambda x. x x)(\lambda x. x x)$$

Solche Terme ohne Normalformen lassen sich endlos weiterreduzieren, ohne daß der Prozeß jemals zum Schluß kommt. Sie entsprechen damit Programmen, die endlos weiterrechnen. Dies ist kein spezieller Defekt des  $\lambda$ -Kalküls: Jeder Kalkül, der mächtig genug ist, um beliebige Computerprogramme zu modellieren, hat diese Eigenschaft.

Eine wichtige Eigenschaft auf dem Weg zur Eindeutigkeit von Normalformen ist der Satz von Church/Rosser:

SATZ 16.7 (CHURCH/ROSSER-EIGENSCHAFT) Die  $\beta$ -Reduktionsregel hat die *Church/Rosser-Eigenschaft*: Für beliebige  $\lambda$ -Terme  $e_1$  und  $e_2$  mit  $e_1 \leftrightarrow_{\beta}^* e_2$ , gibt es immer einen  $\lambda$ -Term  $e'$  mit  $e_1 \rightarrow_{\beta}^* e'$  und  $e_2 \rightarrow_{\beta}^* e'$ .

Abbildung 16.1 stellt die Aussage des Satzes von Church/Rosser grafisch dar. Der Beweis des Satzes ist leider recht umfangreich und technisch. Die einschlägige Literatur über den  $\lambda$ -Kalkül hat ihn vorrätig [HINDLEY and SELDIN, 1986].

Die Church/Rosser-Eigenschaft ebnet den Weg für Benutzung von Normalformen zum Finden von Beweisen im  $\lambda$ -Kalkül:

SATZ 16.8 (EINDEUTIGKEIT DER NORMALFORM) Ein  $\lambda$ -Term  $e$  hat höchstens eine Normalform modulo  $\alpha$ -Reduktion.

**Beweis** Angenommen, es gebe zwei unterschiedliche Normalformen  $e_1$  und  $e_2$  von  $e$ . Nach Satz 16.7 muß es dann aber einen weiteren  $\lambda$ -Term  $e'$  geben mit  $e_1 \rightarrow_{\beta}^* e'$  und  $e_2 \rightarrow_{\beta}^* e'$ . Entweder sind  $e_1$  und  $e_2$  also nicht unterschiedlich, oder zumindest einer von beiden ist keine Normalform im Widerspruch zur Annahme.  $\square$

Satz 16.8 bestätigt, daß der  $\lambda$ -Kalkül ein sinnvoller Mechanismus für die Beschreibung des Verhaltens von Computerprogrammen ist: Bei einem  $\lambda$ -Term ist es gleichgültig, in welcher Reihenfolge die Reduktionen angewendet werden: Jede Reduktionsfolge, die zu einer Normalform führt, führt immer zur gleichen Normalform.

## 16.3 DER $\lambda$ -KALKÜL ALS PROGRAMMIERSPRACHE

Mit dem Normalformsatz ist geklärt, daß Terme im  $\lambda$ -Kalkül, die eine Normalform besitzen, so etwas wie einen «Sinn» haben, der unabhängig von der Reihenfolge der Reduktionsschritte ist. Bleibt die Frage, ob der  $\lambda$ -Kalkül «groß genug» ist, um Computerprogramme abzubilden.

Auf den ersten Blick erscheint das etwas unwahrscheinlich: In der Welt des  $\lambda$ -Kalküls gibt es direkt keine eingebauten booleschen Werte oder Zahlen. Diese lassen sich jedoch durch Funktionen nachbilden. Das heißt, daß der  $\lambda$ -Kalkül ebenso mächtig wie eine ausgewachsene Programmiersprache ist. Dadurch, daß er aber nur eine zentrale Reduktionsregel besitzt, eignet er sich aber viel besser als eine komplizierte Programmiersprache für die formale Manipulation.

Dieser Abschnitt zeigt, wie sich die wichtigsten Elemente einer Programmiersprache im Kalkül nachbilden lassen:

- Verzweigungen und boolesche Werte
- Zahlen
- Rekursion

### 16.3.1 VERZWEIGUNGEN

Verzweigungen haben ihre primäre Daseinsberechtigung in Verbindung mit booleschen Werten und umgekehrt. Die binäre Verzweigung in den Lehrsprachen (`if t k a`) wählt, abhängig vom Wert von  $t$ , entweder die Konsequente  $k$  oder die Alternative  $a$  aus. Die Nachbildung im  $\lambda$ -Kalkül stellt dieses Prinzip auf den Kopf: die Maschinerie für die Auswahl zwischen Konsequente und Alternative wird in die booleschen Werte selbst gesteckt. *true* ist ein  $\lambda$ -Term, der das erste von zwei Argumenten auswählt und das zweite verwirft; *false* selektiert das zweite und verwirft das erste:

$$true \stackrel{\text{def}}{=} \lambda xy.x$$

$$false \stackrel{\text{def}}{=} \lambda xy.y$$

Damit hat die Verzweigung selbst nicht mehr viel zu tun; sie wendet einfach den Test, der einen booleschen Wert ergeben muß, auf Konsequente und Alternative an:

$$if \stackrel{\text{def}}{=} \lambda txy.t x y$$

Daß *if* tatsächlich so funktioniert wie angenommen, läßt sich an einem Beispiel leicht sehen:

$$\begin{aligned} if \ true \ e_1 \ e_2 &= (\lambda txy.t x y) \ true \ e_1 \ e_2 \\ &\rightarrow_{\beta} (\lambda xy.true x y) \ e_1 \ e_2 \\ &\rightarrow_{\beta}^2 true \ e_1 \ e_2 \\ &= (\lambda xy.x) \ e_1 \ e_2 \\ &\rightarrow_{\beta} (\lambda y.e_1) \ e_2 \\ &\rightarrow_{\beta} e_1 \end{aligned}$$

Für *false* geht der Beweis analog.

## 16.3.2 NATÜRLICHE ZAHLEN

Die Nachbildung von Zahlen ist etwas komplizierter als die der booleschen Werte. Eine Methode dafür ist die Verwendung von *Church-Numeralen*. Das Church-Numeral  $\lceil n \rceil$  einer natürlichen Zahl  $n$  ist eine Funktion, die eine  $n$ -fache Applikation vornimmt.

$$\lceil n \rceil \stackrel{\text{def}}{=} \lambda f \lambda x. f^n(x)$$

Für einen  $\lambda$ -Term  $f$  ist  $f^n : \mathcal{L}_\lambda \rightarrow \mathcal{L}_\lambda$  folgendermaßen induktiv definiert:

$$f^n(e) \stackrel{\text{def}}{=} \begin{cases} e & \text{falls } n = 0 \\ f(f^{n-1}(e)) & \text{sonst} \end{cases}$$

$\lceil 0 \rceil$  ist nach dieser Definition  $\lambda f. \lambda x. x$ ,  $\lceil 1 \rceil$  ist  $\lambda f. \lambda x. f\ x$ ,  $\lceil 2 \rceil$  ist  $\lambda f. \lambda x. f(f\ x)$ , usw.

Die Nachfolgeroperation hängt eine zusätzliche Applikation an:

$$\text{succ} \stackrel{\text{def}}{=} \lambda n. \lambda f. \lambda x. n\ f\ (f\ x)$$

Der folgende Term bildet die Vorgängerfunktion ab:

$$\text{pred} \stackrel{\text{def}}{=} \lambda x. \lambda y. \lambda z. x\ (\lambda p. \lambda q. q\ (p\ y))\ ((\lambda x. \lambda y. x)\ z)\ (\lambda x. x)$$

Der Beweis dafür, daß sich  $\text{pred}$  in bezug auf  $\text{succ}$  wie die Vorgängerfunktion verhält, ist Übungsaufgabe 16.1.

In Verbindung mit den booleschen Werten läßt sich eine Zahl daraufhin testen, ob sie 0 ist:

$$\text{zerop} \stackrel{\text{def}}{=} \lambda n. n\ (\lambda x. \text{false})\ \text{true}$$

Die Funktionsweise von  $\text{zerop}$  läßt sich am einfachsten an einem Beispiel erkennen:

$$\begin{aligned} \text{zerop}\ \lceil 0 \rceil &= (\lambda n. n\ (\lambda x. \text{false})\ \text{true})\ \lceil 0 \rceil \\ &\rightarrow_\beta \lceil 0 \rceil\ (\lambda x. \text{false})\ \text{true} \\ &= (\lambda f. \lambda x. x)\ (\lambda x. \text{false})\ \text{true} \\ &\rightarrow_\beta (\lambda x. x)\ \text{true} \\ &\rightarrow_\beta \text{true} \end{aligned}$$

### 16.3.3 REKURSION UND FIXPUNKTSATZ

Schließlich fehlt noch die Rekursion. Das Hauptproblem dabei ist, daß es im  $\lambda$ -Kalkül kein Pendant zu **define** oder **letrec** gibt: Es gibt keine direkte Möglichkeit, eine rekursive Bindung herzustellen. Zur Realisierung von Rekursion ist deshalb ein Kunstgriff notwendig, der sich an der rekursiven Definition der Fakultät zeigen läßt. Schön wäre eine Definition wie folgt, wobei Zahlen ohne  $\lceil \rceil$  für ihre Church-Numerale stehen:

$$fac \stackrel{\text{def}}{=} \lambda x. if \ (zerop \ x) \ 1 \ (* \ x \ (fac \ (pred \ x)))$$

= und \* stehen dabei für  $\lambda$ -Terme, die Church-Numerale vergleichen beziehungsweise multiplizieren. (Ihre Formulierung ist Teil der Übungsaufgabe 16.2.)

Leider ist diese Formulierung von *fac* keine richtige Definition: *fac* taucht sowohl auf der linken als auch auf der rechten Seite auf. Wenn *fac* aus der rechten Seite entfernt wird, bleibt folgender Term übrig:

$$\lambda x. if \ (zerop \ x) \ 1 \ (* \ x \ (? \ (pred \ x)))$$

Immerhin ist zu sehen, daß dieser Term korrekt die Fakultät von 0 ausrechnet, nämlich 1. Für alle Zahlen größer als 0 ist es allerdings schlecht bestellt, da der Term «?» noch unbekannt ist. Weil der obige Term nur für 0 taugt, sei er mit *fac*<sub>0</sub> benannt:

$$fac_0 \stackrel{\text{def}}{=} \lambda x. if \ (zerop \ x) \ 1 \ (* \ x \ (? \ (pred \ x)))$$

Nun wäre es schön, einen Term zu haben, der zumindest auch die Fakultät von 1 ausrechnen kann. Dazu wird *fac*<sub>0</sub> in seine eigene Definition anstelle des ? eingesetzt. Das Ergebnis sieht so aus:

$$\lambda x. if \ (zerop \ x) \ 1 \ (* \ x \ (fac_0 \ (pred \ x)))$$

Da *fac*<sub>0</sub> keinen Selbstbezug enthält, läßt sich seine Definition einsetzen; das Ergebnis soll der Funktion entsprechend *fac*<sub>1</sub> heißen:

$$fac_1 \stackrel{\text{def}}{=} \lambda x. if \ (zerop \ x) \ 1 \ (* \ x \ ((\lambda x. if \ (zerop \ x) \ 1 \ (* \ x \ (? \ (pred \ x)))) \ (pred \ x)))$$

Auf die gleiche Art und Weise läßt sich ein Term konstruieren, der alle Fakultäten bis 2 ausrechnen kann:

$$fac_2 \stackrel{\text{def}}{=} \lambda x. if \ (zerop \ x) \ 1 \ (* \ x \ (fac_1 \ (pred \ x)))$$

Dieses Muster läßt sich immer so weiter fortsetzen. Leider entsteht dabei trotzdem nie ein Term, der

die Fakultäten *aller* natürlichen Zahlen berechnen kann, da die Terme immer endlich groß bleiben.

Immerhin aber enthalten alle  $fac_n$ -Terme das gleiche Muster und unterscheiden sich nur durch Aufruf von  $fac_{n-1}$ . Also ist es sinnvoll, Abstraktion das Problem anzuwenden:

$$\lambda fac. \lambda x. if \ (zerop \ x) \ 1 \ (* \ x \ (fac \ (pred \ x)))$$

Dieser Term soll  $FAC$  heißen. Nun lassen sich die  $fac_n$ -Funktionen mit Hilfe von  $FAC$  einfacher beschreiben:

$$\begin{aligned} fac_0 &\stackrel{\text{def}}{=} \lambda x. if \ (zerop \ x) \ 1 \ (* \ x \ (? \ (pred \ x))) \\ fac_1 &\stackrel{\text{def}}{=} FAC \ fac_0 \\ fac_2 &\stackrel{\text{def}}{=} FAC \ fac_1 \\ fac_3 &\stackrel{\text{def}}{=} FAC \ fac_2 \\ &\dots \end{aligned}$$

$FAC$  ist also eine Fabrik für Fakultätsfunktionen und teilt mit allen  $fac_i$  die Eigenschaft, daß ihre Definition nicht rekursiv ist.

Damit ist zwar die Notation weniger schreibintensiv geworden, aber das fundamentale Problem ist noch nicht gelöst: Eine korrekte Definition von  $fac$  müsste eine unendliche Kette von Applikationen von  $FAC$  enthalten. Da sich ein Term mit einer unendlichen Kette von Applikationen nicht aufschreiben läßt, hilft im Moment nur Wunschdenken weiter. Dafür sei angenommen,  $fac$  wäre bereits gefunden. Dann gilt folgende Gleichung:

$$fac \equiv FAC \ fac$$

Die eine zusätzliche Applikation, die  $FAC$  vornimmt, landet auf einem ohnehin schon unendlichen Stapel, macht diesen also auch nicht größer. Damit ist aber  $fac$  ein sogenannter *Fixpunkt* von  $FAC$ : Wenn  $fac$  hineingeht, kommt es auch genauso wieder heraus. Wenn es nun eine Möglichkeit gäbe, für einen  $\lambda$ -Term einen Fixpunkt zu finden, wäre das Problem gelöst. Der folgende Satz zeigt, daß dies tatsächlich möglich ist:

**SATZ 16.9 (FIXPUNKTSATZ)** Für jeden  $\lambda$ -Term  $F$  gibt es einen  $\lambda$ -Term  $X$  mit  $F \ X \equiv X$ .

**Beweis** Wähle  $X \stackrel{\text{def}}{=} Y \ F$ , wobei

$$Y \stackrel{\text{def}}{=} \lambda f. (\lambda x. f \ (x \ x)) \ (\lambda x. f \ (x \ x)).$$

$$\begin{aligned}
fac\ 3 &= Y\ FAC\ 3 \\
(\text{Satz 16.9}) \quad &\xrightarrow{*}_{\beta} FAC\ (Y\ FAC)\ 3 \\
&\rightarrow_{\beta} (\lambda x. if\ (zerop\ x)\ 1\ (*\ x\ ((Y\ FAC)\ (pred\ x))))\ 3 \\
&\rightarrow_{\beta} if\ (zerop\ 3)\ 1\ (*\ 3\ ((Y\ FAC)\ (pred\ 3))) \\
&\xrightarrow{*}_{\beta} if\ false\ 1\ (*\ 3\ ((Y\ FAC)\ 2)) \\
&\xrightarrow{*}_{\beta} *3\ ((Y\ FAC)\ 2) \\
(\text{Satz 16.9}) \quad &\xrightarrow{*}_{\beta} *3\ (FAC\ (Y\ FAC)\ 2) \\
&\rightarrow_{\beta} *3\ ((\lambda x. if\ (zerop\ x)\ 1\ (*\ x\ ((Y\ FAC)\ (pred\ x))))\ 2) \\
&\rightarrow_{\beta} *3\ (if\ (zerop\ 2)\ 1\ (*\ 2\ ((Y\ FAC)\ (pred\ 2)))) \\
&\xrightarrow{*}_{\beta} *3\ (if\ false\ 1\ (*\ 2\ ((Y\ FAC)\ 1))) \\
&\xrightarrow{*}_{\beta} *3\ (*2\ ((Y\ FAC)\ 1)) \\
(\text{Satz 16.9}) \quad &\xrightarrow{*}_{\beta} *3\ (*2\ (FAC\ (Y\ FAC)\ 1)) \\
&\rightarrow_{\beta} *3\ (*2\ ((\lambda x. if\ (zerop\ x)\ 1\ (*\ x\ ((Y\ FAC)\ (pred\ x))))\ 1)) \\
&\rightarrow_{\beta} *3\ (*2\ (if\ (zerop\ 1)\ 1\ (*\ 1\ ((Y\ FAC)\ (pred\ 1)))) \\
&\xrightarrow{*}_{\beta} *3\ (*2\ (if\ false\ 1\ (*\ 1\ ((Y\ FAC)\ 0)))) \\
&\xrightarrow{*}_{\beta} *3\ (*2\ (*1\ ((Y\ FAC)\ 0))) \\
(\text{Satz 16.9}) \quad &\xrightarrow{*}_{\beta} *3\ (*2\ (*1\ (FAC\ (Y\ FAC)\ 0))) \\
&\rightarrow_{\beta} *3\ (*2\ (*1\ ((\lambda x. if\ (zerop\ x)\ 1\ (*\ x\ ((Y\ FAC)\ (pred\ x))))\ 0))) \\
&\rightarrow_{\beta} *3\ (*2\ (*1\ (if\ (zerop\ 0)\ 1\ (*\ 1\ ((Y\ FAC)\ (pred\ 0)))))) \\
&\xrightarrow{*}_{\beta} *3\ (*2\ (*1\ (if\ true\ 1\ (*\ 1\ ((Y\ FAC)\ (pred\ 0)))))) \\
&\xrightarrow{*}_{\beta} *3\ (*2\ (*1\ 1)) \\
&\xrightarrow{*}_{\beta} 6
\end{aligned}$$

**Abbildung 16.2:** Berechnung der Fakultät von 3 im  $\lambda$ -Kalkül

Dann gilt:

$$\begin{aligned}
Y\ F &= (\lambda f. (\lambda x. f\ (x\ x))\ (\lambda x. f\ (x\ x)))\ F \\
&\rightarrow_{\beta} (\lambda x. F\ (x\ x))\ (\lambda x. F\ (x\ x)) \\
&\rightarrow_{\beta} F\ ((\lambda x. F\ (x\ x))\ (\lambda x. F\ (x\ x))) \\
&\leftarrow_{\beta} F\ ((\lambda f. (\lambda x. f\ (x\ x))\ (\lambda x. f\ (x\ x)))\ F) \\
&= F\ (Y\ F)
\end{aligned}$$

□

Der  $\lambda$ -Term  $Y$ , der Fixpunkte berechnet, heißt *Fixpunktkombinator*. Mit seiner Hilfe läßt sich die



Fakultät definieren:

$$fac \stackrel{\text{def}}{=} Y \text{ FAC}$$

Abbildung 16.2 zeigt, wie die Berechnung der Fakultät von 3 mit dieser Definition funktioniert.

## 16.4 AUSWERTUNGSSTRATEGIEN

Die Definitionen des vorangegangenen Abschnitts zusammen mit dem Satz von Church/Rosser sind wichtige Meilensteine auf dem Weg zur Verwendung des  $\lambda$ -Kalküls als Basis für reale Programmiersprachen. Leider hat die Anwendung des Satzes von Church/Rosser noch einen Haken in der Praxis: Er besagt zwar, daß sich die Äquivalenz von zwei Termen dadurch beweisen läßt, daß ihre Normalformen verglichen werden. Leider sagt er nichts darüber, wie diese Normalformen gefunden werden.

Zum systematischen Finden von Normalformen gehört eine *Auswertungsstrategie*. Eine solche Strategie ist dafür zuständig, von den  $\beta$ -Redexen innerhalb eines  $\lambda$ -Terms denjenigen auszusuchen, der tatsächlich reduziert wird. Für den  $\lambda$ -Kalkül gibt es mehrere populäre Auswertungsstrategien, die jeweils ihre eigenen Vor- und Nachteile haben, was das effektive Finden von Normalformen betrifft.

Eine populäre Auswertungsstrategie ist die Linksaußen-Reduktion, auch *normal-order reduction* oder *leftmost-outermost reduction* genannt:

DEFINITION 16.10 (LINKSAUSSEN-REDUKTION)

Die Relation  $\rightarrow_{\beta o}$ , die *Linksaußen-Reduktion*, ist durch die gleiche Regel wie die  $\beta$ -Reduktion definiert:

$$(\lambda v. e) f \rightarrow_{\beta o} e[v \mapsto f]$$

Diese Regel darf nur auf bestimmte Subterme angewendet werden, nämlich solche  $\beta$ -Redexe, die möglichst weit links außen stehen.

Die Linksaußen-Reduktion hat folgende äußerst angenehme Eigenschaft:

SATZ 16.11 Wenn  $e'$  eine Normalform von  $e$  ist, so gilt  $e \rightarrow_{\beta o}^* e'$ .

Falls es also eine Normalform gibt, so findet die Linksaußen-Reduktion sie auch.

Es gibt allerdings noch weitere Auswertungsstrategien. Die sogenannte Call-by-Name-Auswertung basiert auf dem Konzept der *schwachen Kopfnormalform*:

DEFINITION 16.12 (SCHWACHE KOPFNORMALFORM)

Unter den  $\lambda$ -Termen heißen die Abstraktionen auch *Werte* oder *schwache Kopfnormalformen*. Ein  $\lambda$ -Term, der kein Wert ist, heißt *Nichtwert*.

DEFINITION 16.13 (CALL-BY-NAME-AUSWERTUNG)

Die Relation  $\rightarrow_{\beta n}$ , die *Call-by-Name-Reduktion*, ist durch folgende Regel definiert, die wiederum identisch zur normalen Regel für  $\beta$ -Reduktion ist:

$$(\lambda v.e) f \rightarrow_{\beta n} e[v \mapsto f]$$

Diese Regel darf nur in einem Gesamtterm angewendet werden, wenn dieser noch nicht in schwacher Kopfnormalform ist, und auch dann nur auf Subterme, die  $\beta$ -Redexe sind, die möglichst weit links außen stehen.

Die Call-by-Name-Auswertung ist damit ähnlich zur Linksaußen-Auswertung, aber nicht ganz so aggressiv: sie gibt sich schon mit einer schwachen Kopfnormalform zufrieden anstatt einer «richtigen» Normalform. Dies ist bei der Verwendung als Auswertungsstrategie in Programmiersprachen allerdings schon genug: die weitere Auswertung des Rumpfes einer schwachen Kopfnormalform wird einfach verschoben auf die Zeit der Applikation.

Linksaußen- und Call-by-Name-Auswertung finden zwar immer eine Normalform beziehungsweise eine schwache Kopfnormalform, wenn es eine solche gibt; gelegentlich aber geschieht dies nicht auf die effektivste Art und Weise. Im folgendem Term wird bei Linksaußen- und Call-by-Name-Reduktion zuerst der äußere Redex reduziert:

$$\begin{aligned} (\lambda x.x x) ((\lambda y.y) z) &\rightarrow_{\beta o} ((\lambda y.y) z) ((\lambda y.y) z) \\ &\rightarrow_{\beta o} z ((\lambda y.y) z) \\ &\rightarrow_{\beta o} z z \end{aligned}$$

Bei dieser Reduktionsfolge wurde der Subterm  $((\lambda y.y) z)$  zunächst «verdoppelt» und mußte demnach auch zweimal reduziert werden. Eine andere Auswertungsstrategie verspricht die Vermeidung solcher doppelter Arbeit: Die meisten Programmiersprachen verwenden eine Strategie, die von der sogenannten *Linksinnen-Reduktion*, auch genannt *applicative-order reduction* oder *leftmost-innermost reduction* abgeleitet ist:

## DEFINITION 16.14 (LINKSINNEN-REDUKTION)

In dieser Definition steht  $w$  immer für einen Wert. Die Relation  $\rightarrow_{\beta_i}$ , die *Linksinnen-Reduktion*, ist definiert durch die folgende Regel:

$$(\lambda v.e) w \rightarrow_{\beta_i} e[v \mapsto w].$$

$\rightarrow_{\beta_i}$  ist dabei nur anwendbar auf Subterme, die möglichst weit links innen stehen.

Die Linksinnen-Reduktion ist beim obigen Beispiel effektiver, da zunächst das Argument der äußeren Applikation ausgewertet wird:

$$\begin{aligned} (\lambda x.x x) ((\lambda y.y) (\lambda z.z)) &\rightarrow_{\beta_i} (\lambda x.x x) (\lambda z.z) \\ &\rightarrow_{\beta_i} (\lambda z.z) (\lambda z.z) \\ &\rightarrow_{\beta_i} (\lambda z.z) \end{aligned}$$

Leider führt die Linksinnen-Reduktion nicht immer zu einer Normalform, selbst wenn es die Linksaußen-Reduktion tut. Der Term

$$(\lambda x.\lambda y.y) ((\lambda z.z z) (\lambda z.z z))$$

zum Beispiel hat zwei Redexe, einmal den ganzen Term und dann noch

$$(\lambda z.z z) (\lambda z.z z).$$

Die Linksinnen-Strategie wählt den inneren Subterm als ersten Redex aus:

$$(\lambda z.z z) (\lambda z.z z) \rightarrow_{\beta_i} (\lambda z.z z) (\lambda z.z z).$$

Damit läuft die Linksinnen-Reduktion unendlich im Kreis, während die Linksaußen-Reduktion sofort den gesamten Term reduziert und die Normalform  $\lambda y.y$  liefert.

Eine Ableitung der Linksinnen-Reduktion, die in den meisten Programmiersprachen Anwendung findet, ist die *Call-by-Value-Reduktion*:

## DEFINITION 16.15 (CALL-BY-VALUE-REDUKTION)

In dieser Definition steht  $w$  immer für einen Wert und  $e$  für einen Nichtwert. Die Relation  $\rightarrow_{\beta_v}$ , die *Call-by-Value-Reduktion*, ist definiert durch die folgende Regel:

$$(\lambda v.e) w \rightarrow_{\beta_v} e[v \mapsto w].$$

$\rightarrow_{\beta v}$  darf nur in einem Gesamtterm angewendet werden, wenn dieser keine schwache Kopfnormalform ist, und dann nur auf einen Subterm, der möglichst weit links innen steht.

## 16.5 DIE AUSWERTUNGSSTRATEGIE DER LEHRSPRACHEN

Mit der Call-by-Value-Reduktion ist die Grundlage für die Auswertungsstrategie der Lehrsprachen gelegt. Tatsächlich definiert das Substitutionsmodell eine Variante der Call-by-Value-Auswertung. Zur Erinnerung ist hier noch einmal die Definition der wichtigsten Regel des Substitutionsmodells, nämlich der für Funktionanwendungen der Form  $(p \ o_1 \ \dots \ o_n)$ :

[...] Zunächst werden Operator  $p$  und Operanden  $o_1, \dots, o_n$  ausgewertet. Der Wert von  $p$  muß eine Funktion sein. [...]

Der entscheidende Satz ist dabei der letzte: Er bedeutet, daß innen zuerst ausgewertet wird; treten bei der Auswertung von Operator und Operanden weitere Funktionanwendungen auf, wird das gleiche Prinzip rekursiv angewendet. Damit ist das Substitutionsmodell eng verwandt mit der Call-by-Value-Auswertung im Lambda-Kalkül.

Programmiersprachen, die von innen nach außen auswerten, heißen *Call-by-Value-Sprachen* oder *strikten Sprachen* zu sprechen. Neben den Lehrsprachen gehören auch OCaml, Clojure, Scala, Erlang, Elixir, C, Java, Pascal und ML und viele andere zu den strikten Sprachen.

Es gibt auch *nicht-strikte* Sprachen wie z.B. Haskell, die auf der sogenannten *lazy evaluation* beruhen. Ihre Auswertungsstrategie ist eng mit der Call-by-Name-Auswertung im  $\lambda$ -Kalkül verwandt. Allerdings vermeiden diese Sprachen die mehrfache überflüssige Auswertung von Ausdrücken dadurch, daß sie den Wert beim ersten Mal abspeichern und danach wiederverwenden.

## ÜBUNGSAUFGABEN

### AUFGABE 16.1

Beweise, daß *pred* den Vorgänger eines positiven Church-Numerals berechnet!

### AUFGABE 16.2

Beweise, daß es Lambda-Terme für die folgenden arithmetischen Operationen auf Church-Numeralen gibt:

$$\begin{aligned}
 add[m][n] &= [m + n] \\
 mult[m][n] &= [mn] \\
 exp[m][n] &= [m^n] \text{ für } m > 0 \\
 = [m][n] &= \begin{cases} true & \text{falls } m = n \\ false & \text{sonst} \end{cases}
 \end{aligned}$$

Benutze dazu die folgenden Definitionen:

$$\begin{aligned}
 add &\stackrel{\text{def}}{=} \lambda x. \lambda y. \lambda p. \lambda q. xp(y pq) \\
 mult &\stackrel{\text{def}}{=} \lambda x. \lambda y. \lambda z. x(yz) \\
 exp &\stackrel{\text{def}}{=} \lambda x. \lambda y. yx
 \end{aligned}$$

und gibt eine eigene Definition für  $=$  an. Dabei läßt sich die Korrektheit von *add* direkt beweisen. Für *mult* und *exp* beweise und benutze dazu folgende Hilfslemmata:

$$\begin{aligned}
 ([n]x)^m y &\leftrightarrow_{\beta} x^{nm} y \\
 [n]^m x &\leftrightarrow_{\beta} [n^m] \text{ für } m > 0
 \end{aligned}$$

### AUFGABE 16.3

Der Y-Kombinator ließe sich auch programmieren als:

```

(define y
  (lambda (f)
    ((lambda (x) (f (x x)))
     (lambda (x) (f (x x))))))

```

Zeige durch Ausprobieren, daß *y* mit dieser Definition nicht funktioniert. Warum ist das so? Benutze für die Erklärung das Substitutionsmodell! Zeige, daß die folgende Variante von *y* ein Fixpunktkombinator ist, der funktioniert:

```

(define y
  (lambda (f)

```

```

((lambda (x)
  (f (lambda (y) ((x x) y))))
 (lambda (x)
  (f (lambda (y) ((x x) y))))))

```

#### AUFGABE 16.4

(Quelle: Ralf Hinze, Bonn) Zeige, daß  $F$  mit der folgenden Definition ebenfalls ein Fixpunktkombinator ist:

$$\begin{aligned}
 F &\stackrel{\text{def}}{=} G^{[26]} \\
 G &\stackrel{\text{def}}{=} \lambda abcdefghijklmnopqrstuvwxyzr.(dasisteinfixpunktkombinator)
 \end{aligned}$$

Dabei steht  $G^{[26]}$  für den Lambda-Term, der durch 26faches Hintereinanderschreiben von  $G$  entsteht, also  $GG \dots G = (\dots ((GG)G) \dots G)$ .

## 17 DIE SECD-MASCHINE

Der  $\lambda$ -Kalkül ist als theoretisches Modell für berechenbare Funktionen lange vor der Erfindung des Computers entwickelt worden. Die Reduktionsregeln dienen dabei der Entwicklung von Beweisen über die Äquivalenz von  $\lambda$ -Termen. Damit der  $\lambda$ -Kalkül auch als Modell für die tatsächliche Ausführung von Programmen, auf dem Computer geeignet ist, fehlen noch zwei Zutaten: die direkte Definition von «eingebauten» Werten und Operationen wie Zahlen und booleschen Werten sowie ein formales Auswertungsmodell. Dieses Kapitel stellt zunächst den *angewandten  $\lambda$ -Kalkül* vor, der den normalen  $\lambda$ -Kalkül um primitive Werte und Operationen erweitert, und dann die *SECD-Maschine*, ein klassisches Auswertungsmodell für die Call-by-Value-Reduktion. Angenehmerweise läßt sich die SECD-Maschine auch als Programm implementieren, was ebenfalls in diesem Kapitel geschieht. Die SECD-Maschine kennt keine Zuweisungen; es folgt darum noch die Darstellung der *SECDH-Maschine*, die auch einen *Speicher* kennt und damit Zuweisungen korrekt modelliert.

Ab diesem Kapitel wird die Sprachebene **Schreibe Dein Programm! – fortgeschritten** benötigt.

### 17.1 DER ANGEWANDTE $\lambda$ -KALKÜL

Abschnitt 16.3 zeigte bereits, daß sich auch boolesche Werte und Zahlen im  $\lambda$ -Kalkül durch  $\lambda$ -Terme darstellen lassen. Das ist zwar aus theoretischer Sicht gut zu wissen, auf Dauer aber etwas mühsam: Darum ist es sinnvoll, mit einer erweiterten Version des  $\lambda$ -Kalküls zu arbeiten, die solche «primitiven» Werte direkt kennt. Abschnitt 16.3 hat gezeigt, daß eine solche Erweiterung nur syntaktischer Zucker ist, also die Ausdruckskraft des Kalküls nicht wirklich erhöht. Alle Erkenntnisse aus dem normalen  $\lambda$ -Kalkül bleiben also erhalten.

Ein solcher erweiterter  $\lambda$ -Kalkül heißt auch *angewandter  $\lambda$ -Kalkül*:

DEFINITION 17.1 (SPRACHE DES ANGEWANDTEN  $\lambda$ -KALKÜLS  $\mathcal{L}_{\lambda A}$ )

Sei  $V$  eine abzählbare Menge von Variablen. Sei  $B$  eine Menge von *Basiswerten*. Sei für eine natürliche Zahl  $n$  und  $i \in \{1, \dots, n\}$  jeweils  $\Sigma^i$  eine Menge von  *$i$ -stelligen Primitiva* – die Namen von «eingebauten Operationen». Jedem  $F^i \in \Sigma^i$  ist eine  $i$ -stellige Funktion  $F_B^i : B \times \dots \times B \rightarrow B$  – ihre *Operation* – zugeordnet. Die Sprache des angewandten  $\lambda$ -Kalküls, die Menge der *angewandten  $\lambda$ -Terme*,  $\mathcal{L}_{\lambda A}$ , ist durch folgende Grammatik definiert:

$$\langle \mathcal{L}_{\lambda A} \rangle \rightarrow \langle V \rangle$$

$$\begin{array}{l}
| (\langle \mathcal{L}_{\lambda A} \rangle \langle \mathcal{L}_{\lambda A} \rangle) \\
| (\lambda \langle V \rangle . \langle \mathcal{L}_{\lambda A} \rangle) \\
| \langle B \rangle \\
| (\langle \Sigma^1 \rangle \langle \mathcal{L}_{\lambda A} \rangle) \\
| (\langle \Sigma^2 \rangle \langle \mathcal{L}_{\lambda A} \rangle \langle \mathcal{L}_{\lambda A} \rangle) \\
\vdots \\
| (\langle \Sigma^n \rangle \langle \mathcal{L}_{\lambda A} \rangle \dots \langle \mathcal{L}_{\lambda A} \rangle) \quad (n\text{-mal})
\end{array}$$

Die Grammatik ist abgekürzt notiert: Die letzten Klauseln besagen, daß es für jede Stelligkeit  $i$  eine Klausel mit  $\langle \Sigma^i \rangle$  gibt, bei der jeweils  $i$  Wiederholungen von  $\langle \mathcal{L}_{\lambda A} \rangle$  – entsprechend der Stelligkeit der Primitiva in  $\Sigma^i$ . Dabei heißen Terme der Form  $(F^k e_1 \dots e_k)$  auch *primitive Applikationen*.

In diesem Kapitel dienen normalerweise die Zahlen als Basiswerte mit den üblichen Operationen wie  $+$ ,  $-$ ,  $*$ ,  $/$  etc. Damit sind Terme wie zum Beispiel  $(+ (- 5 3) 17)$  möglich.

Im angewandten  $\lambda$ -Kalkül kommen zu den Werten aus Definition 16.12 die Basiswerte dazu:

DEFINITION 17.2 (WERTE IM ANGEWANDTEN  $\lambda$ -KALKÜL)

Im angewandten  $\lambda$ -Kalkül heißen die Abstraktionen und Basiswerte kollektiv *Werte*. Ein  $\lambda$ -Term, der kein Wert ist, heißt *Nichtwert*.

Damit die primitiven Operationen auch tatsächlich eine Bedeutung bekommen, muß eine spezielle Reduktionsregel für sie eingeführt werden:

DEFINITION 17.3 ( $\delta$ -REDUKTION)

$$(F^k e_1 \dots e_k) \rightarrow_{\delta} F_B^k(e_1, \dots, e_k) \quad e_1, \dots, e_k \in B$$

Diese Regel besagt, daß eine primitive Applikation, wenn alle Operanden Werte sind, durch Anwendung der entsprechenden Operation reduziert werden kann. Damit wird z.B. der obige Beispielterm folgendermaßen reduziert:

$$(+ (- 5 3) 17) \rightarrow_{\delta} (+ 2 17) \rightarrow_{\delta} 19$$



## 17.2 DIE EINFACHE SECD-MASCHINE

Wie schon in Abschnitt 16.5 erwähnt, ist der Call-by-Value- $\lambda$ -Kalkül ein Modell für die Auswertung der Lehrsprachen und viele andere Programmiersprachen. Allerdings ist Definition 16.15 strenggenommen etwas vage: Es wird immer nur der Subterm reduziert, der «möglichst weit links innen steht», aber was das heißt, ist nicht genau definiert. Außerdem ist Reduktion zwar ein mächtiges formales Modell, entspricht aber nicht der Ausführungsmethode tatsächlicher Programmiersprachen auf echten Prozessoren. Ein präzises und echten Maschinen deutlich näheres Modell ist die *SECD-Maschine*, erfunden schon in den 60er Jahren von Peter Landin [LANDIN, 1964], und seitdem die Grundlage für zahllose Implementierungen von Call-by-Value-Sprachen. (Die Darstellung hier ist gegenüber Landins ursprünglicher Formulierung etwas modernisiert.)

Damit ein Programm aus dem angewandten  $\lambda$ -Kalkül mit der SECD-Maschine ausgewertet werden kann, muß es erst einmal in einen speziellen *Maschinencode* übersetzt oder «compiliert» werden. Der Maschinencode besteht, anders als der  $\lambda$ -Kalkül, nicht aus geschachtelten Termen, sondern aus einer Folge von *Instruktionen*.

### DEFINITION 17.4 (MASCHINENCODE)

In der folgenden Definition ist  $I$  die Menge der Instruktionen:

$$\begin{aligned} \langle I \rangle \rightarrow \langle B \rangle & \\ & | \langle V \rangle \\ & | \text{ap} \\ & | \text{prim}_{F^i} \text{ für alle } F^i \in \Sigma^i \\ & | (\langle V \rangle, \langle C \rangle) \end{aligned}$$

Ein Maschinencode-Programm ist eine Folge von Instruktionen:

$$C = I^*$$

Ein Term aus dem angewandten  $\lambda$ -Kalkül wird mit Hilfe folgender Funktion in Maschinen-

code übersetzt:

$$\llbracket \_ \rrbracket : \mathcal{L}_{\lambda A} \rightarrow C$$

$$\llbracket e \rrbracket \stackrel{\text{def}}{=} \begin{cases} b & \text{falls } e = b \in B \\ v & \text{falls } e = v \in V \\ \llbracket e_0 \rrbracket \llbracket e_1 \rrbracket \mathbf{ap} & \text{falls } e = (e_0 \ e_1) \\ \llbracket e_1 \rrbracket \dots \llbracket e_k \rrbracket \mathbf{prim}_{F^k} & \text{falls } e = (F \ e_1 \dots e_k) \\ (v, \llbracket e_0 \rrbracket) & \text{falls } e = \lambda v. e_0 \end{cases}$$

Die Übersetzungsfunktion «linearisiert» einen  $\lambda$ -Term. Zum Beispiel bedeutet die Übersetzung  $\llbracket e_0 \rrbracket \llbracket e_1 \rrbracket \mathbf{ap}$  für einen Term  $(e_0 \ e_1)$ , daß zuerst  $e_0$  ausgewertet wird, danach wird  $e_1$  ausgewertet, und schließlich wird die eigentliche Applikation ausgeführt: Entsprechend steht  $\mathbf{ap}$  für «Applikation ausführen» und  $\mathbf{prim}_{F^k}$  für «Primitiv  $F$  ausführen». Basiswerte und Variablen werden im Maschinencode belassen. Ein  $\lambda$ -Term wird übersetzt in ein Tupel aus seiner Variable und dem Maschinencode für seinen Rumpf.

Durch die Linearisierung sind die Instruktionen schon in einer Liste in der Reihenfolge ihrer Ausführung aufgereiht. Insbesondere hat die Linearisierung den Begriff «links innen» formalisiert: der jeweils am weitesten links innen stehende Redex steht in der Liste der Instruktionen vorn.

Beispiel:

$$\begin{aligned} \llbracket \lambda f. \lambda x. \lambda y. f \ (+ \ x \ (* \ y \ 2)) \rrbracket &= (f, \llbracket \lambda x. \lambda y. f \ (+ \ x \ (* \ y \ 2)) \rrbracket) \\ &= (f, (x, \llbracket \lambda y. f \ (+ \ x \ (* \ y \ 2)) \rrbracket)) \\ &= (f, (x, (y, \llbracket f \ (+ \ x \ (* \ y \ 2)) \rrbracket))) \\ &= (f, (x, (y, \llbracket f \rrbracket \llbracket (+ \ x \ (* \ y \ 2)) \rrbracket \mathbf{ap}))) \\ &= (f, (x, (y, f \llbracket (+ \ x \ (* \ y \ 2)) \rrbracket \mathbf{ap}))) \\ &= (f, (x, (y, f \llbracket x \rrbracket \llbracket (* \ y \ 2) \rrbracket \mathbf{prim}_+ \mathbf{ap}))) \\ &= (f, (x, (y, f \ x \llbracket (* \ y \ 2) \rrbracket \mathbf{prim}_+ \mathbf{ap}))) \\ &= (f, (x, (y, f \ x \llbracket y \rrbracket \llbracket 2 \rrbracket \mathbf{prim}_* \mathbf{prim}_+ \mathbf{ap}))) \\ &= (f, (x, (y, f \ x \ y \llbracket 2 \rrbracket \mathbf{prim}_* \mathbf{prim}_+ \mathbf{ap}))) \\ &= (f, (x, (y, f \ x \ y \ 2 \mathbf{prim}_* \mathbf{prim}_+ \mathbf{ap}))) \end{aligned}$$

Das Beispiel zeigt deutlich, wie der Rumpf der innersten Abstraktion in eine lineare Folge von In-

struktionen übersetzt wird, die genau der Call-by-Value-Reduktionsstrategie entspricht: erst  $f$  auswerten, dann  $x$ , dann  $y$ , dann das Primitiv  $*$  anwenden, dann  $+$ , und schließlich die Applikation durchführen.

Nun zur eigentlichen SECD-Maschine – sie funktioniert ähnlich wie ein Reduktionskalkül, operiert aber auf sogenannten *Maschinenzuständen*: die Maschine überführt also einen Maschinenzustand durch einen Auswertungsschritt in einen neuen Maschinenzustand. Ein Maschinenzustand ist dabei ein 4-Tupel aus der Menge  $S \times E \times C \times D$  (daher der Name der Maschine). Die Buchstaben sind deshalb so gewählt, weil  $S$  der sogenannte *Stack*,  $E$  die sogenannte *Umgebung* beziehungsweise auf englisch das *Environment*,  $C$  der schon bekannte Maschinencode beziehungsweise *Code* und  $D$  der sogenannte *Dump* ist. Die formalen Definitionen dieser Mengen sind wie folgt; dabei ist  $W$  die Menge der Werte:

$$\begin{aligned} S &= W^* \\ E &= \mathcal{P}(V \times W) \\ D &= (S \times E \times C)^* \\ W &= B \cup (V \times C \times E) \end{aligned}$$

Der Stack ist dabei eine Folge von Werten. In der Maschine sind dies die Werte der zuletzt ausgewerteten Terme, wobei der zuletzt ausgewertete Term vorn beziehungsweise «oben» steht. Die Umgebung ist eine partielle Abbildung von Variablen auf Werte: sie ersetzt die Substitution in der Reduktionsrelation des  $\lambda$ -Kalküls. Anstatt daß Werte für Variablen eingesetzt werden, merkt sich die Umgebung einfach, an welche Werte die Variablen gebunden sind. Erst wenn der Wert einer Variablen benötigt wird, holt ihn die Maschine aus der Umgebung. Der Dump schließlich ist eine Liste früherer Zustände der Maschine: er entspricht dem Kontext im Substitutionsmodell.

Die Menge  $W$  schließlich entspricht dem Wertebegriff aus Definition 17.2: Die Basiswerte gehören dazu, außerdem Tripel aus  $(V \times C \times E)$ . Ein solches Tripel, genannt *Closure* – repräsentiert den Wert einer Abstraktion – es besteht aus der Variable einer Abstraktion, dem Maschinencode ihres Rumpfs und der Umgebung, die notwendig ist, um die Abstraktion anzuwenden: Die Umgebung wird benötigt, damit die freien Variablen der Abstraktion entsprechend der lexikalischen Bindung ausgewertet werden können. Dies ist anders als im Substitutionsmodell, wo Variablen bei der Applikation direkt ersetzt werden und damit verschwinden. Eine Closure ist also einfach die Repräsentation einer Funktion.

Im Verlauf der Auswertung werden Umgebungen häufig um neue Bindungen von einer Variable an einen Wert erweitert. Dazu ist die Notation  $e[v \mapsto w]$  nützlich.  $e[v \mapsto w]$  konstruiert aus einer Umgebung  $e$  eine neue Umgebung, in der die Variable  $v$  an den Wert  $w$  gebunden ist. Hier

ist die Definition:

$$e[v \mapsto w] \stackrel{\text{def}}{=} (e \setminus \{(v, w') \mid (v, w') \in e\}) \cup \{(v, w)\}$$

Es wird also zunächst eine eventuell vorhandene alte Bindung entfernt und dann eine neue hinzugefügt.

Um einen  $\lambda$ -Term  $e$  in die SECD-Maschine zu «injizieren», wird er in einen Anfangszustand  $(\epsilon, \emptyset, \llbracket e \rrbracket, \epsilon)$  übersetzt. Dann wird dieser Zustand wiederholt in die Zustandsübergangsrelation  $\hookrightarrow$  gefüttert. In der folgenden Definition von  $\hookrightarrow$  sind Bezeichner mit einem Unterstrich versehen, wenn es sich um Folgen handelt, also z.B.  $\underline{s}$  für einen Stack:

$$\hookrightarrow \in \mathcal{P}((S \times E \times C \times D) \times (S \times E \times C \times D))$$

$$(\underline{s}, e, b\underline{c}, \underline{d}) \hookrightarrow (b\underline{s}, e, \underline{c}, \underline{d}) \quad (17.1)$$

$$(\underline{s}, e, v\underline{c}, \underline{d}) \hookrightarrow (e(v)\underline{s}, e, \underline{c}, \underline{d}) \quad (17.2)$$

$$(b_k \dots b_1 \underline{s}, e, \text{prim}_{F^k} \underline{c}, \underline{d}) \hookrightarrow (b\underline{s}, e, \underline{c}, \underline{d}) \quad (17.3)$$

$$\text{wobei } F^k \in \Sigma^k \text{ und } F_B^k(b_1, \dots, b_k) = b$$

$$(\underline{s}, e, (v, \underline{c}')\underline{c}, \underline{d}) \hookrightarrow ((v, \underline{c}', e)\underline{s}, e, \underline{c}, \underline{d}) \quad (17.4)$$

$$(w(v, \underline{c}', e')\underline{s}, e, \text{ap } \underline{c}, \underline{d}) \hookrightarrow (\epsilon, e'[v \mapsto w], \underline{c}', (\underline{s}, e, \underline{c})\underline{d}) \quad (17.5)$$

$$(w, e, \epsilon, (\underline{s}', e', \underline{c}')\underline{d}) \hookrightarrow (w\underline{s}', e', \underline{c}', \underline{d}) \quad (17.6)$$

Die Regeln definieren eine Fallunterscheidung nach der ersten Instruktion der Code-Komponente des Zustands, beziehungsweise greift die letzte Regel, wenn der Code leer ist. Der Reihe nach arbeiten die Regeln wie folgt:

- Regel 17.1 (die *Literalregel*) schiebt einen Basiswert direkt auf den Stack.
- Regel 17.2 (die *Variablenregel*) ermittelt den Wert einer Variable aus der Umgebung und schiebt diesen auf den Stack.
- Regel 17.3 ist die *Primitivregel*. Bei einer primitiven Applikation müssen soviele Basiswerte oben auf dem Stack liegen wie die Stelligkeit des Primitivs. Dann ermittelt die Primitivregel das Ergebnis der primitiven Applikation und schiebt es oben auf den Stack.
- Regel 17.4 ist die *Abstraktionsregel*: Das Tupel  $(v, \underline{c}')$  ist bei der Übersetzung aus einer Abstraktion entstanden. Die Regel ergänzt  $v$  und  $\underline{c}'$  mit  $e$  zu einer Closure, die auf den Stack geschoben wird.
- Regel 17.5 ist die *Applikationsregel*: Bei einer Applikation müssen oben auf dem Stack ein Wert sowie eine Closure liegen. (Zur Erinnerung: Eine Applikation kann nur ausgewertet werden, wenn eine Abstraktion vorliegt. Abstraktionen werden zu Closures ausgewertet.) In einem sol-

chen Fall «sichert» die Applikation den aktuellen Zustand auf den Dump, und die Auswertung fährt mit einem leeren Stack, der Umgebung aus der Closure – erweitert um eine Bindung für die Variable – und dem Code aus der Closure fort.

- Regel 17.6 ist die *Rückkehrregel*: Sie ist anwendbar, wenn das Ende des Codes erreicht ist. Das heißt, daß gerade die Auswertung einer Applikation fertig ist. Auf dem Dump liegt aber noch ein gesicherter Zustand, der jetzt «zurückgeholt» wird.

Hier ein Beispiel für den Ablauf der SECD-Maschine für den Term  $((\lambda x. \lambda y. (+ x y)) 1) 2$ :

$(\epsilon,$	$\emptyset,$	$(x, (y, x y \text{ prim}_+)) 1 \text{ ap } 2 \text{ ap},$	$\epsilon)$
$\hookrightarrow ((x, (y, x y \text{ prim}_+), \emptyset),$	$\emptyset,$	$1 \text{ ap } 2 \text{ ap},$	$\epsilon)$
$\hookrightarrow (1 (x, (y, x y \text{ prim}_+), \emptyset),$	$\emptyset,$	$\text{ap } 2 \text{ ap},$	$\epsilon)$
$\hookrightarrow (\epsilon,$	$\{(x, 1)\},$	$(y, x y \text{ prim}_+),$	$(\epsilon, \emptyset, 2 \text{ ap}))$
$\hookrightarrow ((y, x y \text{ prim}_+, \{(x, 1)\}),$	$\{(x, 1)\},$	$\epsilon,$	$(\epsilon, \emptyset, 2 \text{ ap}))$
$\hookrightarrow ((y, x y \text{ prim}_+, \{(x, 1)\}),$	$\emptyset,$	$2 \text{ ap},$	$\epsilon)$
$\hookrightarrow (2 (y, x y \text{ prim}_+, \{(x, 1)\}),$	$\emptyset,$	$\text{ap},$	$\epsilon)$
$\hookrightarrow (\epsilon,$	$\{(x, 1), (y, 2)\},$	$x y \text{ prim}_+,$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow (1,$	$\{(x, 1), (y, 2)\},$	$y \text{ prim}_+,$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow (2 1,$	$\{(x, 1), (y, 2)\},$	$\text{prim}_+,$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow (3,$	$\{(x, 1), (y, 2)\},$	$\epsilon,$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow (3,$	$\emptyset,$	$\epsilon,$	$\epsilon)$

Die Zustandsübergangsrelation  $\hookrightarrow$  ist nun die Grundlage für die *Auswertungsfunktion* der SECD-Maschine, die für einen  $\lambda$ -Term dessen Bedeutung ausrechnet. Dies ist scheinbar ganz einfach:

$$\begin{aligned} \text{eval}_{SECD} &: \mathcal{L}_{\lambda A} \rightarrow B \\ \text{eval}_{SECD}(e) &= x \text{ wenn } (\epsilon, \emptyset, \llbracket e \rrbracket, \epsilon) \hookrightarrow^* (x, e, \epsilon, \epsilon) \end{aligned}$$

Diese Definition hat jedoch zwei Haken:

- Die Auswertung von  $\lambda$ -Termen terminiert nicht immer (wie zum Beispiel für den «Endlos-Term»  $(\lambda x. (x x)) (\lambda x. (x x))$ ), es kommt also nicht immer dazu, daß die Zustandsübergangsrelation bei einem Zustand der Form  $(\epsilon, \emptyset, \llbracket e \rrbracket, \epsilon)$  terminiert.
- Das  $x$  aus dieser Definition ist nicht immer ein Basiswert – es kann auch eine Closure sein.

Der erste Haken sorgt dafür, daß die Auswertungsfunktion nur eine Relation im Sinne einer «partiellen Funktion» ist. Meist wird trotzdem von einer Auswertungsfunktion gesprochen. Beim zweiten Haken, wenn  $x$  eine Closure ist, läßt sich mit dem Resultat nicht viel anfangen: Um die genaue Bedeutung der Closure herauszubekommen, müsste sie angewendet werden – das Programm ist

aber schon fertig gelaufen. Es ist also gar nicht sinnvoll, zwischen verschiedenen Closures zu unterscheiden. Darum wird für die Zwecke der Auswertungsfunktion eine Menge  $Z$  der *Antworten* definiert, die einen designierten Spezialwert für Closures enthält:

$$Z = B \cup \{\text{function}\}$$

Damit läßt sich die Evaluationsfunktion wie folgt definieren:

$$\begin{aligned} eval_{SECD} &\in \mathcal{L}_{\lambda A} \times Z \\ eval_{SECD}(e) &= \begin{cases} b & \text{falls } (\epsilon, \emptyset, \llbracket e \rrbracket, \epsilon) \hookrightarrow^* (b, e, \epsilon, \epsilon) \\ \text{function} & \text{falls } (\epsilon, \emptyset, \llbracket e \rrbracket, \epsilon) \hookrightarrow^* ((v, \underline{c}, e'), e, \epsilon, \epsilon) \end{cases} \end{aligned}$$

### 17.3 QUOTE UND SYMBOLE

Dieses Kapitel wird ab hier Gebrauch von einer weiteren Sprachebene in DrRacket machen, nämlich **Die Macht der Abstraktion – fortgeschritten**. Diese Ebene muß mit dem DrRacket-Menü **Sprache** unter **Sprache** auswählen aktiviert sein, damit die Programme dieses Kapitels funktionieren.

Die entscheidende Änderung gegenüber den früheren Sprachebenen ist die Art, mit der die REPL Werte ausdrückt. (Diese neue Schreibweise, ermöglicht, die Programme des Interpreters, die als Werte repräsentiert sind, korrekt auszudrucken.) Bei Zahlen, Zeichenketten und booleschen Werten bleibt alles beim alten:

```
5
↪ 5
"Mike ist doof"
↪ "Mike ist doof"
#t
↪ #t
```

Bei Listen sieht es allerdings anders aus:

```
(list 1 2 3 4 5 6)
↪ (1 2 3 4 5 6)
```

Die REPL druckt also eine Liste aus, indem sie zuerst eine öffnende Klammer ausdrückt, dann die Listenelemente (durch Leerzeichen getrennt) und dann eine schließende Klammer.

Das funktioniert auch für die leere Liste:

```
empty
↪ ()
```

Mit der neuen Sprachebene bekommt außerdem der Apostroph, der dem Literal für die leere Liste voransteht, eine erweiterte Bedeutung. Unter anderem kann der Apostroph benutzt werden, um Literale für Listen zu formulieren:

```
'(1 2 3 4 5 6)
↪ (1 2 3 4 5 6)
'(1 #t "Mike" (2 3) "doof" 4 #f 17)
↪ (1 #t "Mike" (2 3) "doof" 4 #f 17)
'()
↪ ()
```

In der neuen Sprachebene benutzen die Literale und die ausgedruckten externen Repräsentationen für Listen also die gleiche Notation. Sie unterscheiden sich nur dadurch, daß beim Literal der Apostroph voransteht. Der Apostroph funktioniert auch bei Zahlen, Zeichenketten und booleschen Werten:

```
'5
↪ 5
'"Mike ist doof"
↪ "Mike ist doof"
'#t
↪ #t
```

Der Apostroph am Anfang eines Ausdrucks kennzeichnet diesen also als Literal. Der Wert des Literals wird genauso ausgedruckt, wie es im Programm steht. (Abgesehen von Leerzeichen und Zeilenumbrüchen.) Der Apostroph heißt auf englisch «quote», und deshalb ist diese Literalschreibweise auch unter diesem Namen bekannt. Bei Zahlen, Zeichenketten und booleschen Literalen ist auch ohne Quote klar, daß es sich um Literale handelt. Das Quote ist darum bei ihnen rein optional; sie heißen *selbstquotierend*. Bei Listen hingegen sind Mißverständnisse mit anderen zusammengesetzten Formen möglich, die ja auch mit einer öffnenden Klammer beginnen: <sup>1</sup>

<sup>1</sup> Tatsächlich ist die neue Schreibweise für externe Repräsentationen die Standard-Repräsentation in Scheme. Die früheren Sprachebenen benutzten die alternative Schreibweise, um die Verwirrung zwischen Listenliteralen und zusammengesetzten Formen zu vermeiden.

```
(1 2 3 4 5 6)
```

```
↪ procedure application: expected procedure, given: 1;  
   arguments were: 2 3 4 5 6
```

Mit der Einführung von Quote kommt noch eine völlig neue Sorte Werte hinzu: die *Symbole*. Symbole sind Werte ähnlich wie Zeichenketten und bestehen aus Text. Sie unterscheiden sich allerdings dadurch, daß sie als Literal mit Quote geschrieben und in der REPL ohne Anführungszeichen ausgedruckt werden:

```
'mike
```

```
↪ mike
```

```
'doof
```

```
↪ doof
```

Symbole lassen sich mit dem Prädikat `symbol?` von anderen Werten unterscheiden:

```
(symbol? 'mike)
```

```
↪ #t
```

```
(symbol? 5)
```

```
↪ #f
```

```
(symbol? "Mike")
```

```
↪ #f
```

Vergleichen lassen sich Symbole mit `equal?` (siehe Abbildung ??):

```
(equal? 'mike 'herb)
```

```
↪ #f
```

```
(equal? 'mike 'mike)
```

```
↪ #t
```

Symbole können nicht aus beliebigem Text bestehen. Leerzeichen sind zum Beispiel verboten. Tatsächlich entsprechen die Namen der zulässigen Symbole genau den Namen von Variablen:

```
'karl-otto
```

```
↪ karl-otto
```

```
'mehrwertsteuer
```

```
↪ mehrwertsteuer
```



```
'duftmarke
⇨ duftmarke
'lambda
⇨ lambda
'+
⇨ +
'*
⇨ *
```

Diese Entsprechung wird in diesem Kapitel noch eine entscheidene Rolle spielen. Symbole können natürlich auch in Listen und damit auch in Listenliteralen vorkommen:

```
'(karl-otto mehrwertsteuer duftmarke)
⇨ (karl-otto mehrwertsteuer duftmarke)
```

Mit Hilfe von Symbolen können Werte konstruiert werden, die in der REPL ausgedruckt wie Ausdrücke aussehen:

```
'(+ 1 2)
⇨ (+ 1 2)
'(lambda (n) (+ n 1))
⇨ (lambda (n) (+ n 1))
```

Auch wenn diese Werte wie Ausdrücke so aussehen, sind sie doch ganz normale Listen: der Wert von `'(+ 1 2)` ist eine Liste mit drei Elementen: das Symbol `+`, die Zahl `1` und die Zahl `2`. Der Wert von `'(lambda (n) (+ n 1))` ist ebenfalls eine Liste mit drei Elementen: das Symbol `lambda`, eine Liste mit einem einzelnen Element, nämlich dem Symbol `n`, und einer weiteren Liste mit drei Elementen: dem Symbol `+`, dem Symbol `n` und der Zahl `1`.

Quote hat noch eine weitere verwirrende Eigenheit:

```
''()
⇨ '()
```

Dieses Literal bezeichnet nicht die leere Liste (dann würde nur `()` ausgedruckt, ohne Quote), sondern etwas anderes:

```
(cons? ''())
⇨ #t
(first ''())
```

```

↪ quote
(rest ' '())
↪ (())

```

Der Wert des Ausdrucks `' '()` ist also eine Liste mit zwei Elementen: das erste Element ist das Symbol `quote` und das zweite Element ist die leere Liste. `'t` ist selbst also nur syntaktischer Zucker, und zwar für `(quote t)`:

```

(equal? (quote ()) '())
↪ #t
(equal? (quote (quote ())) ' '())
↪ #t

```

Quote erlaubt die Konstruktion von Literalen für viele Werte, aber nicht für alle. Ein Wert, für den Quote ein Literal konstruieren kann, heißt *repräsentierbarer Wert*. Die folgende induktive Definition spezifiziert, was ein repräsentierbarer Wert ist:

- Zahlen, boolesche Werte, Zeichenketten und Symbole sind repräsentierbare Werte.
- Eine Liste aus repräsentierbaren Werten ist ihrerseits ein repräsentierbarer Wert.
- Nichts sonst ist ein repräsentierbarer Wert.

## 17.4 IMPLEMENTIERUNG DER SECD-MASCHINE

Die SECD-Maschine ist ein Modell für die Implementierung des  $\lambda$ -Kalküls. Eine solche Implementierung lässt sich in einfach bauen – dieser Abschnitt zeigt, wie. Der grobe Fahrplan ergibt sich dabei aus der Struktur der SECD-Maschine selbst: Nach den obligatorischen Datendefinitionen müssen zunächst Terme in Maschinencode übersetzt werden. Dann kommt die Zustandsübergangsfunktion und schließlich die Auswertungsfunktion an die Reihe.

### 17.4.1 DATENANALYSE

Die erste Aufgabe ist dabei zunächst, wie immer, die Datenanalyse: Am Anfang stehen die Terme des angewandten  $\lambda$ -Kalküls. Eine geeignete Repräsentation mit Listen und Symbolen lässt dabei die Terme in der «fortgeschrittenen» Sprachebene genau wie entsprechenden Programm-Terme aussehen:

<code>(+ 1 2)</code>	steht für	$(+ 1 2)$
<code>(lambda (x) x)</code>	steht für	$\lambda x.x$
<code>((lambda (x) (x x)) (lambda (x) (x x)))</code>	steht für	$(\lambda x.(x x)) (\lambda x.(x x))$

etc.

Die Datendefinition dafür orientiert sich direkt an Definition 17.1:

Ein Lambda-Term ist eins der folgenden:

- ; - ein Symbol (für eine Variable)
- ; - eine zweielementige Liste (für eine reguläre Applikation)
- ; - eine Liste der Form `(lambda (x) e)` (für eine Abstraktion)
- ; - ein Basiswert
- ; - eine Liste mit einem Primitiv als erstem Element
- ;       (für eine primitive Applikation)

Hier die dazu passende Signatur-Definition:

```
(define term
  (signature
    (mixed symbol
      application
      abstraction
      base
      primitive-application)))
```

Die Signaturen für `application` etc. müssen noch definiert werden.

Um Verzweigungen über die Sorte `term` zu ermöglichen, müssen Prädikate für die einzelnen Teilsorten geschrieben werden. Diese können dann für die Definition der entsprechenden Signaturen benutzt werden.

```
(: application? (%a -> boolean))
(define application?
  (lambda (t)
    (and (cons? t)
      (not (equal? 'lambda (first t)))
      (not (primitive? (first t))))))

(define application (signature (predicate application?)))
```

```

; Prädikat für Abstraktionen
(: abstraction? (%a -> boolean))
(define abstraction?
  (lambda (t)
    (and (cons? t)
         (equal? 'lambda (first t)))))

(define abstraction (signature (predicate abstraction?)))

; Prädikat für primitive Applikationen
(: primitive-application? (%a -> boolean))
(define primitive-application?
  (lambda (t)
    (and (cons? t)
         (primitive? (first t)))))

(define primitive-application (signature (predicate primitive-application?)))

```

Die Definition läßt noch offen, was genau ein «Basiswert» und was ein «Primitiv» ist. Auch hierfür werden noch Datendefinitionen benötigt, zuerst für Basiswerte. Der Einfachheit halber beschränkt sich die Implementierung erst einmal auf boolesche Werte und Zahlen:

```

; Ein Basiswert ist ein boolescher Wert oder eine Zahl

```

Damit Basiswerte in Fallunterscheidungen von den anderen Arten von Termen unterschieden werden können, wird ein Prädikat benötigt:

```

; Prädikat für Basiswerte
(: base? (%a -> boolean))
(define base?
  (lambda (v)
    (or (boolean? v) (number? v))))

(define base (signature (predicate base?)))

```

Als Nächstes sind Primitive gefragt: Am obigen Beispiel ist zu erkennen, daß z.B. + ein Primitiv sein sollte. Die Datendefinition für eine kleine beispielhafte Menge von Primitiven ist wie folgt:

```

; Ein Primitiv ist eins der Symbole +, -, *, /, =

```

Da die Primitive genau wie die Variablen Symbole sind, stehen die Primitive als Variablen nicht mehr zur Verfügung: Alle Symbole, die keine Primitive sind, sind also Variablen. Das dazugehörige Prädikat ist das folgende:

```
; Prädikat für Primitive
(: primitive? (%a -> boolean))
(define primitive?
  (lambda (s)
    (or (equal? '+ s)
        (equal? '- s)
        (equal? '* s)
        (equal? '/ s)
        (equal? '= s))))

(define primitive (signature (predicate primitive?)))
```

Bevor nun ein die SECD-Maschine einen Term verarbeiten kann, muß dieser erst in Maschinencode übersetzt werden. Dabei entsteht aus Definition 17.4 direkt Daten- und Signatur-Definitionen für Instruktionen und Maschinencode:

```
; Eine Instruktion ist eins der folgenden:
; - ein Basiswert
; - eine Variable
; - eine Applikations-Instruktion
; - eine Instruktion für eine primitive Applikation
; - eine Abstraktion
(define instruction
  (signature
    (mixed base
      symbol
      ap
      tailap
      prim
      abs)))

; Eine Maschinencode-Programm ist eine Liste von Instruktionen.
(define machine-code (signature (list-of instruction)))
```

Bei der Definition von Instruktionen ist wieder einiges Wunschenken im Spiel. Basiswerte und Variablen sind wie bei den Termen. Die restlichen Fälle werden durch eigene Datendefinitionen abgebildet. Wie schon bei den leeren Bäumen sind Record-Definitionen ohne Felder im Spiel, die Fallunterscheidungen möglich machen:

```
; Eine Applikations-Instruktion ist ein Wert
;   (make-ap)
(define-record-procedures ap
  make-ap ap?
  ())
(: make-ap (-> ap))

; Die Instruktion für eine primitive Applikation
; ist ein Wert
;   (real-make-prim op arity)
; wobei op ein Symbol und arity die Stelligkeit
; ist
(define-record-procedures prim
  real-make-prim prim?
  (prim-operator prim-arity))
(: make-prim (symbol natural -> prim))

; Eine Abstraktions-Instruktion ist ein Wert
;   (make-abs v c)
; wobei v ein Symbol (für eine Variable) und c
; Maschinencode ist
(define-record-procedures abs
  make-abs abs?
  (abs-variable abs-code))
(: make-abs (symbol machine-code -> abs))
```

Da die Stelligkeit eines Primitivs dem Primitiv fest zugeordnet ist, ist eine Hilfsfunktion nützlich, die bei der Erzeugung eines Werts der Sorte `prim` die Stelligkeit ergänzt. Glücklicherweise haben alle oben eingeführten Primitive die gleiche Stelligkeit:

```
; Primitiv erzeugen
(: make-prim (symbol -> prim))
```

```
(define make-prim
  (lambda (s)
    (real-make-prim s 2)))
```

Die Einführung von Primitive mit anderen Stelligkeiten ist Gegenstand von Aufgabe 17.6.

#### 17.4.2 ÜBERSETZUNG IN MASCHINENCODE

Nun, da sowohl Terme als auch der Maschinencode Datendefinitionen haben, ist es möglich, die Übersetzung zu programmieren. Hier sind Kurzbeschreibung, Signatur und Gerüst:

```
; Term in Maschinencode übersetzen
(: term->machine-code (term -> machine-code))
(define term->machine-code
  (lambda (e)
    ...))
```

Da es sich bei `term` um gemischte Daten handelt, muß – wie immer – eine Verzweigung den Rumpf der Funktion bilden:

```
(define term->machine-code
  (lambda (e)
    (cond
      ((symbol? e) ...)
      ((application? e) ...)
      ((abstraction? e) ...)
      ((base? e) ...)
      ((primitive-application? e) ...))))
```

Die Implementierung entspricht in den einzelnen Fällen genau der Übersetzungsfunktion  $\llbracket \_ \rrbracket$ . Die Fälle für Variablen und Basiswerte sind, genau wie dort, trivial:

```
(define term->machine-code
  (lambda (e)
    (cond
      ((symbol? e) (list e))
      ((base? e) (list e))
      ...)))
```

Bei regulären Applikationen werden Operator und Operand übersetzt, und das ganze zusammen mit einer `ap`-Instruktion zu einer Liste zusammengesetzt:

```
(define term->machine-code
  (lambda (e)
    (cond
      ...
      ((application? e)
       (append (term->machine-code (first e))
                (append (term->machine-code (first (rest e)))
                        (list (make-ap))))))
    ...)))
```

Bei den primitiven Applikationen werden erst einmal die Operanden in Maschinencode übersetzt, die Resultate aneinandergehängt, und schließlich kommt noch eine `prim`-Instruktion ans Ende:

```
(define term->machine-code
  (lambda (e)
    (cond
      ...
      ((primitive-application? e)
       (append
        (append-lists
         (map term->machine-code (rest e)))
        (list (make-prim (first e)))))
      ...)))
```

Dieses Stück Code benutzt die Hilfsfunktion `append-lists`, die aus einer Liste von Listen eine einzelne Liste macht, indem die Elemente aneinandergehängt werden:

```
; die Elemente einer Liste von Listen aneinanderhängen
(: append-lists ((list-of (list-of %a)) -> (list-of %a)))
(define append-lists
  (lambda (l)
    (fold '() append l)))
```

Zurück zur Übersetzung: Eine Abstraktionen wird direkt in eine `abs`-Instruktion übersetzt, wobei der Rumpf selbst noch in Maschinencode übersetzt wird:



```
(define term->machine-code
  (lambda (e)
    (cond
      ...
      ((abstraction? e)
       (list
        (make-abs (first (first (rest e)))
                  (term->machine-code
                   (first (rest (rest e)))))))))))
```

### 17.4.3 ZUSTANDSÜBERGANG UND AUSWERTUNG

Da nun alle  $\lambda$ -Terme in Maschinencode-Programme übersetzt werden können, ist jetzt die eigentliche SECD-Maschine an der Reihe. Hier sind erst einmal einige neue Datendefinitionen fällig. Zunächst einmal die Menge  $S$  der Stacks:

```
; Ein Stack ist eine Liste von Werten
(define stack (signature (list-of value)))
```

Die Definition von Werten  $W$  kommt etwas später an die Reihe.

Umgebungen aus der Menge  $E$  sind mathematisch gesehen Mengen aus Tupeln. In der Implementierung werden sie dargestellt aus Listen von *Bindungen*, wobei jede Bindung einem Tupel aus der mathematischen Definition entspricht:

```
; Eine Umgebung ist eine Liste von Bindungen.
; Dabei gibt es für jede Variable nur eine Bindung.
(define environment (signature (list-of binding)))

; Eine Bindung (Name: binding) ist ein Wert
; (make-binding v x)
; wobei v der Name einer Variablen und x der dazugehörige Wert ist.
```

```
(define-record-procedures binding
  make-binding binding?
  (binding-variable binding-value))
(: make-binding (symbol value -> binding))
```

Die leere Umgebung wird öfter benötigt und wird darum schon vordefiniert:

```
; die leere Umgebung
(define the-empty-environment empty)
```

Zwei Operationen gibt es für eine Umgebung  $e$ : die Erweiterung um eine Bindung  $e[v \mapsto w]$  und das Nachschauen einer Bindung  $e(v)$ . Zunächst die Erweiterung: die Implementierung entspricht genau der mathematischen Definition: zunächst wird eine eventuell vorhandene Bindung für  $v$  entfernt, dann eine neue Bindung hinzugefügt:

```
; eine Umgebung um eine Bindung erweitern
(: extend-environment (environment symbol value -> environment))
(define extend-environment
  (lambda (e v w)
    (cons (make-binding v w)
          (remove-environment-binding e v))))
```

Für das Entfernen der alten Bindung ist die Hilfsfunktion `remove-environment-binding` zuständig. Sie folgt einmal mehr strikt der Konstruktionsanleitung für Funktionen, die Listen akzeptieren:

```
; die Bindung für eine Variable aus einer Umgebung entfernen
(: remove-environment-binding (environment symbol -> environment))
(define remove-environment-binding
  (lambda (e v)
    (cond
      ((empty? e) empty)
      ((cons? e)
       (if (equal? v (binding-variable (first e)))
           (rest e)
           (cons (first e)
                  (remove-environment-binding (rest e) v)))))))
```

Auch die zweite Operation, das Nachschauen einer Bindung in der Umgebung, folgt der Konstruktionsanleitung:

```
; die Bindung für eine Variable in einer Umgebung finden
(: lookup-environment (environment symbol -> value))
(define lookup-environment
```

```

(lambda (e v)
  (cond
    ((empty? e) (violation "unbound variable"))
    ((cons? e)
     (if (equal? v (binding-variable (first e)))
         (binding-value (first e))
         (lookup-environment (rest e) v))))))

```

Damit sind die Operationen auf Umgebungen abgeschlossen. Als Nächstes sind Dumps an der Reihe:  $D$  ist als Folge von Tupeln  $S \times E \times C$  definiert, auch genannt *Frames*. Hier sind Daten- und Record-Definition:

```

; Ein Dump ist eine Liste von Frames

; Ein Frame ist ein Wert
; (make-frame s e c)
; wobei s ein Stack, e eine Umgebung und c Maschinencode ist.
(define-record-procedures frame
  make-frame frame?
  (frame-stack frame-environment frame-code))
(: make-frame (stack environment machine-code -> frame))

```

Schließlich fehlt noch eine Repräsentation für die Menge  $W$  der Werte: Ein Wert ist entweder ein Basiswert oder eine Closure. Basiswerte wurden bereits in Abschnitt 17.4.1 definiert; es fehlen noch Closures, die Tupel aus  $V \times C \times E$  sind. Hier sind die entsprechenden Definitionen:

```

; Ein SECD-Wert ist ein Basiswert oder eine Closure
(define value (signature (mixed base closure)))

; Eine Closure ist ein Wert
; (make-closure v c e)
; wobei v die Variable der Lambda-Abstraktion,
; c der Code der Lambda-Abstraktion
; und e ein Environment ist.
(define-record-procedures closure
  make-closure closure?
  (closure-variable closure-code closure-environment))
(: make-closure (symbol machine-code environment -> closure))

```

Mit Hilfe dieser Definitionen ist es möglich, eine Daten- und eine Record-Definition für die Zustände der SECD-Maschine anzugeben, also die Tupel aus  $S \times E \times C \times D$ :

```
; Ein SECD-Zustand ist ein Wert
; (make-secd s e c d)
; wobei s ein Stack, e eine Umgebung, c Maschinencode
; und d ein Dump ist
(define-record-procedures secd
  make-secd secd?
  (secd-stack secd-environment secd-code secd-dump))
(: make-secd (stack environment machine-code dump -> secd))
```

Damit kann es an die Zustandsübergangsfunktion gehen. Sie wird als Funktion realisiert, die einen SECD-Zustand akzeptiert und einen neuen liefert. Hier sind Kurzbeschreibung, Signatur und Gerüst:

```
; Zustandsübergang berechnen
(: secd-step (secd -> secd))
(define secd-step
  (lambda (state)
    ...))
```

Entsprechend den Regeln der SECD-Maschine muß der Rumpf der Funktion eine Verzeigung zwischen den verschiedenen Fällen bei der Code-Komponente von `state` sein. Diese folgen den Konstruktionsanleitungen für Listen und für gemischte Daten. Es ist bereits an den Regeln abzulesen, daß alle Regeln Zugriff auf die Komponenten von `state` benötigen. Für diese werden gleich am Anfang lokale Variablen angelegt:

```
(define secd-step
  (lambda (state)
    (let ((stack (secd-stack state))
          (environment (secd-environment state))
          (code (secd-code state))
          (dump (secd-dump state)))
      (cond
        ((cons? code)
         (cond
          ((base? (first code)) ...)
```

```

      ((symbol? (first code)) ...)
      ((prim? (first code)) ...)
      ((abs? (first code)) ...)
      ((ap? (first code)) ...)))
    ((empty? code) ...))))

```

In diesem Gerüst werden nun die Regeln direkt abgebildet. Hier zur Erinnerung noch einmal die erste Regel für Basiswerte:

$$(\underline{s}, e, \underline{bc}, \underline{d}) \hookrightarrow (b\underline{s}, e, \underline{c}, \underline{d})$$

Hier der passende Code dafür:

```

(define secd-step
  (lambda (state)
    ...
    (cond
      ((base? (first code))
       (make-secd (cons (first code) stack)
                   environment
                   (rest code)
                   dump)))
      ...))
  ...))

```

Hier die Regel für Variablen:

$$(\underline{s}, e, v\underline{c}, \underline{d}) \hookrightarrow (e(v)\underline{s}, e, \underline{c}, \underline{d})$$

Hier der entsprechende Code:

```

(define secd-step
  (lambda (state)
    ...
    (cond
      ((symbol? (first code))
       (make-secd (cons
                    (lookup-environment environment (first code))
                    stack)
                   environment
                   ...)))
      ...))

```

```

                (rest code)
            dump))
        ...)
    ...))

```

Die Regel für primitive Applikationen ist etwas aufwendiger:

$$(b_k \dots b_1 \underline{s}, e, \text{prim}_{F^k} \underline{c}, \underline{d}) \hookrightarrow (b \underline{s}, e, \underline{c}, \underline{d})$$

wobei  $F^k \in \Sigma^k$  und  $F_B(b_1, \dots, b_k) = b$

Für die Implementierung werden Hilfsfunktionen gebraucht, welche die Argumente vom Stack holen und in der Reihenfolge umdrehen, die Argumente vom Stack entfernen und schließlich die eigentliche  $\delta$ -Transition berechnen:

```

(define secd-step
  (lambda (state)
    ...
    (cond
      ...
      ((prim? (first code))
       (make-secd (cons
                    (apply-primitive
                     (prim-operator (first code))
                     (take-reverse (prim-arity (first code)) stack))
                    (drop (prim-arity (first code)) stack))
                  environment
                  (rest code)
                  dump)))
      ...))
    ...))

```

Die Funktion drop ist gerade die in Aufgabe ?? geforderte Funktion:

```

; die ersten Elemente einer Liste weglassen
(: drop (natural (list-of %a) -> (list-of %a)))

```

Die take-reverse-Funktion ist das Pendant zu drop, das die ersten  $n$  Elemente einer Liste in umgekehrter Reihenfolge liefert. Dies ist am einfachsten über eine endrekursive Hilfsfunktion zu

erledigen – aus Kapitel 10 ist ja bekannt, daß bei endrekursiver Konstruktion von Listen gerade immer die Reihenfolge umgedreht wird:

```
; die ersten Elemente einer Liste in umgekehrter Reihenfolge berechnen
(: take-reverse (natural (list-of %a) -> (list-of %a)))
(define take-reverse
  (lambda (n l)
    ;; (: loop (natural (list-of %a) (list-of %a) -> (list-of %a)))
    (letrec ((loop (lambda (n l r)
                      (if (= n 0)
                          r
                          (loop (- n 1) (rest l) (cons (first l) r))))))
      (loop n l '()))))
```

Aus einem Primitiv und einer Liste von Argumenten berechnet `apply-primitive` das Resultat der primitiven Applikation. Dabei handelt es sich bei `primitive` um eine Fallunterscheidung, der Rumpf der Funktion ist also eine entsprechende Verzweigung:

```
; Delta-Transition berechnen
(: apply-primitive (primitive (list-of value) -> value))
(define apply-primitive
  (lambda (p args)
    (cond
      ((equal? p '+)
       (+ (first args) (first (rest args))))
      ((equal? p '-')
       (- (first args) (first (rest args))))
      ((equal? p '=)
       (= (first args) (first (rest args))))
      ((equal? p '*)
       (* (first args) (first (rest args))))
      ((equal? p '/')
       (/ (first args) (first (rest args))))))
```

Die Regel für Abstraktionen macht aus einer Abstraktion eine Closure:

$$(\underline{s}, e, (v, \underline{c'})\underline{c}, \underline{d}) \hookrightarrow ((v, \underline{c'}, e)\underline{s}, e, \underline{c}, \underline{d})$$

Der Code macht dies genauso:





```

                                environment (rest code))
                                dump))))
                                ...))
                                ...))

```

Schließlich bleibt noch der Code für die Rückgabe eines Wertes von einer Funktion. Hier ist die Regel:

$$(w, e, \epsilon, (\underline{s'}, e', \underline{c'}) \underline{d}) \hookrightarrow (w \underline{s'}, e', \underline{c'}, \underline{d})$$

Hier ist der Code dazu:

```

(define secd-step
  (lambda (state)
    ...
    (cond
      ...
      ((empty? code)
       (let ((f (first dump)))
         (make-secd
          (cons (first stack)
                (frame-stack f))
          (frame-environment f)
          (frame-code f)
          (rest dump))))))
    ...))

```

Damit die SECD-Maschine in Betrieb genommen werden kann, muß ein Term  $e$  noch in einen Anfangszustand  $(\epsilon, \emptyset, \llbracket e \rrbracket, \epsilon)$  übersetzt werden. Das erledigt folgende Hilfsfunktion:

```

; Aus Term SECD-Anfangszustand machen
(: inject-secd (term -> secd))
(define inject-secd
  (lambda (e)
    (make-secd empty
                the-empty-environment
                (term->machine-code e)
                empty)))

```

Damit läßt sich die Maschine schon ausprobieren:

```

(secd-step (inject-secd '(+ 1 2)))
↪ #<record:secd (1) () (2 #<record:prim + 2>) ()>
(secd-step (secd-step (inject-secd '(+ 1 2))))
↪ #<record:secd (2 1) () (#<record:prim + 2>) ()>
(secd-step (secd-step (secd-step (inject-secd '(+ 1 2)))))
↪ #<record:secd (3) () () ()>

```

Es fehlt noch die Auswertungsfunktion  $\text{eval}_{\text{SECD}}$ , die eine Hilfsfunktion benötigt, um die reflexiv-transitive Hülle des Zustandsübergangs  $\hookrightarrow^*$  benötigt:

```

; bis zum Ende Zustandsübergänge berechnen
(: secd-step* (secd -> secd))
(define secd-step*
  (lambda (state)
    (if (and (empty? (secd-code state))
             (empty? (secd-dump state)))
        state
        (secd-step* (secd-step state)))))

```

Die Auswertungsfunktion orientiert sich direkt an der mathematischen Definition:

```

; Evaluationsfunktion zur SECD-Maschine berechnen
(: eval-secd (term -> (mixed value (enum 'function))))
(define eval-secd
  (lambda (e)
    (let ((val (first
                  (secd-stack
                   (secd-step*
                    (inject-secd e)))))))
      (if (base? val)
          val
          'proc))))

```

Damit läuft die SECD-Maschine:

```

(eval-secd '(((lambda (x) (lambda (y) (+ x y))) 1) 2))
↪ 3

```

## 17.5 DIE ENDREKURSIVE SECD-MASCHINE

Die SECD-Maschine hat einen Schönheitsfehler: Bei endkursiven Applikationen sollte sie eigentlich, wie in Lehrsprachen-Programmen, keinerlei zusätzlichen Platz verbrauchen, da kein Kontext anfällt. Folgende Beispielauswertung für den Term  $(\lambda x.x x) (\lambda x.x x)$  zeigt aber, daß der Zustand mit fortschreitender Auswertung immer größer wird:

$(\epsilon,$	$\emptyset,$	$(x, x x \text{ ap}) (x, x x \text{ ap}) \text{ ap},$	$\epsilon)$
$\hookrightarrow ((x, x x \text{ ap}, \emptyset),$	$\emptyset,$	$(x, x x \text{ ap}) \text{ ap},$	$\epsilon)$
$\hookrightarrow ((x, x x \text{ ap}, \emptyset) (x, x x \text{ ap}, \emptyset),$	$\emptyset,$	$\text{ap},$	$\epsilon)$
$\hookrightarrow (\epsilon,$	$\{(x, (x, x x \text{ ap}, \emptyset))\},$	$x x \text{ ap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow ((x, x x \text{ ap}, \emptyset),$	$\{(x, (x, x x \text{ ap}, \emptyset))\},$	$x \text{ ap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow ((x, x x \text{ ap}, \emptyset) (x, x x \text{ ap}, \emptyset),$	$\{(x, (x, x x \text{ ap}, \emptyset))\},$	$\text{ap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow (\epsilon,$	$\{(x, (x, x x \text{ ap}, \emptyset))\},$	$x x \text{ ap},$	$(\epsilon, \{(x, (x, x x \text{ ap}, \emptyset))\}, \epsilon) (\epsilon, \emptyset, \epsilon))$
$\hookrightarrow ((x, x x \text{ ap}, \emptyset),$	$\{(x, (x, x x \text{ ap}, \emptyset))\},$	$x \text{ ap},$	$(\epsilon, \{(x, (x, x x \text{ ap}, \emptyset))\}, \epsilon) (\epsilon, \emptyset, \epsilon))$
$\hookrightarrow ((x, x x \text{ ap}, \emptyset) (x, x x \text{ ap}, \emptyset),$	$\{(x, (x, x x \text{ ap}, \emptyset))\},$	$\text{ap},$	$(\epsilon, \{(x, (x, x x \text{ ap}, \emptyset))\}, \epsilon) (\epsilon, \emptyset, \epsilon))$
$\hookrightarrow (\epsilon,$	$\{(x, (x, x x \text{ ap}, \emptyset))\},$	$x x \text{ ap},$	$(\epsilon, \{(x, (x, x x \text{ ap}, \emptyset))\}, \epsilon) (\epsilon, \{(x, (x, x x \text{ ap}, \emptyset))\}, \epsilon) (\epsilon,$
$\hookrightarrow ((x, x x \text{ ap}, \emptyset),$	$\{(x, (x, x x \text{ ap}, \emptyset))\},$	$x \text{ ap},$	$(\epsilon, \{(x, (x, x x \text{ ap}, \emptyset))\}, \epsilon) (\epsilon, \{(x, (x, x x \text{ ap}, \emptyset))\}, \epsilon) (\epsilon,$
$\dots$			

Damit ist die SECD-Maschine, so wie ist, als Ausführungsmodell für die Lehrsprachen ungeeignet. Dieses Manko läßt sich zum Glück reparieren: Die SECD-Maschine muß endrekursive und «normale» Applikationen unterschiedlich behandeln. Dazu wird eine neue Instruktion namens **tailap** eingeführt, die wie **ap** eine Applikation durchführt, aber eine endrekursive Applikation signalisiert:

$$I = \dots \cup \{\text{tailap}\}$$

Als Nächstes muß die Übersetzungsfunktion von Termen in Maschinencode geändert werden: Applikationen, die Kontext um sich herum haben, sollen mit **ap** übersetzt werden, solche ohne Kontext mit **tailap**. Da der Applikation allein der Kontext nicht anzusehen ist, sondern nur dem Term «drumherum», wird die Übersetzungsfunktion  $\llbracket \cdot \rrbracket$  in zwei Teile aufgespalten: für einen Term  $e$  wird die Auswertungsfunktion  $\llbracket \cdot \rrbracket$  immer dann benutzt, wenn um  $e$  Kontext steht. Eine weitere Funktion  $\llbracket \cdot \rrbracket'$  wird immer dann aufgerufen, wenn *kein* Kontext drumherum steht.

Kontext entsteht seinerseits immer durch Funktionsapplikationen. Bei der Auswertung eines Terms  $(e_0 e_1)$  muß *nach*  $e_0$  noch  $e_1$  ausgewertet werden, und nach Auswertung von  $e_1$  muß noch die Applikation durchgeführt werden. Sowohl  $e_0$  als auch  $e_1$  stehen in Kontext. Ähnlich ist es bei den Argumenten von primitiven Applikationen.

Auf der anderen Seite schneiden Abstraktionen für ihren Rumpf den Kontext erst einmal ab: Der Rumpf einer Abstraktion kommt schließlich bei der Auswertung der Abstraktion noch gar nicht zum Zug. Ob er Kontext hat oder nicht, entscheidet sich erst bei der Applikation. Dement-

sprechend schalten Applikationen und Abstraktionen zwischen den beiden Funktionen  $\llbracket \cdot \rrbracket$  und  $\llbracket \cdot \rrbracket'$  hin und her:

$$\begin{aligned}
 \llbracket \cdot \rrbracket &: \mathcal{L}_{\lambda A} \rightarrow C \\
 \llbracket e \rrbracket &\stackrel{\text{def}}{=} \begin{cases} b & \text{falls } e = b \in B \\ v & \text{falls } e = v \in V \\ \llbracket e_0 \rrbracket \llbracket e_1 \rrbracket \mathbf{ap} & \text{falls } e = (e_0 e_1) \\ \llbracket e_1 \rrbracket \dots \llbracket e_k \rrbracket \mathbf{prim}_{F^k} & \text{falls } e = (F e_1 \dots e_k) \\ (v, \llbracket e_0 \rrbracket') & \text{falls } e = \lambda v. e_0 \end{cases} \\
 \llbracket \cdot \rrbracket' &: \mathcal{L}_{\lambda A} \rightarrow C \\
 \llbracket e \rrbracket' &\stackrel{\text{def}}{=} \begin{cases} b & \text{falls } e = b \in B \\ v & \text{falls } e = v \in V \\ \llbracket e_0 \rrbracket \llbracket e_1 \rrbracket \mathbf{tailap} & \text{falls } e = (e_0 e_1) \\ \llbracket e_1 \rrbracket \dots \llbracket e_k \rrbracket \mathbf{prim}_{F^k} & \text{falls } e = (F e_1 \dots e_k) \\ (v, \llbracket e_0 \rrbracket') & \text{falls } e = \lambda v. e_0 \end{cases}
 \end{aligned}$$

Die Übersetzungsfunktion hat die eigentliche Arbeit geleistet: Jetzt muß nur noch eine Zustandsübergangsregel her, die **tailap** verarbeitet. Diese ergibt sich direkt aus den Regeln für **ap** und die Rückgabe eines Wertes: **tailap** funktioniert so, wie **ap** direkt gefolgt von der Rückgaberegeln. Hier sind die beiden Regeln noch einmal zur Erinnerung:

$$\begin{aligned}
 (w(v, \underline{c}', e') \underline{s}, e, \mathbf{ap} \underline{c}, \underline{d}) &\hookrightarrow (\epsilon, e'[v \mapsto w], \underline{c}', (\underline{s}, e, \underline{c}) \underline{d}) \\
 (w, e, \epsilon, (\underline{s}', e', \underline{c}') \underline{d}) &\hookrightarrow (w \underline{s}', e', \underline{c}', \underline{d})
 \end{aligned}$$

Da die erste Regel ein neues Dump-Frame erzeugt und die zweite ein Dump-Frame «vernichtet», entfällt diese Arbeit in der Regel für **tailap**:

$$(w(v, \underline{c}', e') \underline{s}, e, \mathbf{tailap} \underline{c}, \underline{d}) \hookrightarrow (\underline{s}, e'[v \mapsto w], \underline{c}', \underline{d})$$

Damit läuft das Beispiel zwar immer noch endlos, aber immerhin, ohne immer mehr Platz zu ver-

brauchen:

$\hookrightarrow (\epsilon,$	$\emptyset,$	$(x, x x \text{ tailap}) (x, x x \text{ tailap}) \text{ ap},$	$\epsilon)$
$\hookrightarrow ((x, x x \text{ tailap}, \emptyset),$	$\emptyset,$	$(x, x x \text{ tailap}) \text{ ap},$	$\epsilon)$
$\hookrightarrow ((x, x x \text{ tailap}, \emptyset) (x, x x \text{ tailap}, \emptyset),$	$\emptyset,$	$\text{ap},$	$\epsilon)$
$\hookrightarrow (\epsilon,$	$\{(x, (x, x x \text{ tailap}, \emptyset))\},$	$x x \text{ tailap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow ((x, x x \text{ tailap}, \emptyset),$	$\{(x, (x, x x \text{ tailap}, \emptyset))\},$	$x \text{ tailap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow ((x, x x \text{ tailap}, \emptyset) (x, x x \text{ tailap}, \emptyset),$	$\{(x, (x, x x \text{ tailap}, \emptyset))\},$	$\text{tailap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow (\epsilon,$	$\{(x, (x, x x \text{ tailap}, \emptyset))\},$	$x x \text{ tailap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow ((x, x x \text{ tailap}, \emptyset),$	$\{(x, (x, x x \text{ tailap}, \emptyset))\},$	$x \text{ tailap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow ((x, x x \text{ tailap}, \emptyset) (x, x x \text{ tailap}, \emptyset),$	$\{(x, (x, x x \text{ tailap}, \emptyset))\},$	$\text{tailap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow (\epsilon,$	$\{(x, (x, x x \text{ tailap}, \emptyset))\},$	$x x \text{ tailap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow ((x, x x \text{ tailap}, \emptyset),$	$\{(x, (x, x x \text{ tailap}, \emptyset))\},$	$x \text{ tailap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow ((x, x x \text{ tailap}, \emptyset) (x, x x \text{ tailap}, \emptyset),$	$\{(x, (x, x x \text{ tailap}, \emptyset))\},$	$\text{tailap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow (\epsilon,$	$\{(x, (x, x x \text{ tailap}, \emptyset))\},$	$x x \text{ tailap},$	$(\epsilon, \emptyset, \epsilon))$
$\hookrightarrow ((x, x x \text{ tailap}, \emptyset),$	$\{(x, (x, x x \text{ tailap}, \emptyset))\},$	$x \text{ tailap},$	$(\epsilon, \emptyset, \epsilon))$

Die Implementierung der endrekursiven SECD-Maschine ist Gegenstand von Übungsaufgabe 17.3.

## 17.6 DER $\lambda$ -KALKÜL MIT ZUSTAND

Der bisher vorgestellte  $\lambda$ -Kalkül liefert keinerlei Erklärung für das Verhalten von Zuweisungen. Tatsächlich hat sich schon in Abschnitt ?? angedeutet, daß Zuweisungen die Formalisierung deutlich erschweren. Möglich ist es trotzdem, und dieser Abschnitt zeigt, wie es geht.

Als erstes muß wieder einmal die Sprache des  $\lambda$ -Kalküls erweitert werden, diesmal um **set !**-Ausdrücke:

DEFINITION 17.5 (SPRACHE DES ANGEWANDTEN  $\lambda$ -KALKÜLS MIT ZUSTAND  $\mathcal{L}_{\lambda S}$ )

Sei  $V$  eine abzählbare Menge von Variablen. Sei  $B$  eine Menge von *Basiswerten* mit  $\text{void} \in B$ . Sei für eine natürliche Zahl  $n$  und  $i \in \{1, \dots, n\}$  jeweils  $\Sigma^i$  eine Menge von *i-stelligen Primitiva*. Jedem  $F^i \in \Sigma^i$  ist eine *i-stellige Funktion*  $F_B^i : B \times \dots \times B \rightarrow B$  – ihre *Operation* – zugeordnet. Sei  $A$  eine abzählbare Menge von Adressen mit  $V \cap A = \emptyset$ .

Die Sprache des angewandten  $\lambda$ -Kalküls mit Zustand, die Menge der *angewandten  $\lambda$ -Terme mit Zustand*,  $\mathcal{L}_{\lambda S}$ , ist durch folgende Grammatik definiert:

$$\begin{aligned}
 \langle \mathcal{L}_{\lambda S} \rangle &\rightarrow \langle V \rangle \\
 &\mid \langle \langle \mathcal{L}_{\lambda S} \rangle \ \langle \mathcal{L}_{\lambda S} \rangle \rangle \\
 &\mid \langle \lambda \langle V \rangle . \langle \mathcal{L}_{\lambda S} \rangle \rangle \\
 &\mid \langle B \rangle \\
 &\mid \langle \Sigma^1 \rangle \ \langle \mathcal{L}_{\lambda S} \rangle
 \end{aligned}$$

$$\begin{array}{l}
| (\langle \Sigma^2 \rangle \langle \mathcal{L}_{\lambda S} \rangle \langle \mathcal{L}_{\lambda S} \rangle) \\
\ldots \\
| (\langle \Sigma^n \rangle \langle \mathcal{L}_{\lambda S} \rangle \ldots \langle \mathcal{L}_{\lambda S} \rangle) \quad (n\text{-mal}) \\
| (\text{set! } \langle V \rangle \langle \mathcal{L}_{\lambda S} \rangle)
\end{array}$$

Der `void`-Wert wird als Rückgabewert von `set!`-Ausdrücken dienen.

Um Reduktionsregeln für Zuweisungen zu bilden, ist es notwendig, den Begriff des *Speichers* in den  $\lambda$ -Kalkül einzuführen: Im  $\lambda$ -Kalkül mit Zustand stehen Variablen nicht mehr für Werte, die für sie eingesetzt werden können, sondern für *Speicherzellen*. Eine Speicherzelle ist ein Ort im Speicher, der einen Wert aufnimmt, der auch wieder verändert werden kann. Dabei wird jede Speicherzelle durch eine *Adresse* identifiziert. Eine Adresse ist eine abstrakte Größe, es kommt also gar nicht darauf an, um was für eine Art Wert es sich handelt – im realen Computer ist eine Adresse in der Regel einfach eine Zahl. In diesem Abschnitt steht  $A$  für die Menge der Adressen, die abzählbar sein sollte.

Ein Speicher aus der Menge  $M$  ist eine Zuordnung zwischen Adressen aus  $A$  und Werten. Die Werte sind wie schon im normalen  $\lambda$ -Kalkül die Basiswerte und die Abstraktionen – hier bekommen sie, weil sie eine Rolle in den Reduktionsregeln spielen, den Namen  $X$ :

$$\begin{aligned}
M &= \mathcal{P}(A \times X) \\
X &= B \cup \{\lambda v.e \mid \lambda v.e \in \mathcal{L}_{\lambda S}\}
\end{aligned}$$

Um Reduktionsregeln für den  $\lambda$ -Kalkül mit Zustand zu formulieren, muß  $\mathcal{L}_{\lambda S}$  noch erweitert werden, damit die Adressen ins Spiel kommen: Adressen werden Terme und sind auf der linken Seite von Zuweisungen zulässig:

$$\begin{array}{l}
\langle \mathcal{L}_{\lambda S} \rangle \rightarrow \ldots \\
| \langle A \rangle \\
| (\text{set! } \langle A \rangle \langle \mathcal{L}_{\lambda S} \rangle)
\end{array}$$

Adressen tauchen dabei nur als Zwischenschritte bei der Reduktion auf; sie sind nicht dafür gedacht, daß sie der Programmierer in ein Programm schreibt.

Da das bisherige Substitutionsprinzip bei Zuweisungen nicht mehr funktioniert, reicht es nicht, die Reduktionsregeln für den  $\lambda$ -Kalkül mit Zustand einfach nur auf Termen auszudrücken: Ein Term, der ja Adressen enthalten kann, ergibt nur Sinn, wenn er mit einem Speicher kombiniert wird. Die Reduktionsregeln überführen somit immer ein Paar, bestehend aus einem Term und einem Speicher in ein ebensolches Paar. Hierbei wird der Einfachheit halber kein Unterschied mehr

zwischen den verschiedenen Arten der Reduktion gemacht:

$$\begin{aligned}
 b, m &\rightarrow a, m[a \mapsto b] \text{ wobei } a \text{ frisch} \\
 \lambda v.e, m &\rightarrow a, m[a \mapsto \lambda v.e] \text{ wobei } a \text{ frisch} \\
 (a_0 a_1), m &\rightarrow e[v \mapsto a], m[a \mapsto m(a_1)] \text{ wobei } m(a_0) = \lambda v.e \text{ und } a \text{ frisch} \\
 (\text{set! } a_0 a_1), m &\rightarrow \text{void}, m[a_0 \mapsto m(a_1)] \\
 (F^k a_1 \dots a_k), m &\rightarrow a, m[a \mapsto F_B(b_1, \dots, b_k)] \text{ wobei } b_i = m(a_i) \in B, a \text{ frisch}
 \end{aligned}$$

Die Formulierung « $a$  frisch» bedeutet dabei, daß  $a$  eine Adresse sein sollte, die in  $m$  bisher noch nicht benutzt wurde. Die Operation  $m[a \mapsto x]$  ist ähnlich wie bei Umgebungen definiert: der alte Speicherinhalt bei  $a$  wird zunächst entfernt, und dann eine neue Zuordnung für  $a$  nach  $x$  hinzugefügt:

$$m[a \mapsto x] \stackrel{\text{def}}{=} (e \setminus \{(a, x') \mid (a, x') \in m\}) \cup \{(a, x)\}$$

Die Regeln sind immer noch über Substitution definiert, allerdings werden für Variablen jetzt nicht mehr Werte sondern Adressen eingesetzt. Sie werden, wie beim normalen Call-by-Value-Kalkül auch, auf Subterme fortgesetzt, die möglichst weit links innen stehen.

Im folgenden Beispiel stehen fettgedruckte Zahlen **0**, **1** für Adressen. Die Redexe sind jeweils unterstrichen:

$$\begin{aligned}
 &((\lambda x.((\lambda y.x)(\text{set! } x (+ x 1)))) 12), \emptyset \\
 &\rightarrow (\underline{\mathbf{0} \ 12}), \{(\mathbf{0}, \lambda x.((\lambda y.x)(\text{set! } x (+ x 1))))\} \\
 &\rightarrow (\underline{\mathbf{0} \ 1}), \{(\mathbf{0}, \lambda x.((\lambda y.x)(\text{set! } x (+ x 1))))\}, (\mathbf{1}, 12)\} \\
 &\rightarrow ((\lambda y.\underline{\mathbf{2}})(\text{set! } \mathbf{2} (+ \mathbf{2} \ 1))), \{(\mathbf{0}, \lambda x.((\lambda y.x)(\text{set! } x (+ x 1))))\}, (\mathbf{1}, 12), (\mathbf{2}, 12)\} \\
 &\rightarrow (\mathbf{3} (\text{set! } \mathbf{2} (+ \mathbf{2} \ 1))), \\
 &\quad \{(\mathbf{0}, \lambda x.((\lambda y.x)(\text{set! } x (+ x 1))))\}, (\mathbf{1}, 12), (\mathbf{2}, 12), (\mathbf{3}, (\lambda y.\mathbf{2}))\} \\
 &\rightarrow (\mathbf{3} (\text{set! } \mathbf{2} (+ \mathbf{2} \ 4))), \\
 &\quad \{(\mathbf{0}, \lambda x.((\lambda y.x)(\text{set! } x (+ x 1))))\}, (\mathbf{1}, 12), (\mathbf{2}, 12), (\mathbf{3}, (\lambda y.\mathbf{2})), (\mathbf{4}, 1)\} \\
 &\rightarrow (\mathbf{3} (\text{set! } \mathbf{2} \ 5)), \\
 &\quad \{(\mathbf{0}, \lambda x.((\lambda y.x)(\text{set! } x (+ x 1))))\}, (\mathbf{1}, 12), (\mathbf{2}, 12), (\mathbf{3}, (\lambda y.\mathbf{2})), (\mathbf{4}, 1), (\mathbf{5}, 13)\} \\
 &\rightarrow (\mathbf{3} \text{ void}), \\
 &\quad \{(\mathbf{0}, \lambda x.((\lambda y.x)(\text{set! } x (+ x 1))))\}, (\mathbf{1}, 12), (\mathbf{2}, 13), (\mathbf{3}, (\lambda y.\mathbf{2})), (\mathbf{4}, 1), (\mathbf{5}, 13)\} \\
 &\rightarrow (\underline{\mathbf{3} \ 6}), \\
 &\quad \{(\mathbf{0}, \lambda x.((\lambda y.x)(\text{set! } x (+ x 1))))\}, (\mathbf{1}, 12), (\mathbf{2}, 13), (\mathbf{3}, (\lambda y.\mathbf{2})), (\mathbf{4}, 1), (\mathbf{5}, 13), (\mathbf{6}, \text{void})\} \\
 &\rightarrow \mathbf{2}, \\
 &\quad \{(\mathbf{0}, \lambda x.((\lambda y.x)(\text{set! } x (+ x 1))))\}, (\mathbf{1}, 12), (\mathbf{2}, 13), (\mathbf{3}, (\lambda y.\mathbf{2})), (\mathbf{4}, 1), (\mathbf{5}, 13), (\mathbf{6}, \text{void})\}
 \end{aligned}$$

Der Endausdruck steht für die Speicherzelle an Adresse **2**, wo der Wert 13 steht. Es ist sichtbar,

daß die Auswertungsmaschinerie durch die Einführung von Zustand deutlich komplizierter wird.

## 17.7 DIE SECDH-MASCHINE

Die SECD-Maschine ist nicht mächtig genug, um den  $\lambda$ -Kalkül mit Zustand zu modellieren: Es fehlt ein Speicher. Darum muß das Maschinen-Pendant zum  $\lambda$ -Kalkül mit Zustand um eine Speicher-Komponente erweitert werden: Heraus kommt die *SECDH-Maschine*, um die es in diesem Abschnitt geht.

Der Maschinencode für die SECDH-Maschine ist dabei genau wie bei der SECD-Maschine, nur daß eine spezielle Zuweisungsoperation hinzukommt:

$\langle I \rangle \rightarrow \dots$   
 $\quad | :=$

Die Übersetzungsfunktion produziert diese neue Instruktion bei **set!**-Ausdrücken:

$$\llbracket e \rrbracket \stackrel{\text{def}}{=} \begin{cases} \dots \\ v \llbracket e' \rrbracket := & \text{falls } e = (\mathbf{set!} \ v \ e') \end{cases}$$

Der Begriff der Adresse aus der Menge  $A$  wird direkt aus dem Kalkül übernommen. Ähnlich wie im Kalkül landen Zwischenergebnisse nicht mehr direkt auf dem Stack, sondern stattdessen landen ihre Adressen im Speicher. Dementsprechend bilden nun Umgebungen Variablen auf Adressen ab. Die neue Komponente  $H$  ist gerade der Speicher, auch genannt *Heap*, der die Adressen auf Werte abbildet:

$$\begin{aligned} S &= A^* \\ E &= \mathcal{P}(V \times A) \\ D &= (S \times E \times C)^* \\ H &= \mathcal{P}(A \times W) \\ W &= B \cup (V \times C \times E) \end{aligned}$$

Die Regeln für die SECDH-Maschine sind analog zu den Regeln für die SECD-Maschine. Zwei Hauptunterschiede gibt es dabei:



- Der Heap aus  $H$  gehört nun zum Zustand dazu. Anders als die Umgebung wird er nicht bei der Bildung von Closures «eingepackt»: Stattdessen wird der Heap stets linear von links nach rechts durch Regeln durchgefädelt.
- Zwischenergebnisse nehmen stets den Umweg über den Heap: Immer, wenn ein neues Zwischenergebnis entsteht, wird es bei einer neuen Adresse im Heap abgelegt. Auf dem Stack landen die Adressen der Zwischenergebnisse.

$$\begin{aligned}
& \hookrightarrow \in \mathcal{P}((S \times E \times C \times D \times H) \times (S \times E \times C \times D \times H)) \\
(\underline{s}, e, b\underline{c}, \underline{d}, h) & \hookrightarrow (a\underline{s}, e, \underline{c}, \underline{d}, h[a \mapsto b]) \\
& \text{wobei } a \text{ frisch} \\
(\underline{s}, e, v\underline{c}, \underline{d}, h) & \hookrightarrow (e(v)\underline{s}, e, \underline{c}, \underline{d}, h) \\
(a_k \dots a_1 \underline{s}, e, \mathbf{prim}_{F^k} \underline{c}, \underline{d}, h) & \hookrightarrow (a\underline{s}, e, \underline{c}, \underline{d}, h[a \mapsto b]) \\
& \text{wobei } a \text{ frisch, } b_i = h(a_i) \text{ und } F^k \in \Sigma^k \text{ und } F_B^k(b_1, \dots, b_k) = b \\
(a_1 a_0 \underline{s}, e, :=\underline{c}, \underline{d}, h) & \hookrightarrow (a\underline{s}, e, \underline{c}, \underline{d}, h[a_0 \mapsto h(a_1)] [a \mapsto \mathbf{void}]) \\
& \text{wobei } a \text{ frisch} \\
(\underline{s}, e, (v, \underline{c}')\underline{c}, \underline{d}, h) & \hookrightarrow (a\underline{s}, e, \underline{c}, \underline{d}, h[a \mapsto (v, \underline{c}', e)]) \\
& \text{wobei } a \text{ frisch} \\
(a_1 a_0 \underline{s}, e, \mathbf{apc} \underline{c}, \underline{d}, h) & \hookrightarrow (\epsilon, e'[v \mapsto a], \underline{c}', (\underline{s}, e, \underline{c})\underline{d}, h[a \mapsto h(a_1)]) \\
& \text{wobei } a \text{ frisch und } h(a_0) = (v, \underline{c}', e') \\
(a, e, \epsilon, (\underline{s}', e', \underline{c}')\underline{d}, h) & \hookrightarrow (a\underline{s}', e', \underline{c}', \underline{d}, h)
\end{aligned}$$

Entsprechend muß die Auswertungsfunktion das Endergebnis im Heap nachschauen:

$$\begin{aligned}
eval_{SECD} & \in \mathcal{L}_{\lambda S} \times Z \\
eval_{SECD}(e) & = \begin{cases} h(a) & \text{falls } (\epsilon, \emptyset, \llbracket e \rrbracket, \epsilon, \emptyset) \hookrightarrow^* (a, e, \epsilon, \epsilon, h), h(a) \in B \\ \mathbf{proc} & \text{falls } (\epsilon, \emptyset, \llbracket e \rrbracket, \epsilon, \emptyset) \hookrightarrow^* (a, e, \epsilon, \epsilon, h), h(a) = (v, \underline{c}, e') \end{cases}
\end{aligned}$$

## 17.8 IMPLEMENTIERUNG DER SECDH-MASCHINE

Für die Implementierung der SECDH-Maschine werden einige der Funktionen wiederverwendet, die für die SECD-Maschine programmiert wurden. Zunächst einmal muß – genau wie bei der SECD-Maschine – erst einmal die Übersetzung von Termen in Maschinencode realisiert werden. Zuweisungsterme haben wie in den Lehrsprachen die folgende Form:

```
(set! v e)
```

Das dazu passende Prädikat ist das folgende:

```
; Prädikat für Zuweisungen
(: assignment? (%a -> boolean))
(define assignment?
  (lambda (t)
    (and (cons? t)
         (equal? 'set! (first t)))))
```

```
(define assignment (signature (predicate assignment?)))
```

Mit Hilfe dieser Definition kann die Signatur-Definition von `term` erweitert werden:

```
(define term
  (signature
    (mixed symbol
            application
            abstraction
            base
            primitive-application
            assignment)))
```

Um zu vermeiden, daß Zuweisungen mit regulären Applikationen verwechselt werden, muß das Prädikat `application?` erweitert werden:

```
(define application?
  (lambda (t)
    (and (cons? t)
         (not (equal? 'set! (first t)))
         (not (equal? 'lambda (first t)))
         (not (primitive? (first t))))))
```

Als Nächstes wird die zusätzliche `:=`-Instruktion repräsentiert. Hier sind Daten- und Record-Definition:

```
; Eine Zuweisungs-Instruktion ist ein Wert
; (make-:=)
(define-record-procedures :=
  make-:= :=?
  ())
(: make-:= (-> :=))
```

Die Signatur-Definition für Maschinen-Instruktionen kann um `:=` erweitert werden:

```
(define instruction
  (signature
    (mixed base
      symbol
      ap
      tailap
      prim
      abs
      :=)))
```

Bei der Übersetzung in Maschinencode kommt in `term->machine-code` ein weiterer Zweig hinzu:

```
; Term in Maschinencode übersetzen
(: term->machine-code (term -> machine-code))
(define term->machine-code
  (lambda (e)
    (cond
      ...
      ((assignment? e)
       (cons (first (rest e))
              (append (term->machine-code (first (rest (rest e))))
                      (list (make-:=)))))))
```

Wie bei der SECD-Maschine werden die verschiedenen Mengendefinitionen erst einmal in Daten- und Record-Definitionen übersetzt. Das ist für Stacks, Umgebungen und Speicheradressen ganz einfach:

```
; Ein Stack ist eine Liste aus Adressen.
(define stackh (signature (list-of address)))
```

```
; Eine Umgebung bildet Variablen auf Adressen ab.
```

```
; Eine Adresse ist eine ganze Zahl.
(define address (signature natural))
```

Die Änderung in der Definition von Umgebungen bedingt eine Änderung der Signatur von `make-binding`:

```
(: make-binding (symbol address -> binding))
```

Bei der Repräsentation des Heaps ist wichtig, daß eine Operation zur Beschaffung frischer Adressen eingebaut wird. Aus diesem Grund enthält der Heap zusätzlich zu den Zellen auch noch einen Zähler mit der nächsten frischen Adresse:

```
; Ein Heap ist ein Wert
;   (make-heap s n)
; wobei n die nächste freie Adresse ist und s eine Liste
; von Zellen.
(define-record-procedures heap
  make-heap heap?
  (heap-cells heap-next))
(: make-heap ((list-of cell) natural -> heap))
```

Der leere Heap wird schon einmal vorfabriziert:

```
(define the-empty-heap (make-heap empty 0))
```

Jede Zelle ordnet einer Adresse einen Wert zu:

```
; Eine Zelle ist ein Wert
;   (make-cell a w)
; wobei a eine Adresse und w ein Wert ist
(define-record-procedures cell
  make-cell cell?
  (cell-address cell-value))
(: make-cell (address value -> cell))
```

Die Funktion `heap-store`, erweitert den Heap um eine Zelle entsprechend der mathematischen Definition:

```
; Wert im Speicher ablegen
(: heap-store (heap address value -> heap))
(define heap-store
  (lambda (h a w)
    (make-heap (cons (make-cell a w)
                     (remove-cell a (heap-cells h)))
              ...)))
```

Die Ellipse steht für die nächste frische Adresse: Wenn die bisherige frische Adresse in `heap-store` belegt wird, so muß eine neue frische Adresse gewählt werden:

```
(define heap-store
  (lambda (h a w)
    (make-heap (cons (make-cell a w)
                     (remove-cell a (heap-cells h)))
              (let ((next (heap-next h)))
                (if (= a next)
                    (+ next 1)
                    next))))))
```

Es fehlt noch die Hilfsfunktion `remove-cell`:

```
; Zelle zu einer Adresse entfernen
(: remove-cell (address (list-of cell) -> (list-of cell)))
(define remove-cell
  (lambda (a c)
    (cond
      ((empty? c) empty)
      ((cons? c)
       (if (= a (cell-address (first c)))
           (rest c)
           (cons (first c)
                 (remove-cell a (rest c))))))))
```

Als Nächstes ist die Operation an der Reihe, die den Wert, der an einer Adresse im Heap gespeichert ist. Die Funktion `heap-lookup` benutzt eine Hilfsfunktion `cells-lookup`, um in der Liste von Zellen nach der richtigen zu suchen:

```
; den Wert an einer Adresse im Heap nachschauen
```

```
(: heap-lookup (heap address -> value))
(define heap-lookup
  (lambda (h a)
    (cells-lookup (heap-cells h) a)))

; den Wert an einer Adresse in einer Liste von Zellen nachschauen
(: cells-lookup ((list-of cell) address -> value))
(define cells-lookup
  (lambda (c a)
    (cond
      ((empty? c) (violation "unassigned address"))
      ((cons? c)
       (if (= a (cell-address (first c)))
           (cell-value (first c))
           (cells-lookup (rest c) a))))))
```

Schließlich fehlt noch eine Repräsentation für den void-Wert:

```
; Ein void-Wert ist ein Wert
; (make-void)
(define-record-procedures void
  make-void void?
  ())
(: make-void (-> void))
```

Auch hier wird nur ein void-Wert benötigt, der vorfabriziert wird:

```
(define the-void (make-void))
```

Der Zustand für die SECDH-Maschine wird genau wie bei der SECD-Maschine repräsentiert, ergänzt um die Komponente für den Heap:

```
; Ein SECDH-Zustand ist ein Wert
; (make-secd s e c d h)
; wobei s ein Stack, e eine Umgebung, c Maschinencode,
; d ein Dump und h ein Speicher ist.
(define-record-procedures secdh
  make-secdh secdh?
  (secdh-stack secdh-environment secdh-code secdh-dump secdh-heap))
```

```
(: make-secdh (stackh environment machine-code dump heap -> secdh))
```

Die Implementierung der Zustandsübergangsfunktion hat exakt die gleiche Struktur wie die Implementierung der SECD-Maschine und hält sich eng an die mathematische Definition der Regeln:

```
; eine Zustandstransition berechnen
(: secdh-step (secdh -> secdh))
(define secdh-step
  (lambda (state)
    (let ((stack (secdh-stack state))
          (environment (secdh-environment state))
          (code (secdh-code state))
          (dump (secdh-dump state))
          (heap (secdh-heap state)))
      (cond
        ((cons? code)
         (cond
           ((base? (first code))
            (let ((a (heap-next heap)))
              (make-secdh
               (cons a stack)
               environment
               (rest code)
               dump
               (heap-store heap a (first code))))))
           ((symbol? (first code))
            (make-secdh
             (cons (lookup-environment environment (first code))
                   stack)
             environment
             (rest code)
             dump
             heap)))
           ((prim? (first code))
            (let ((a (heap-next heap)))
              (make-secdh
               (cons a
```

---

```

                                (drop (prim-arity (first code)) stack))
environment
(rest code)
dump
(heap-store heap a
  (apply-primitive
    (prim-operator (first code))
    (map (lambda (address)
          (heap-lookup heap address))
      (take-reverse (prim-arity (first code)) stack))))))
((:=? (first code))
 (let ((a (heap-next heap)))
  (make-secdh
   (cons a (rest (rest stack)))
   environment
   (rest code)
   dump
   (heap-store
    (heap-store heap
      (first (rest stack))
      (heap-lookup heap (first stack)))
    a the-void))))
((abs? (first code))
 (let ((a (heap-next heap)))
  (make-secdh
   (cons a stack)
   environment
   (rest code)
   dump
   (heap-store heap a
    (make-closure (abs-variable (first code))
                  (abs-code (first code))
                  environment))))))
((ap? (first code))
 (let ((closure (heap-lookup heap (first (rest stack)))))

```



```

        (a (heap-next heap)))
(make-secdh empty
  (extend-environment
    (closure-environment closure)
    (closure-variable closure)
    a)
  (closure-code closure)
  (cons
    (make-frame (rest (rest stack)) environment (rest code)
      dump)
    (heap-store heap a (heap-lookup heap (first stack)))))
((tailap? (first code))
  (let ((closure (heap-lookup heap (first (rest stack)))))
    (a (heap-next heap)))
    (make-secdh (rest (rest stack))
      (extend-environment
        (closure-environment closure)
        (closure-variable closure)
        a)
      (closure-code closure)
      dump
      (heap-store heap a
        (heap-lookup heap (first stack)))))))
((empty? code)
  (let ((f (first dump)))
    (make-secdh
      (cons (first stack)
        (frame-stack f))
      (frame-environment f)
      (frame-code f)
      (rest dump)
      heap))))))

```

Es bleibt die Auswertungsfunktion, die ebenfalls genau wie bei der SECD-Maschine realisiert wird:

```

; aus Term SECDH-Anfangszustand machen
(: inject-secdh (term -> secdh))

```

```

(define inject-secdh
  (lambda (e)
    (make-secdh empty
      the-empty-environment
      (term->machine-code e)
      empty
      the-empty-heap)))

; bis zum Ende Zustandsübergänge berechnen
(: secdh-step* (secdh -> secdh))
(define secdh-step*
  (lambda (state)
    (if (and (empty? (secdh-code state))
      (empty? (secdh-dump state)))
      state
      (secdh-step* (secdh-step state)))))

; Evaluationsfunktion zur SECD-Maschine berechnen
(: eval-secdh (term -> (mixed value (enum 'function))))
(define eval-secdh
  (lambda (e)
    (let ((final (secdh-step* (inject-secdh e))))
      (let ((val (heap-lookup (secdh-heap final)
        (first (secdh-stack final)))))
        (if (base? val)
          val
          'proc))))))

```

## ÜBUNGSAUFGABEN

### AUFGABE 17.1

Übersetzen Sie folgende Lambda-Terme in die Zwischenrepräsentation der SECD-Maschine:

- I.  $(\lambda xy. (+ x y)) (* 5 6) 23$

2.  $(\lambda x.(! x)) (\lambda xy.(\&\& x y)) ((\lambda xy.(> x y)) 23\ 42)\ true$
3.  $(\lambda xy. y\ x\ x) (\lambda z. z) (\lambda yz. (y\ y) (y\ z))$

Dabei steht ! für das boolesche not und && für das boolesche and.

#### AUFGABE 17.2

Betrachten Sie folgendes SECD-Programm:

$$(f, (x, (y, f\ x\ \mathbf{ap}\ y\ \mathbf{ap}))) (a, (b, a\ b\ \mathbf{prim}_+))\ \mathbf{ap}\ 23\ \mathbf{ap}\ 42\ \mathbf{ap}$$

1. Übersetzen Sie das SECD-Programm in den entsprechenden  $\mathcal{L}_{\lambda A}$ -Term.
2. Werten Sie das SECD-Programm aus und geben Sie die einzelnen Auswertungsschritte an!

#### AUFGABE 17.3

Erweitern Sie die Implementierung der SECD-Maschine um korrekte Behandlung der Endrekursion! Erweitern Sie dazu zunächst die Datendefinition für Maschinencode. Implementieren Sie dann die Übersetzung von  $\lambda$ -Termen für die endrekursive SECD-Maschine. Erweitern Sie schließlich die Zustandsübergangsfunktion um einen Fall für die `tailap`-Instruktion.

#### AUFGABE 17.4

Die um Endrekursion erweiterte SECD-Maschine führt eine neue Maschinencode-Instruktion `tailap` ein. Dies ist aber nicht unbedingt nötig. Formulieren Sie die Zustandsübergangsregeln der SECD-Maschine mit Endrekursion so um, daß die Funktionalität, also insbesondere die richtige Behandlung endrekursiver Applikationen, auch ohne das zusätzliche Schlüsselwort `tailap` erhalten bleibt.

#### AUFGABE 17.5

Zeigen Sie in der um Endrekursion erweiterten SECD-Maschine, daß `tailap` immer am Ende steht, also tatsächlich keinen Kontext besitzt.

#### AUFGABE 17.6

Erweitern Sie die SECD-Maschine um Primitive anderer Stelligkeiten, z.B. `abs` oder `odd?`.

#### AUFGABE 17.7

Ändern Sie die Implementierung der SECDH-Maschine dahingehend, daß sie Endrekursion korrekt behandelt.

#### AUFGABE 17.8

Abstrahieren Sie über `remove-environment-binding` und `remove-cell`.

#### AUFGABE 17.9

Erweitern Sie den angewandten  $\lambda$ -Kalkül um Abstraktionen und Applikationen mit mehr als einem Parameter. Erweitern Sie die SECD-Maschine und ihre Implementierung entsprechend.

#### AUFGABE 17.10

Erweitern Sie den angewandten  $\lambda$ -Kalkül um binäre Verzweigungen analog zu `if`. Erweitern Sie entsprechend die SECD-Maschine und ihre Implementierung.

#### AUFGABE 17.11

`Begin` läßt sich im angewandten  $\lambda$ -Kalkül als syntaktischer Zucker auffassen: Wie müssten `begin`-Ausdrücke in die Sprache des Kalküls übersetzt werden?

#### AUFGABE 17.12

Anstatt Umgebungen durch Listen von Bindungen zu repräsentieren, ist es auch möglich, Funktionen zu verwenden, so daß `lookup-environment` folgendermaßen aussieht:

```
(define lookup-environment
  (lambda (e v)
    (e v)))
```

Ergänzen Sie eine passende Definition für `extend-environment`.

## AUFGABE 17.13

Auf den ersten Blick erscheint es etwas aufwendig, jedesmal bei der Auswertung einer Abstraktion die gesamte Umgebung in die Closure einzupacken. Was würde sich ändern, wenn dieser Schritt weggelassen würde, Closures also nur Variable und Maschinencode für den Rumpf enthalten würden? Formulieren Sie die entsprechenden Regeln für die SECD-Maschine und ändern Sie die Implementierung entsprechend. Funktioniert die SECD-Maschine nach der Änderung noch korrekt?



## A MATHEMATISCHE GRUNDLAGEN

Dieser Anhang erläutert die in diesem Buch verwendeten Begriffe und Notationen aus der Mathematik.

### A.1 AUSSAGENLOGIK

Eine *Aussage* ist ein Satz, der prinzipiell einen *Wahrheitswert* W (für «wahr») oder F (für «falsch») besitzt.

Aus *primitiven (elementaren) Aussagen* werden mit Hilfe sogenannter aussagenlogischer *Junktoren* zusammengesetzte Aussagen aufgebaut. Es gibt zwei vordefinierte primitive Aussagen  $\top$  und  $\perp$  mit den Wahrheitswerten W beziehungsweise F. Die wichtigsten Junktoren sind:

«und» ( $\wedge$ ):  $a \wedge b$  hat den Wahrheitswert W genau dann, wenn  $a$  und  $b$  beide den Wert W haben.

«oder» ( $\vee$ ):  $a \vee b$  hat den Wahrheitswert W genau dann, wenn von  $a$  und  $b$  mindestens eins den Wert W hat.

«nicht» ( $\neg$ ):  $\neg a$  hat den Wahrheitswert W genau dann, wenn  $a$  den Wert F hat.

In der Aussagenlogik gilt das Prinzip, daß der Wahrheitswert einer zusammengesetzten Aussage durch die Wahrheitswerte seiner Bestandteile bestimmt ist.

Statt  $\neg a$  wird gelegentlich auch  $\bar{a}$  geschrieben.

Meistens werden logische Junktoren durch sogenannte *Wahrheitstafeln* definiert:

$\wedge$	W	F
W	W	F
F	F	F

$\vee$	W	F
W	W	W
F	W	F

$\neg$	
W	F
F	W

Andere Junktoren, die ebenfalls häufig verwendet werden, sind:

$\Rightarrow$	W	F
W	W	F
F	W	W

«impliziert» ( $\Rightarrow$ ):

$a \Rightarrow b$  spricht sich als «wenn  $a$ , dann  $b$ » oder «aus  $a$  folgt  $b$ ».  $F \Rightarrow W$  besitzt ebenso wie  $F \Rightarrow F$  den Wahrheitswert W! In der formalen Aussagenlogik folgt aus einer falschen Voraussetzung jede Folgerung.

«genau-dann-wenn» ( $\Leftrightarrow$ ):

$\Leftrightarrow$	W	F
W	W	F
F	F	W

Häufig werden Wahrheitstafeln auch in einer etwas ausführlicheren Form notiert, wie im folgenden gezeigt. Dabei sind die Wahrheitstafeln für alle vorgestellten Junktoren in einer Tabelle zusammengefaßt:

$a$	$b$	$a \wedge b$	$a \vee b$	$\neg a$	$a \Rightarrow b$	$a \Leftrightarrow b$
W	W	W	W	F	W	W
W	F	F	W	F	F	F
F	W	F	W	W	W	F
F	F	F	F	W	W	W

Zur Einsparung von Klammern wird vereinbart, daß  $\neg$  am stärksten bindet, gefolgt von  $\wedge$ , dann  $\vee$ , dann  $\Rightarrow$  und zum Schluß  $\Leftrightarrow$ .

Eine zusammengesetzte Aussage heißt *allgemeingültig* oder eine *Tautologie*, wenn sie stets den Wahrheitswert W besitzt, unabhängig vom Wahrheitswert ihrer elementaren Aussagen. Beispiele für Tautologien sind etwa  $a \vee \bar{a}$  («Satz vom ausgeschlossenen Dritten») und  $\bar{a} \wedge \bar{\bar{a}}$  («Satz vom Widerspruch»). Zwei Aussagen  $a$  und  $b$  heißen *äquivalent*, wenn  $a \Leftrightarrow b$  eine Tautologie ist.

Es ist möglich, jede aussagenlogische Aussage durch Wahrheitstafeln auf ihre Allgemeingültigkeit hin zu überprüfen. Auch die Äquivalenz von Ausdrücken läßt sich durch Wahrheitstafeln überprüfen. In der Regel ist es jedoch einfacher, mit diesen Aussagen formal zu rechnen. Die folgenden Tautologien stellen Rechenregeln für die Aussagenlogik dar:

#### LEMMA A.1

Für Aussagen  $a, b, c$  gilt:

$a \wedge a \Leftrightarrow a$	$a \vee a \Leftrightarrow a$	Idempotenzgesetze
$(a \wedge b) \wedge c \Leftrightarrow a \wedge (b \wedge c)$	$(a \vee b) \vee c \Leftrightarrow a \vee (b \vee c)$	Assoziativgesetze
$a \wedge b \Leftrightarrow b \wedge a$	$a \vee b \Leftrightarrow b \vee a$	Kommutativgesetze
$a \wedge (a \vee b) \Leftrightarrow a$	$a \vee (a \wedge b) \Leftrightarrow a$	Absorptivgesetze
$a \wedge (b \vee c) \Leftrightarrow (a \wedge b) \vee (a \wedge c)$	$a \vee (b \wedge c) \Leftrightarrow (a \vee b) \wedge (a \vee c)$	Distributivgesetze
$\overline{a \wedge b} \Leftrightarrow \bar{a} \vee \bar{b}$	$\overline{a \vee b} \Leftrightarrow \bar{a} \wedge \bar{b}$	DeMorgan'sche Gesetze
$\bar{\bar{a}} \Leftrightarrow a$		



## A.2 MENGEN

Die ursprüngliche Definition des Begriffs «Menge» lautet:

«Unter einer Menge verstehen wir eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.» (G. CANTOR)

Die Objekte einer Menge  $M$  heißen *Elemente* von  $M$ . Die Notation  $x \in M$  bedeutet, daß  $x$  ein Element von  $M$  ist,  $x \notin M$ , daß  $x$  kein Element von  $M$  ist.

In der Informatik wie in der Mathematik werden häufig Mengen von *Zahlen* gebraucht. Für die wichtigsten Zahlenmengen gibt es feste Bezeichnungen. So bezeichnet  $\mathbb{N}$  die Menge der *natürlichen Zahlen*; in diesem Buch gilt  $0 \in \mathbb{N}$ .  $\mathbb{Z}$  bezeichnet die Menge der *ganzen Zahlen* und  $\mathbb{R}$  die Menge der *reellen Zahlen*.

Endliche Mengen, also Mengen mit endlich vielen Elementen können als Aufreihung ihrer Elemente aufgeschrieben werden:

$$M = \{11, 13, 17, 19\}.$$

Häufig werden Mengen jedoch auch durch eine bestimmte Eigenschaft definiert, die ihre Elementen haben:

$$M = \{x \mid x \text{ ist Primzahl, } 10 \leq x \leq 20\}.$$

Die *leere Menge* ist die Menge, die keine Elemente besitzt und wird durch  $\emptyset$  bezeichnet.

$A$  heißt *Teilmenge* von  $B$ , in Zeichen  $A \subseteq B$ , wenn jedes Element von  $A$  auch Element von  $B$  ist:

$$A \subseteq B \stackrel{\text{def}}{\iff} a \in A \Rightarrow a \in B.$$

Zwei Mengen sind gleich, wenn sie die gleichen Elemente besitzen; dies läßt sich mit Hilfe der Teilmengenbeziehung auch so ausdrücken:

$$A = B \stackrel{\text{def}}{\iff} A \subseteq B \text{ und } B \subseteq A.$$

Hieraus ergibt sich für die oben erwähnte Darstellung endlicher Mengen z.B.

$$\{11, 13, 17, 19\} = \{17, 13, 19, 11\},$$

das heißt die *Reihenfolge* der Elemente ist unerheblich (beziehungsweise es gibt gar keine ausgezeichnete Reihenfolge) und

$$\{11, 13, 17, 19\} = \{11, 13, 11, 17, 17, 11, 13, 19\},$$

das heißt es spielt keine Rolle, wie oft ein bestimmtes Element erwähnt wird; es ist trotzdem nur einmal in der Menge enthalten.

Die Notation  $A \not\subseteq B$  bedeutet, daß  $A \subseteq B$  nicht gilt,  $A \neq B$ , daß  $A = B$  nicht gilt.  $A$  heißt *echte Teilmenge* von  $B$ , wenn  $A \subseteq B$ , aber  $A \neq B$ . Die Notation dafür ist  $A \subset B$ . Es bedeutet  $B \supseteq A$ , daß  $A \subseteq B$  gilt, ebenso für  $B \supset A$ .

Die *Vereinigung*  $A \cup B$  zweier Mengen  $A$  und  $B$  ist definiert durch

$$A \cup B \stackrel{\text{def}}{=} \{a \mid a \in A \vee a \in B\}.$$

Der *Durchschnitt*  $A \cap B$  zweier Mengen  $A$  und  $B$  ist definiert durch

$$A \cap B \stackrel{\text{def}}{=} \{a \mid a \in A \wedge a \in B\}.$$

Die *Differenz*  $A \setminus B$  zweier Mengen  $A$  und  $B$  ist definiert durch

$$A \setminus B \stackrel{\text{def}}{=} \{a \mid a \in A \wedge a \notin B\}.$$

#### DEFINITION A.2

Das *kartesische Produkt*  $A \times B$  zweier Mengen  $A$  und  $B$  ist definiert durch

$$A \times B \stackrel{\text{def}}{=} \{(a, b) \mid a \in A, b \in B\}.$$

Für  $n \geq 2$  Mengen  $A_1, \dots, A_n$  ist definiert:

$$A_1 \times \dots \times A_n \stackrel{\text{def}}{=} \{(a_1, \dots, a_n) \mid a_i \in A_i\}.$$

Für eine Menge  $A$  und eine natürliche Zahl  $n \geq 2$  ist

$$A^n \stackrel{\text{def}}{=} A \times \overset{n}{\dots} \times A.$$

Um die Fälle  $n = 0$  und  $n = 1$  nicht immer ausschließen zu müssen, wird außerdem definiert:

$$\begin{aligned} A^1 &\stackrel{\text{def}}{=} A \\ A^0 &\stackrel{\text{def}}{=} \{()\}. \end{aligned}$$

$A^0$  ist also eine einelementige Menge, deren einziges Element in Übereinstimmung mit der Tupel-schreibweise  $(a_1, \dots, a_n)$  mit  $()$  bezeichnet wird.

Für eine Menge  $A$  wird die Anzahl ihrer Elemente, ihre *Mächtigkeit*, als  $|A|$  geschrieben. Für unendliche Mengen wird  $|A| = \infty$  definiert.

Für eine Menge  $A$  heißt

$$\mathcal{P}(A) \stackrel{\text{def}}{=} \{T \mid T \subseteq A\}$$

die *Potenzmenge* von  $A$ . Für endliche Mengen gilt  $|\mathcal{P}(A)| = 2^{|A|}$ .

#### DEFINITION A.3

Für  $T \in \mathcal{P}(A)$  ist das *Komplement* von  $T$  in  $A$  definiert durch

$$\overline{T} \stackrel{\text{def}}{=} A \setminus T.$$

#### LEMMA A.4

Für  $A, B \in \mathcal{P}(M)$  gelten die sogenannten *DeMorgan'schen Gesetze*:

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

Es ist üblich, Teilmengen  $T \subseteq \mathcal{P}(A)$  durch sog. *charakteristische Funktionen* darzustellen:

#### DEFINITION A.5

Sei  $A$  eine Menge,  $T \in \mathcal{P}(A)$ . Die *charakteristische Funktion von  $T$*  ist definiert durch

$$\chi_T : A \longrightarrow \{0, 1\}$$

$$\chi_T(x) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{falls } x \in T \\ 0 & \text{falls } x \notin T \end{cases}$$

Ist umgekehrt  $f : A \rightarrow \{0, 1\}$  eine (totale) Abbildung, so läßt sich hieraus eine Menge  $T_f \in \mathcal{P}(A)$  ableiten durch

$$T_f \stackrel{\text{def}}{=} \{x \in A \mid f(x) = 1\}.$$

Die Zuordnung  $T \mapsto \chi_T$  ist bijektiv.

### A.3 PRÄDIKATENLOGIK

Viele Aussagen haben die Form «es gibt (mindestens) ein Objekt (*Individuum*) mit der Eigenschaft...» oder «für alle Objekte aus einem bestimmten Bereich gilt...». Solche Aussagen heißen *Prädikate* und für ihre mathematische Behandlung werden sogenannte *Quantoren* eingeführt, und zwar der *Allquantor* (Universalquantor)  $\forall$  und der *Existenzquantor*  $\exists$ . Im folgenden werden Großbuchstaben zur Bezeichnung von Prädikaten und Kleinbuchstaben für Individuen verwendet. Die *Stelligkeit* eines Prädikats ist die Anzahl der Individuen, über die hier eine Aussage gemacht wird.

Ist  $Q$  ein  $n$ -stelliges Prädikat und sind  $x_1, \dots, x_n$  Individuen eines Individuenbereichs, so ist die Behauptung, daß  $Q$  auf  $x_1, \dots, x_n$  zutrifft (abgekürzt  $Qx_1 \dots x_n$ ) eine prädikatenlogische Aussage. Mathematische Ausdrücke wie  $x \in M$  oder  $n < m$  oder aussagenlogische Aussagen sind ebenfalls prädikatenlogische Aussagen, nur in anderer Schreibweise.

Statt der Individuen selbst kommen in den Quantorenausdrücken *Individuenvariablen* vor; diese sind Platzhalter für die Individuen selbst. Die prädikatenlogische Aussage  $\forall x : Qx$  bedeutet: Für alle Individuen  $a$  gilt die Aussage (das Prädikat)  $Q$ , wobei für die Variable  $x$  das Individuum  $a$  eingesetzt wird. Dementsprechend heißt  $\exists x : Qx$ : Es gibt mindestens ein Individuum  $a$ , so daß die Aussage  $Qa$  wahr ist. Die Variable  $x$  heißt in beiden Fällen eine *gebundene Variable* des Prädikatsausdrucks. Variablen, die nicht gebunden sind, heißen *freie Variablen*.

Eine wichtige Eigenschaft der Quantoren beschreibt der folgende Satz:

**SATZ A.6** Ist  $Q$  ein einstelliges Prädikat, so ist die Aussage  $\neg \forall x : Qx$  genau dann wahr, wenn  $\exists x : \neg Qx$  wahr ist. Umgekehrt ist die Aussage  $\neg(\exists x : Qx)$  genau dann wahr, wenn  $\forall x : \neg Qx$  wahr ist.

Im Prinzip wäre es also möglich, mit nur einem der beiden Quantoren auszukommen.

Häufig werden Einschränkungen an den Individuenbereich gemacht, z.B. in der Form

$$\forall x \in M : \exists y \in N : Pxy .$$

Dies dient als Abkürzung für die kompliziertere Aussage

$$\forall x : (x \in M \Rightarrow \exists y \in N : Pxy)$$

oder die noch kompliziertere Aussage

$$\forall x : (x \in M \Rightarrow \exists y : (y \in N \wedge Pxy)) .$$

## A.4 MULTIMENGEN

Mengen sind so definiert, daß sie jedes Element «nur einmal enthalten». In einer *Multimenge*, kann jedes Element mit einer bestimmten *Multiplizität* (Vielfachheit) vorkommen. Technisch werden Multimengen als Kreuzprodukte  $M \subseteq G \times (\mathbb{N} \setminus \{0\})$  beschrieben, wobei  $G$  die *Grundmenge* heißt und die natürliche Zahl die Multiplizität jedes Elements der Grundmenge angibt.

Die Schreibweise  $x \in M$  ist eine Abkürzung für die Tatsache  $(x, n) \in M$ . Die Multiplizität  $\mathcal{M}(M, x)$  eines Elements beziehungsweise Nicht-Elements  $x$  ist so definiert:

$$\mathcal{M}(M, x) \stackrel{\text{def}}{=} \begin{cases} 0 & x \notin M \\ n & (x, n) \in M \end{cases}$$

Die mengentheoretischen Operationen sind wie folgt definiert:

$$\begin{aligned} P \cup Q &\stackrel{\text{def}}{=} \{(x, m) \mid x \in P \vee x \in Q, m = \max(\mathcal{M}(P, x), \mathcal{M}(Q, x))\} \\ P \cap Q &\stackrel{\text{def}}{=} \{(x, m) \mid (x, p) \in P, (x, q) \in Q, m = \min(p, q), m > 0\} \\ P \setminus Q &\stackrel{\text{def}}{=} \{(x, m) \mid (x, p) \in P, m = p \ominus \mathcal{M}(Q, x), m > 0\} \end{aligned}$$

Dabei ist  $\ominus$  die positive Differenz mit dem Wert 0 für  $q > p$ .

## A.5 RELATIONEN UND ABBILDUNGEN

### DEFINITION A.7

Eine (*binäre*) *Relation* ist eine Teilmenge  $\rho \subseteq A \times B$ . Eine alternative Notation für  $(a, b) \in \rho$  ist  $a\rho b$ . Für eine Relation  $\rho$  heißt

$$\rho^{-1} \stackrel{\text{def}}{=} \{(b, a) \mid (a, b) \in \rho\}$$

die *Umkehrrelation* von  $\rho$ .

### DEFINITION A.8

Eine Relation  $\rho \subseteq A \times A$  heißt

- *reflexiv*  $\stackrel{\text{def}}{\iff}$  für alle  $a \in A$  gilt  $a\rho a$ ,
- *symmetrisch*  $\stackrel{\text{def}}{\iff}$  aus  $a\rho b$  folgt  $b\rho a$ ,
- *antisymmetrisch*  $\stackrel{\text{def}}{\iff}$  aus  $a\rho b$  und  $b\rho a$  folgt  $a = b$ ,
- *transitiv*  $\stackrel{\text{def}}{\iff}$  aus  $a\rho b$  und  $b\rho c$  folgt  $a\rho c$ ,
- *Äquivalenzrelation*  $\stackrel{\text{def}}{\iff}$   $\rho$  ist reflexiv, symmetrisch und transitiv.

Äquivalenzrelationen werden oft dazu verwendet, Elemente einer Menge in *Äquivalenzklassen* einzuteilen. Die Äquivalenzklasse  $[a]$  eines Elements  $a \in A$  bezüglich einer Äquivalenzrelation  $\cong$  ist die Menge aller Elemente  $b \in A$  mit der Eigenschaft  $b \cong a$ .

Für eine Relation  $\rho \subseteq A \times A$  wird häufig eine verwandte Relation  $\rho' \supseteq \rho$  betrachtet, die eine oder mehrere der oberen Eigenschaften zusätzlich zu  $\rho$  besitzt. Eine solche Relation heißt jeweils *Abschluß*:

### DEFINITION A.9 (ABSCHLÜSSE ÜBER MENGEN)

Gegeben sei eine Relation  $\rho \subseteq A \times A$ . Eine Relation  $\rho' \subseteq A \times A$  heißt

- *reflexiver Abschluß* von  $\rho \stackrel{\text{def}}{\iff} \rho'$  ist die kleinste reflexive Relation mit  $\rho \subseteq \rho'$ ,
- *symmetrischer Abschluß* von  $\rho \stackrel{\text{def}}{\iff} \rho'$  ist die kleinste symmetrische Relation mit  $\rho \subseteq \rho'$ ,
- *transitive Abschluß* von  $\rho \stackrel{\text{def}}{\iff} \rho'$  ist die kleinste transitive Relation mit  $\rho \subseteq \rho'$ ,
- *reflexiv-transitiver Abschluß* von  $\rho \stackrel{\text{def}}{\iff} \rho'$  ist die kleinste reflexive und transitive Relation mit  $\rho \subseteq \rho'$ .

Dabei heißt «kleinste Relation» jeweils, daß für jede andere Relation  $\rho''$ , welche die jeweilige Eigenschaft erfüllt, gilt  $\rho' \subseteq \rho''$ .

Die transitive Abschluß von  $\rho$  wird  $\rho^+$  geschrieben, der reflexiv-transitive Abschluß  $\rho^*$ .

#### DEFINITION A.10

Eine *Abbildung* ist ein Tripel  $f = (A, \rho_f, B)$ , wobei  $A$  und  $B$  Mengen sind und  $\rho_f \subseteq A \times B$  eine Relation, so daß für jedes  $a \in A$  genau ein  $b \in B$  existiert mit  $a\rho_fb$ .  $A$  heißt *Vorbereich* von  $f$ ,  $B$  heißt *Nachbereich* und  $\rho_f$  der *Graph* von  $f$ . Für  $f = (A, \rho_f, B)$  steht auch  $f : A \rightarrow B$  oder  $A \xrightarrow{f} B$ . Für  $(a, b) \in \rho_f$  steht normalerweise  $f(a) = b$ .

Nach dieser Definition sind zwei Abbildungen  $f = (A, \rho_f, B)$  und  $g = (C, \rho_g, D)$  *gleich* genau dann, wenn  $A = C$ ,  $B = D$  und  $\rho_f = \rho_g$ .

#### DEFINITION A.11

Für eine Menge  $A$  ist die *Identitätsabbildung* durch  $\text{id}_A \stackrel{\text{def}}{=} (A, \rho_{\text{id}_A}, A)$  mit  $\rho_{\text{id}_A} \stackrel{\text{def}}{=} \{(a, a) \mid a \in A\}$  definiert.

#### DEFINITION A.12

Eine Abbildung (auch: Funktion)  $A \xrightarrow{f} B$  heißt

- *surjektiv*  $\stackrel{\text{def}}{\iff}$  für alle  $b \in B$  gibt es ein  $a \in A$ , so daß  $f(a) = b$ ,
- *injektiv*  $\stackrel{\text{def}}{\iff}$  aus  $f(a_1) = f(a_2)$  folgt  $a_1 = a_2$  für beliebige  $a_1, a_2 \in A$ ,
- *bijektiv*  $\stackrel{\text{def}}{\iff} A \xrightarrow{f} B$  ist injektiv und surjektiv.

Ist  $A \xrightarrow{f} B$  bijektiv, so heißen  $A$  und  $B$  *isomorph*, in Zeichen  $A \cong B$ .

#### DEFINITION A.13

Für zwei Abbildungen  $A \xrightarrow{f} B$  und  $B \xrightarrow{g} C$  definiere durch

$$g \circ f \stackrel{\text{def}}{=} (A, \rho_{g \circ f}, C)$$

die *Komposition* von  $g$  und  $f$  mit

$$\rho_{g \circ f} \stackrel{\text{def}}{=} \{(a, c) \mid \exists b \in B : (a, b) \in \rho_f \wedge (b, c) \in \rho_g\}.$$

#### LEMMA A.14

Die Komposition von Abbildungen ist assoziativ, das heißt für  $A \xrightarrow{f} B$ ,  $B \xrightarrow{g} C$  und  $C \xrightarrow{h} D$  gilt

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Es ist deshalb möglich, die Klammern ganz wegzulassen und  $h \circ g \circ f$  zu schreiben.

LEMMA A.15

Für eine bijektive Abbildung  $A \xrightarrow{f} B$  mit  $f = (A, \rho_f, B)$  existiert eine *Umkehrabbildung*  $B \xrightarrow{f^{-1}} A$  mit  $f^{-1} \stackrel{\text{def}}{=} (B, \rho_{f^{-1}}, A)$ , wobei  $\rho_{f^{-1}} \stackrel{\text{def}}{=} \rho_f^{-1}$ . Es gilt  $f^{-1} \circ f = \text{id}_A$  und  $f \circ f^{-1} = \text{id}_B$ .

DEFINITION A.16

Sei  $A$  eine Menge,  $f : A \rightarrow A$  eine Abbildung. Dann ist definiert:

$$\begin{aligned} f^0 &\stackrel{\text{def}}{=} \text{id}_A \\ f^{n+1} &\stackrel{\text{def}}{=} f \circ f^n \end{aligned}$$

## A.6 ORDNUNGEN

DEFINITION A.17

Sei  $M$  eine Menge. Eine Relation  $\rho \subseteq M \times M$  ist eine *partielle Ordnung* (auch *Halbordnung* genannt) auf  $M \iff \rho$  ist reflexiv, transitiv und antisymmetrisch (vgl. Definition A.8).

Ein Beispiel für eine partielle Ordnung ist etwa die Relation  $\ll \subseteq \gg$  zwischen Mengen. Eine *partiell geordnete Menge*  $(M; \rho)$  ist eine Menge  $M$  mit einer partiellen Ordnung  $\rho$ . In einer partiell geordneten Menge kann es *unvergleichbare Elemente* geben, das heißt Elemente  $x, y \in M$ , für die weder  $x\rho y$  noch  $y\rho x$  gilt. In der halbgeordneten Menge  $(M; \subseteq)$  mit

$$M = \{\{1, 2\}, \{2, 3\}, \{1, 2, 3\}\}$$

sind z.B. die Elemente  $\{1, 2\}$  und  $\{2, 3\}$  unvergleichbar.



## DEFINITION A.18

Eine *totale Ordnung* auf einer Menge  $M$  ist eine partielle Ordnung  $\rho$ , bei der zusätzlich gilt

$$\forall x, y \in M : x\rho y \vee y\rho x .$$

## DEFINITION A.19

Sei  $(M; \rho)$  eine halbgeordnete Menge. Ein Element  $x \in M$  heißt

*minimales Element*:  $\stackrel{\text{def}}{\iff} \forall y \in M : y\rho x \Rightarrow y = x$

*kleinstes Element*:  $\stackrel{\text{def}}{\iff} \forall y \in M : x\rho y$

*maximales Element*:  $\stackrel{\text{def}}{\iff} \forall y \in M : x\rho y \Rightarrow y = x$

*größtes Element*:  $\stackrel{\text{def}}{\iff} \forall y \in M : y\rho x$

Die Begriffe «minimales» und «kleinstes Element» werden gern verwechselt. Das liegt vielleicht daran, daß die Relation « $\leq$ » auf Zahlen, die oft als «Modell» für eine partielle Ordnung erhalten muß, in Wirklichkeit eine totale Ordnung ist. Bei totalen Ordnungen fallen die Begriffe «minimales» und «kleinstes Element» jedoch zusammen. Deshalb hier noch einmal eine verbale Definition:

*minimales Element* heißt, daß es kein Element gibt, das kleiner ist. Es kann aber Elemente geben, die unvergleichbar mit einem minimalen Element sind. (In der oben angegebenen Menge  $M$  sind  $\{1, 2\}$  und  $\{2, 3\}$  beide minimal.)

*kleinstes Element* heißt, daß alle anderen Elemente größer sind. Damit ist auch die Vergleichbarkeit gegeben. Es gibt in einer Menge höchstens ein kleinstes Element. (In der oben angegebenen Menge  $M$  gibt es kein kleinstes Element.)

## DEFINITION A.20

Eine totale Ordnung heißt *wohlfundierte Ordnung*, wenn es keine unendlich langen absteigenden Ketten gibt, d. h. keine unendliche Folge  $(a_i \mid i \in \mathbb{N})$  mit der Eigenschaft

$$\forall i \in \mathbb{N} : a_i \rho^{-1} a_{i+1} .$$

In einer wohlfundierten Menge besitzt jede nichtleere Teilmenge ein kleinstes Element.

## AUFGABEN

## AUFGABE A.1

$A \xrightarrow{f} B \xrightarrow{g} C$  sei injektiv (surjektiv). Welche Aussagen über  $A \xrightarrow{f} B$  beziehungsweise  $B \xrightarrow{g} C$  gelten dann mit Bestimmtheit?

## AUFGABE A.2

Zeige, daß die Umkehrfunktion  $B \xrightarrow{f^{-1}} A$  einer bijektiven Funktion  $A \xrightarrow{f} B$  stets injektiv ist.

## AUFGABE A.3

Für eine endliche Menge  $A$  beweise

$$|\mathcal{P}(A)| = 2^{|A|}$$

## AUFGABE A.4

Sei  $A$  eine endliche Menge. Beweise, daß für  $T \in \mathcal{P}(A)$  gilt

$$|T| + |\overline{T}| = |A|$$

## AUFGABE A.5

Beweise anhand der Wahrheitstafeln für die Aussagenlogik:

- Die DeMorgan'schen Gesetze:

$$\neg(A \wedge B) = \neg A \vee \neg B$$

$$\neg(A \vee B) = \neg A \wedge \neg B$$

- Das erste Distributivgesetz:

$$A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$$

- Die Kontraposition:

$$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$$

- Folgenden Satz über die Implikation:

$$(A \Rightarrow (B \Rightarrow C)) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$$

## AUFGABE A.6

In einem an Schulen populären Lehrbuch der Mathematik wird folgende Aussage gemacht:

Für eine partielle Ordnung  $\sqsubseteq$  auf einer Menge  $D$  und  $d_1, d_2 \in D$  mit  $d_1 \not\sqsubseteq d_2$  gilt  $d_2 \sqsubseteq d_1$ .

Gib eine partielle Ordnung an, für welche die Aussage nicht gilt!



## LITERATURVERZEICHNIS

- ABELSON, HAROLD, G. J. SUSSMAN, and J. SUSSMAN (1996). *Structure and Interpretation of Computer Programs*. MIT Press, Cambridge, Mass., second ed.
- ADAMS, STEPHEN (1993). *Efficient sets—a balancing act*. *Journal of Functional Programming*, 3(4):553–561.
- BARENDREGT, HENK P. (1990). *Functional programming and lambda calculus*. In VAN LEEUWEN, JAN, ed.: *Handbook of Theoretical Computer Science—Formal Models and Semantics*, vol. B, chap. 7. Elsevier Science Publishers.
- BAUER, FRIEDRICH L. and H. WÖSSNER (1984). *Algorithmische Sprache und Programmentwicklung*. Springer-Verlag.
- BIRD, RICHARD and J. GIBBONS (2020). *Algorithm Design with Haskell*. Cambridge University Press.
- BROOKS, FREDERICK P. (1995). *The Mythical Man-Month, Anniversary Edition: Essays on Software Engineering*. Addison-Wesley, 2nd ed.
- CLAESSEN, KOEN and J. HUGHES (2000). *QuickCheck: A lightweight tool for random testing of Haskell programs*. In *Proceedings of the International Conference on Functional Programming (ICFP)*, pp. 268–279.
- FELLEISEN, MATTHIAS, R. B. FINDLER, M. FLATT, and S. KRISHNAMURTHI (2014). *How to Design Programs*. MIT Press, 2nd ed.
- FRIEDMAN, DANIEL P., M. WAND, and C. T. HAYNES (2001). *Essentials of Programming Languages*. MIT Press and McGraw-Hill, 2nd ed.
- GOLDBERG, DAVID (1991). *What every computer scientist should know about floating-point arithmetic*. *ACM Computing Surveys*, 23(1):5–48.
- GOOS, GERHARD (1996). *Vorlesungen über Informatik*, vol. 1: Grundlagen und funktionales Programmieren. Springer-Verlag.
- HAILPERIN, MAX, B. KAISER, and K. KNIGHT (1999). *Concrete Abstractions*. Brooks/Cole.
- HINDLEY, J. R. and J. P. SELDIN (1986). *Introduction to Combinators and  $\lambda$ -Calculus*, vol. 1 of *Mathematical Sciences Student Texts*. Cambridge University Press, London.

HINZE, RALF (1991). *Einführung in die funktionale Programmierung mit Miranda*. Teubner Verlag.

HOFSTADTER, DOUGLAS R. (1979). *Gödel, Escher, Bach: An Eternal Golden Braid*. Basic Books, New York.

HUTH, MICHAEL and M. RYAN (2004). *Logic in Computer Science*. Cambridge University Press, 2nd ed.

KELSEY, RICHARD, W. CLINGER, and J. REES (1998). *Revised<sup>5</sup> report on the algorithmic language Scheme*. Higher-Order and Symbolic Computation, 11(1):7–105.

KLAEREN, HERBERT (1983). *Algebraische Spezifikation — Eine Einführung*. Lehrbuch Informatik. Springer-Verlag, Berlin-Heidelberg-New York.

LANDIN, PETER (1964). *The mechanical evaluation of expressions*. The Computer Journal, 6:308–320.

MEYER, BERTRAND (1997). *Object-Oriented Software Construction*. Prentice-Hall, second ed.

PEYTON JONES, SIMON and J.-M. EBER (2003). *How to write a financial contract*. In GIBBONS, JEREMY and O. DE MOOR, eds.: *The Fun of Programming*. Palgrave Macmillan.

RAYMOND, ERIC S., ed. (1996). *The New Hacker's Dictionary*. MIT Press, 3rd ed.

THIEMANN, PETER (1994). *Grundlagen der funktionalen Programmierung*. Teubner Verlag, Stuttgart.

## INDEX

- $()$ , 561
- $M^*$ , 221
- $\Leftrightarrow$ , 558
- $\Rightarrow$ , 557
- $\perp$ , 557
- $\cap$ , 560
- $\chi_T$ , 561
- $\circ$ , 565
- $\cup$ , 560
- $\epsilon$ , 221
- $\exists$ , 562
- $\forall$ , 562
- $\in$ , 559, 563
- $\lceil \rceil$ , 499
- $\mapsto$ , 492
- $\|$ , 561
- $\neg$ , 557
- $\notin$ , 559
- $+$
- $\rho$ , 565
- $*$
- $\rho$ , 565
- $\rightarrow_\alpha$ , 494
- $\rightarrow_\beta$ , 494
- $\backslash$ , 560, 561
- $\subset$ , 560
- $\subseteq$ , 559
- $\supset$ , 560
- $\supseteq$ , 560
- $\times$ , 560
- $T$ , 557
- $\emptyset$ , 559
- $V$ , 557
- $\wedge$ , 557
- $\leq$ , 44
- $<$ , 44
- $=$ , 44
- $\Rightarrow$ , 434
- $>$ , 44
- $>=$ , 44
- DrRacket, 3
- Abbildung, 565
- Abdeckung, 29, 48
- abgeleitete Form, 72
- abgeschlossen, 491
- abs, 341
- Abschluß, 564
- Absolutbetrag, 71, 341
- Absorptivgesetz, 558
- Abstraktion, 18, 22, 34, 489
- Abstraktionsregel, 514
- Adresse, 539, 540
- Akkumulator, 287
- Algebra, 60, 425
- Algorithmus, 378
- allgemeingültig, 558
- Allquantor, 562
- $\alpha$ -Reduktion, 494
- Alternative, 71
- and, 44
- angewandter  $\lambda$ -Kalkül, 509, 539
- animate, 330
- Ansicht, 321
- antisymmetrisch, 434, 564
- Antwort, 516
- append, 185
- applicative-order reduction, 504

- 
- Applikation, 22, 489
  - Applikationsregel, 514
  - Argument, 22
  - `articles-weight`, 465
  - Assoziativgesetz, 558
  - Assoziativität, 423, 427
  - Aufzählung, 49, 67
  - Ausdruck, 8
  - Aussage, 557
    - elementar, 557
    - primitiv, 557
  - Aussagenlogik, 557
  - Auswertungsfunktion, 515
  - Auswertungsstrategie, 503
  - balancieren
    - von Suchbäumen, 378
  - Basis, 199
  - Basiselement, 220
  - Basiswert, 509, 539
  - Baum, 353
    - größenannotiert, 379
  - Bedingung, 44
  - $\beta$ -Reduktion, 494
  - `big-bang`, 345
  - bijektiv, 565
  - Bindung, 268, 527
    - dynamisch, 270
    - lexikalisch, 269
    - statisch, 270
  - Binärbaum, 353, 358
  - binäre Relation, 433, 564
  - binäre Verzweigung, 70
  - binäres Prädikat, 433
  - Blatt, 359
  - `boolean`, 45
  - `boolean?`, 121
  - boolesche Fallunterscheidung, 70
  - boolescher Ausdruck, 43
  - boolescher Wert, 43, 498
  - `bound`, 491
  - Bruch, 27
  - Call-by-Value-Reduktion, 505, 511
  - Call-by-Value-Sprache, 506
  - `cat`, 224
  - charakteristische Funktion, 561
  - `check-error`, 75
  - `check-expect`, 28
  - `check-property`, 400
  - $\chi_T$ , 561
  - Church-Numeral, 499
  - Church/Rosser-Eigenschaft, 496
  - `circle`, 312
  - Closure, 513
  - Code, 513
  - `color`, 312
  - Computer, 80
  - `concatenate`, 183
  - `textttconcatenate`, 435
  - `cond`, 61
  - `cons`, 167
  - Cons-Liste, 166
  - `cons?`, 167
  - `curry`, 443
  - curryfizieren, 265
  - Daten, 23
  - Datenanalyse, 65
  - Datendefinition, 49, 65
  - `define-record`, 81, 88, 114, 175
  - Definition, 11



- lokal, 187, 188
- Definitionsfenster, 7
- $\delta$ -Reduktion, 510
- DeMorgan'sches Gesetz, 432, 558, 561
- Differenz, 560
- Distributivgesetz, 558
- Domänenlogik, 321
- drawn-games, 245
- DrRacket, 7
- Dump, 513
- Durchschnitt, 560
- dynamische Bindung, 270
  
- echte Teilmenge, 560
- Editor, 7
- Eigenschaft, 399, 562
- Eingabe, 9
- eingebaute Prädikate, 121
- Element, 559
- elementare Aussage, 557
- ellipse, 312
- Ellipse, 33
- else, 60
- else-Zweig, 61
- empty-image, 425
- endliche Folge, 221
- endliche Menge, 559
- Endrekursion, 302
- entarteter Suchbaum, 377
- enum, 49
- Environment, 513
- even?, 212
- exakte Zahlen, 428
- exchange-for-cute, 70
- Existenzquantor, 562
- exklusives Oder, 45
  
- expect, 440
- expect-within, 442
- Exponent, 199
- extensionale Gleichheit, 444
  
- F, 557
- fac, 500
- FAC, 501
- factorial, 448
- Fakultät, 211, 297
- Fallunterscheidung, 43, 46, 66, 110
- falsch, 43, 557
- false, 43
- false, 498
- false, 366
- falsifizierbar, 399, 427
- feed-parrot, 115
- Feld, 80
- Fibonacci-Folge, 301
- Fibonacci-Function, 235
- filter, 251
- first, 167
- Fixpunkt, 501
- Fixpunktkombinator, 281, 502
- Fixpunktsatz, 501
- fold, 262
- Folge, 221
- for-all, 398, 399
- Form
  - abgeleitet, 72
- Frame, 529
- free, 491
- freie Variable, 491, 562
- Funktion, 22
  - höherer Ordnung, 237, 242
- Funktionsanwendung, 22

- Funktionsaufruf, 22
- ganze Zahlen, 27, 559
- Gaußsche Summenformel, 213, 233
- gebundene Variable, 491, 562
- gebundenes Vorkommen, 268
- Gegenbeispiel, 427
- gemischte Daten, 109, 110
- Gerüst, 25, 32
- Gleichheit
- extensional, 444
  - intensional, 444
- globale Variable, 268
- Grammatik, 225
- Graph, 565
- Grundmenge, 563
- größenannotierter Baum, 379
- größtes Element, 567
- Halbgruppe, 425
- Halbordnung, 566
- Heap, 542
- Higher-Order-Funktion, 237, 242
- home-points, 239
- Höhe
- eines Baums, 361
- idempotent, 281
- Idempotenzgesetz, 558
- Identität, 490
- Identitätsabbildung, 565
- if*, 498
- if, 71
- image-height, 315
- image-width, 315
- image.rkt, 311
- Implikation, 434, 557
- Index, 203
- Individuenvariable, 562
- Individuum, 562
- Induktion, 217
- noethersch, 235
  - strukturell, 229
  - vollständig, 217
  - wohlfundiert, 235
- Induktionsanfang, 217, 223
- Induktionsaxiom, 232
- Induktionsschluss, 217, 223
- Induktionsverankerung, 220
- induktive Definition, 220
- induktive Menge, 220
- induktiver Abschluss, 220, 232
- inexakte Zahl, 403
- inexakte Zahlen, 428
- Information, 23
- injektiv, 565
- inklusives Oder, 45
- Insertionsort, 406
- Instruktion, 511
- integer, 27
- integer-from-to, 47
- intensionale Gleichheit, 444
- Interaktionsfenster, 8
- Invariante, 291, 452
- Inverse, 405
- inverses Element, 432
- invert, 283
- isomorph, 565
- Isomorphismus, 267
- Iteration, 303
- Junktor, 557

- 
- Kalkül, 489
  - kartesisches Produkt, 560
  - Kette, 567
  - Klammer, 9
  - Klasse, 110
  - kleinstes Element, 567
  - Knoten, 358
  - Kombinator, 38, 131, 134, 140, 424, 491
  - Kommentar, 26
  - Kommutativgesetz, 558
  - Kommutativität, 426
  - Komplement, 561
  - komplexe Zahlen, 27
  - Komponente, 80
  - Komposition, 565
  - Konsequente, 71
  - Konstruktionsanleitung, 30
  - Konstruktor, 81
  - Kontext, 302, 513
  - kontextfreie Grammatik, 225
  - Kopfnormalform
    - schwach, 504
  - Kurzbeschreibung, 25, 31
  
  - $\mathcal{L}_{\lambda A}$ , 509
  - $\mathcal{L}_{\lambda S}$ , 539
  - $\mathcal{L}_{\lambda}$ , 489
  - lambda, 19
  - $\lambda$ -Kalkül, 489
  - $\lambda$ -Term, 489
  - lazy evaluation, 506
  - leere Folge, 221
  - leere Liste, 165
  - leere Menge, 559
  - leftmost-innermost reduction, 504
  - leftmost-outermost reduction, 503
  
  - Lehrsprachen, 7
  - length, 178, 185
  - lexikalische Bindung, 269, 492, 494, 513
  - line, 313
  - lineares Wachstum, 289
  - Linksaußen-Reduktion, 503
  - Linksinnen-Reduktion, 505
  - list, 176
  - list-length, 177
  - list-sum, 441
  - Liste, 165
    - leer, 165
  - Literal, 13, 517
  - Literalregel, 514
  - load-list, 462
  - Logarithmus, 376
  - logischer Kalkül, 489
  - lokale Variable, 187, 188, 268
  
  - Mantra, 36
  - map, 257
  - Markierung, 358
  - Maschinencode, 511
  - maximales Element, 567
  - Menge, 559
    - endlich, 559
    - minimales Element, 567
  - minutes-since-midnight, 89
  - minutes-since-midnight->wallclock-time, 90
  - mixed, 111
  - mode, 312
  - Modell, 321
  - Modularität, 483
  - Monoid, 426
  - Multimenge, 563

- 
- Multiplizität, 563
  - Mächtigkeit, 561
  - $\mathbb{N}$ , 559
  - Nachbereich, 565
  - Nachfolger, 232, 499
  - natural, 27
  - natural?, 121
  - natürliche Zahl, 26, 199, 499, 559
  - natürliche Zahlen, 27
  - Negation, 45
  - neutrales Element, 172, 425, 432
  - nicht-strikte Sprache, 506
  - Nichtwert, 504, 510
  - noethersche Induktion, 235
  - noethersche Ordnung, 235
  - normal-order reduction, 503
  - Normalform, 495
  - not, 45
  - number, 27
  - textttnumber-list=?, 436
  - number?, 121
  - odd?, 212
  - on-key, 345
  - Operand, 9, 489
  - Operator, 9, 489
  - or, 44
  - Ordnung
    - Halb-, 566
    - noethersch, 235
    - partiell, 566
    - total, 567
    - wohlfundiert, 235, 567
  - overlay, 316
  - overlay/xy, 317
  - $\mathcal{P}(A)$ , 561
  - Papagei, 109
  - Parameter, 22, 489
  - parametrische Polymorphie, 177
  - partielle Ordnung, 566
  - Peano-Axiome, 232
  - Periode, 10
  - pet, 49
  - Pfad, 360
  - place-image, 317
  - place-image/align, 318
  - polygon, 313
  - Polymorphie, 177
  - positive natürliche Zahlen, 200
  - positive?, 200
  - posn, 313
  - Potenz, 199
  - Potenzmenge, 561
  - pred, 499
  - primitive Applikation, 510
  - primitive Aussage, 557
  - Primitivregel, 514
  - Primitivum, 509, 539
  - property-based testing, 400
  - Prädikat, 112, 562
    - binär, 433
  - Prädikate
    - eingebaut, 121
  - Prädikatenlogik, 562
  - quadratisches Wachstum, 286
  - Quantor, 562
  - quote, 517
  - $\mathbb{R}$ , 559
  - Racket, 4, 7

- rational, 27
- rational?, 121
- rationale Zahl, 26
- rationale Zahlen, 27
- real, 27
- real?, 121
- Record-Definition, 81, 88
- Record-Signatur, 81
- Record-Typ, 80
- Records, 80
- rectangle, 311
- Red-Black-Trees, 394
- Redundanz, 411
- reelle Zahlen, 27, 559
- reflexiv, 433, 564
- Reihenfolge, 560
- Rekursion, 500
  - strukturell, 227
- rekursiver Aufruf, 137
- Relation, 564
  - antisymmetrisch, 564
  - binär, 433
  - reflexiv, 564
  - symmetrisch, 564
  - transitiv, 564
  - Äquivalenz-, 564
- relationales Problem, 409
- REPL, 8
- Repräsentation, 23, 517
- repräsentierbarer Wert, 520
- rest, 167
- reverse, 289
- Rosenbaum, 396
- Rotation
  - eines Baumes, 379
  - eines Suchbaums, 388
- Rumpf, 22, 25, 33, 489
- run-over-dillo, 97
- Rückkehrregel, 515
- say-number, 46
- Schablone, 68
- Scheme, 4
- Schleife, 303
- schrittweise Verfeinerung, 469
- schwache Kopfnormalform, 504
- schönfinkeln, 263, 265, 490
- sec:SECDH-Maschine, 542
- SECD-Maschine, 511
- Selbstbezug, 131, 133, 220
- selbstquotierend, 517
- Selektor, 82
- Semantikklammern, 228
- Sichtbarkeit, 268
- Signatur, 26, 27
- Signatur-Definition, 49
- Signatur-Deklaration, 27, 31
- Signatur-Konstruktor, 173
- Signatur-Parameter, 173
- signature, 49
- Signaturvariable, 174
- Signaturverletzung, 30
- smart constructor, 161
- Sorte, 26
- Speicher, 540
- Speicherzelle, 540
- Sprachebene, 176
  - Anfänger, 7
  - fortgeschritten, 516
- sqrt, 162
- Stack, 513

- 
- statische Bindung, 270
  - Stelligkeit, 562
  - Stepper, 19
  - strikte Sprache, 506
  - string, 27
  - String, 14
  - string->number, 14
  - string-append, 14
  - string-length, 14
  - string=?, 44, 50
  - string?, 121
  - strukturelle Induktion, 229
  - strukturelle Rekursion, 227
  - Substitution, 492
  - succ*, 499
  - Suchbaum, 371
    - entartet, 377
  - Suchbaumeigenschaft, 454
  - Summe, 8
  - surjektiv, 565
  - symbol?, 518
  - Symbol, 518
  - symmetrisch, 433, 564
  - syntaktischer Zucker, 72
  - tail call*, 302
  - Tautologie, 558
  - Teachpack, 15, 311
  - Teilbaum, 359
  - Teilmenge, 559
    - echt, 560
  - Test, 28
  - Testfall, 28
  - Tests, 32
  - text, 313
  - Text, 13
  - Ticks, 330
  - Tiefe
    - eines Baums, 361
  - Tier, 110
  - title, 461
  - to-draw, 345
  - total-memory, 83
  - totale Ordnung, 567
  - transitiv, 433, 564
  - triangle, 313
  - true, 43
  - true*, 498
  - true, 366
  - Umgebung, 513
  - Umkehrabbildung, 566
  - Umkehrrelation, 564
  - uncurry, 443
  - Unit-Test, 398
  - Universalquantor, 562
  - unvergleichbare Elemente, 566
  - var, 491
  - Variable, 12
    - frei, 491, 562
    - gebunden, 491, 562
    - global, 268
    - Individuen-, 562
    - lokal, 268
  - Variablenregel, 514
  - Vereinigung, 560
  - Verfeinerung
    - schrittweise, 469
  - Verzweigung, 43, 46, 61, 498
  - View, 321
  - Vorbereich, 565

- 
- Vorgänger, 499  
Vorkommen, 268, 492
- W, 557  
wahr, 43, 557  
Wahrheitstafel, 557  
Wahrheitswert, 43, 557  
wallclock-time, 86  
Wert, 504, 510, 540  
Widerspruch, 558  
wohlfundierte Induktion, 235  
wohlfundierte Ordnung, 567  
Wurzel, 360
- $Y$ , 502
- $\mathbb{Z}$ , 559  
Zeichenkette, 13, 14  
zero?, 200  
zerop, 499  
Zucker, syntaktischer, 72  
zusammengesetzten Daten, 79  
Zustand, 539  
Zweierlogarithmus, 376  
Zweig, 47, 61  
Zwischenergebnis, 286, 289
- Äquivalenz  
    im  $\lambda$ -Kalkül, 495  
Äquivalenzklasse, 564  
Äquivalenzrelation, 433, 564