Learning to program with F#

Jon Sporring

Department of Computer Science,
University of Copenhagen

August 26, 2017

Contents

1 | Preface

This book has been written as an introduction to programming for novice programmers. It is used on the first programming course at the University of Copenhagen's bachelor in computer science program. It has been typeset in LaTeX, and all programs have been developed and tested in Mono version 5.2.0.

Jon Sporring
Associate Professor, Ph.d.
Department of Computer Science,
University of Copenhagen
August 26, 2017

2 | Introduction

Programming is a creative process in which exciting problems may be solved and new tools and applications may be created. With programming skills, you can create high-level applications to run on a mobile device that interacts with other users, databases, and artificial intelligences; you may create programs that run on super computers for simulating weather systems on alien planets or social phenomena in the internet economy; and you may create programs that run on small custom-made hardware for controlling your home appliances.

2.1  ow to learn to program

In order to learn how to program there are a couple of steps that are useful to follow:

1. Choose a programming language: It is possible to program without a concrete language, but your ideas and thoughts must be expressed in some fairly rigorous way. Actually, theoretical computer science typically does not rely on computers nor programming languages, but uses mathematics to prove properties of algorithms. However, most computer scientists program, and with a real language, you have the added benefit of checking your algorithm and hence your thoughts rigorously on a real computer. This book teaches a subset of F#. The purpose is not to be a reference guide to this language, but to use it as a vessel to teach you, the reader, how to convert your ideas into programs.

2. Learn the language: A computer language is a structure for thought, and it influences which thoughts you choose to implement as a program, and how you choose to do it. Any conversion requires you to acquirer a sufficient level of fluency, for you to be able to make programs. You do not need to be a master in F# nor to know every corner of the language, and you will expand your knowledge as you expose yourself to solving problems in the language, but you must invest an initial amount of time and energy in order to learn the basics of the language. This book aims at getting you started quickly, which is why we intentionally are teaching a small subset of F#. On the net and through other works, you will be able to learn much more.

3. Practice: If you want to be a good programmer, then there is only one way: practice, practice, practice! It has been estimated that to master anything, then you have to have spent at least 10000 hours of practice, so get started logging hours! It of course matters, what you practice. This book teaches 3 different programming themes. The point is that programming is thinking, and the scaffold that you use, shapes your thoughts. It is therefore important to recognize this scaffold, and to have the ability to choose that which suits your ideas and your goals best. And the best way to expand your abilities is to sharpen your present abilities, push yourself into new territory, and trying something new. Do not be afraid to make errors or be frustrated at first. These are the experiences that make you grow.

4. Solve real problems: I have found that using my programming skills in real situations with customers demanding solutions, that work for them, has allowed me to put into perspective the

programming tools and techniques that I use. Often customers want solutions that work, are secure, are cheap, and delivered fast, which has pulled me as a programmer in the direction of "if it works, then sell it", while on the longer perspective customers also want bug fixes, upgrades, and new features, which require carefully designed code, well written test-suites, and good documentation. And as always, the right solution is somewhere in between. Regardless, real problems create real programmers.

2.2 How to solve problems

Programming is the act of solving a problem by writing a program to be executed on a computer. A general method for solving problems was given by George Pólya [9] and adapted to programming is:

Understand the problem: To solve any problem it is crucial that the problem formulation is understood: What is to be solved? Do you understand everything in the description of the problem? Is all information for finding the solution available or is something missing?

Design a plan: Good designs mean that programs are faster to program easier to debug and maintain. So, before you start typing a program consider things like: What are the requirements and constraints for the program? Which components should the program have? How are these components supposed to work together? Designing often involves drawing a diagram of the program, and writing pseudo-code on paper.

Implement the plan: Implementation is the act of transforming a program design into code. A crucial part of any implementation is choosing which programming language to use. Also, the solution to many problems will have a number of implementations which vary in how much code they require, to which degree they rely on external libraries, which programming style they are best suited for, what machine resources they require, and what their running times are. With a good design, then the coding is usually easy, since the design will have uncovered the major issues and found solutions for these, but sometimes implementation reveals new problems, which requires rethinking the design. Most implementations also include writing documentation of the code.

Reflect on the result: A crucial part in any programming task is ensuring that the program solves the problem sufficiently. E.g., what are the program's bugs, is the documentation of the code sufficient and relevant for its intended use. Is the code easily maintainable and extendable by other programmers? Are there any general lessons to be learned from or general code developed by the programming experience, which may be used for future programming sessions?

Programming is a very complicated process, and Pólya's list is a useful guide, but not a fail-safe approach. Always approach problem solving with an open mind.

2.3 Approaches to programming

This book focuses on 3 fundamentally different approaches to programming:

Imperative programming, which is a type of programming where *statements* are used to change the program's *state*. Imperative programming emphasizes *how a program shall accomplish a solution* and less on *what the solution is*. A cooking recipe is an example of the spirit of imperative programming. Almost all computer hardware is designed to execute low-level programs written in imperative style. Imperative programming builds on the Turing machine [10]. The first major language was FORTRAN [6] which emphasized an imperative style of programming.

· Imperative programming
· statements
· state

Declarative programming, which emphasizes *what a program shall accomplish* but not *how*. We will consider Functional programming as an example of declarative programming. A *functional programming* language evaluates *functions* and avoids state changes. The program consists of *expressions* instead of statements. As a consequence, the output of functions only depends on its arguments. Functional programming has its roots in lambda calculus [1], and the first language emphasizing functional programming was Lisp [7].

- Functional programming
- functional programming
- functions
- expressions

Structured programming, which emphasizes organization of code in units with well-defined interfaces and isolation of internal states and code from other parts of the program. We will focus on Object-oriented programming as the example of structured programming. *Object-oriented programming* is a type of programming, where the states and programs are structured into *objects*. A typical object-oriented design takes a problem formulation and identifies key nouns as potential objects and verbs as potential actions to be taken on objects. The first object-oriented programming language was Simula 67 developed by Dahl and Nygaard at the Norwegian Computing Center in Oslo [2].

- Structured programming
- Object-oriented programming
- objects

Most programs do not follow a single programming paradigm as, e.g., one of the above, but are a mix. Nevertheless, this book will treat each paradigm separately to emphasize their advantages and disadvantages.

2.4 Why use F#

This book uses F# also known as Fsharp, which is a functional first programming language that also supports imperative and object-oriented programming. It was originally developed for Microsoft's .Net platform, but is available as open source for many operating systems through Mono. As an introduction to programming, F# is a young programming language still under development, with syntax that at times is a bit complex, but it offers a number of advantages:

Interactive and compile mode F# has an interactive and a compile mode of operation: In interactive mode you can write code that is executed immediately in a manner similarly to working with a calculator, while in compile mode, you combine many lines of code possibly in many files into a single application, which is easier to distribute to non F# experts and is faster to execute.

Indentation for scope F# uses indentation to indicate scope: Some lines of code belong together, e.g., should be executed in a certain order and may share data, and indentation helps in specifying this relationship.

Strongly typed F# is strongly typed, reducing the number of runtime errors: F# is picky, and will not allow the programmer to mix up types such as integers and strings. This is a great advantage for large programs.

Multi-platform F# is available on Linux, Mac OS X, Android, iOS, Windows, GPUs, and browsers via the Mono platform.

Free to use and open source F# is supported by the Fsharp foundation (`http://fsharp.org`) and sponsored by Microsoft.

Assemblies F# is designed to be able to easily communicate with other .Net and Mono programs through the language-independent, platform-independent bytecode called Common Intermediate Language (CIL) organized as assemblies. Thus, if you find that certain parts of a program are easy to express in F# and others in C++, then you will be able to combine these parts later into a single program.

Modern computing F# supports all aspects of modern computing including Graphical User Interfaces, Web programming, Information rich programming, Parallel algorithms, ...

Integrated development environments (IDE) F# is supported by major IDEs such as Visual Studio (`https://www.visualstudio.com`) and Xamarin Studio (`https://www.xamarin.com`).

2.5 How to read this book

Learning to program requires mastering a programming language, however most programming languages contains details that are rarely used or used in contexts far from a specific programming topic. Hence, this book only includes a subset of F#, but focuses on language structures necessary to understand 4 common programming paradigms: Imperative programming mainly covered in Chapters 6 to 11, functional programming mainly covered in Chapters 13 to 16, object oriented programming in Chapters 20 and 21, and event driven programming in Chapter 22. A number of general topics given in the appendix for reference. The disadvantage of this approach is that no single part contains a reference guide to F# and F# topics are revisited and expanded across the book. For further reading please consult `http://fsharp.org`.

3 | Executing F# code

3.1 Source code

F# is a functional first programming language, meaning that it has strong support for functional programming, but F# also supports imperative and object-oriented programming. It also has strong support for parallel programming and information rich programs. It was originally developed for Microsoft's .Net platform, but is available as open source for many operating systems through Mono. In this text, we consider F# 4.1 and its Mono implementation, which is different from .Net mainly in terms of the number of libraries accessible. The complete language specification is described in `http://fsharp.org/specs/language-spec/`.

F# has 2 modes of execution, *interactive* and *compiled*. Interactive mode is well suited for small experiments or back-of-an-envelope calculations, but not for programming in general. Both modes can be accessed via the *console*, see Appendix A for more information on the console. The interactive system is started by calling `fsharpi` at the command prompt in the console, while compilation is performed with `fsharpc`, and execution of the compiled code is performed using the `mono` command.

· interactive mode
· compile mode
· console

F# programs comes in many forms, which are identified by suffixes. The *source code* is an F# program written in human readable form using an editor. F# recognises the following types of source code files:

· source code

`.fs` An *implementation file*, e.g., `myModule.fs` · implementation file

`.fsi` A *signature file*, e.g., `myModule.fsi` · signature file

`.fsx` A *script file*, e.g., `gettingStartedStump.fsx` · script file

`.fsscript` Same as `.fsx`, e.g., `gettingStartedStump.fsscript`

Compiled code is source code translated into a machine readable language, which can be executed by a machine. Compiled F# code is either:

`.dll` A *library file*, e.g., `myModule.dll` · library file

`.exe` A stand-alone *executable file*, e.g., `gettingStartedStump.exe` · executable file

The implementation, signature, and script files are all typically compiled to produce an executable file, in which case they are called *scripts*, but can also be entered into the interactive system, in which case these are called *script-fragments*. The implementation and signature files are special kinds of script files used for building libraries. Libraries in F# are called modules, and they are collections of smaller programs used by other programs, which will be discussed in detail in Chapter 9.

· scripts
· script-fragments

3.2 Executing programs

Programs may either be executed by the interpreter or by compiling and executing the compiled code. In Mono the interpreter is called `fsharpi` and can be used in two ways: interactively, where a user enters one or more script-fragments separated by the ";;" characters, or to execute a script file treated as a single script-fragment. [1]

To illustrate the difference between interactive and compile mode, consider the program in Listing 3.1.

Listing 3.1 gettingStartedStump.fsx: A simple demonstration script.

```
1 let a = 3.0
2 printfn "%g" a
```

The code declares a value `a` to be the decimal value 3.0 and finally prints it to the console. An interactive session is obtained by starting the console, typing the `fsharpi` command, typing the lines of the program, and ending the script-fragment with ";;". The dialogue in Listing 3.2 demonstrates the workflow. What the user types has been highlighted by a box.

Listing 3.2: An interactive session.

```
1  $ fsharpi
2
3  F# Interactive for F# 4.1 (Open Source Edition)
4  Freely distributed under the Apache 2.0 Open Source License
5
6  For help type #help;;
7
8  > let a = 3.0
9  - printfn "%g" a;;
10 3
11
12 val a : float = 3.0
13 val it : unit = ()
14
15 > #quit;;
```

We see that after typing `fsharpi`, then the program starts by stating details about itself followed by `>` indicating that it is ready to receive commands. The user then types `let a = 3.0` and presses `enter`, to which the interpreter responds with `-`. This indicates that the line has been received, that the script-fragment is not yet completed, and that it is ready to receive more input. When the user types `printfn "\%g" a;;` followed by `enter`, then by ";;" the interpreter knows that the script-fragment is completed, it interprets the script-fragment, responds with `3` and extra type information about the entered code, and with `>` to indicate, that it is ready for more script-fragments. The interpreter is stopped, when the user types `#quit;;`. It is also possible to stop the interpreter by typing `ctrl-d`.

Instead of running `fsharpi` interactively, we can write the script-fragment from Listing 3.1 into a file, here called `gettingStartedStump.fsx`. This file can be interpreted directly by `fsharpi` as shown in Listing 3.3.

[1] Jon: **Too early to introduce lexeme: "F# uses many characters which at times are given special meanings, e.g., the characters ";;" is compound character denoting end of a script-fragment. Such possibly compound characters are called lexemes."**

Listing 3.3: Using the interpreter to execute a script.

```
$ fsharpi gettingStartedStump.fsx
3
```

Notice that in the file, "`;;`" is optional. We see that the interpreter executes the code and prints the result on screen without the extra type information.

Finally, the file containing Listing 3.1 may be compiled into an executable file with the program `fsharpc`, and run using the program `mono` from the console. This is demonstrated in Listing 3.4.

Listing 3.4: Compiling and executing a script.

```
$ fsharpc gettingStartedStump.fsx
F# Compiler for F# 4.1 (Open Source Edition)
Freely distributed under the Apache 2.0 Open Source License
$ mono gettingStartedStump.exe
3
```

The compiler takes `gettingStartedStump.fsx` and produces `gettingStarted.exe`, which can be run using `mono`.

Both the interpreter and the compiler translates the source code into a format, which can be executed by the computer. While the compiler performs this translation once and stores the result in the executable file, the interpreter translates the code every time the code is executed. Thus, to run the program again with the interpreter, then it must be retranslated as "`$fsharpi gettingStartedStump.fsx`". In contrast, compiled code does not need to be recompiled to be run again, only re-executed using "`$ mono gettingStartedStump.exe`". On a MacBook Pro, with a 2.9 Ghz Intel Core i5, the time the various stages take for this script are:

Command	Time
`fsharpi gettingStartedStump.fsx`	1.88s
`fsharpc gettingStartedStump.fsx`	1.90s
`mono gettingStartedStump.exe`	0.05s

I.e., executing the script with `fsharpi` is slightly faster than by first compiling it with `fsharpc` and then executing the result with `mono`, $1.88\text{s} < 0.05\text{s} + 1.90\text{s}$, if the script were to be executed only once, but every future execution of the script using the compiled version requires only the use of `mono`, which is much faster than `fsharpi`, $1.88\text{s} \gg 0.05\text{s}$.

The interactive session results in extra output on the *type inference* performed, which is very useful for *debugging* and development of code-fragments, but both executing programs with the interpreted directly from a file and compiling and executing the program is much preferred for programming complete programs, since the starting state is well defined, and since this better supports *unit-testing*, which is a method for debugging programs. Thus, **prefer compiling over interpretation**.

· type inference
· debugging
· unit-testing
Advice

4 | Quick-start guide

Programming is the art of solving problems by writing a program to be executed by a computer. For example, to solve the following problem,

Problem 4.1

What is the sum of 357 and 864?

we have written the program in F# shown in Listing 4.1.

Listing 4.1 quickStartSum.fsx:
A script to add 2 numbers and print the result to the console.

```
1 let a = 357
2 let b = 864
3 let c = a + b
4 do printfn "%A" c
```

```
$ fsharpc --nologo quickStartSum.fsx && mono quickStartSum.exe
1221
```

In box the above, we see our program was saved as a script in a file called `quickStartSum.fsx`, and in the console we executed the program by typing the command `fsharpi quickStartSum.fsx`. The result is then printed in the console to be `1221`. Here, as in the rest of this book, we have used the optional flag `--nologo`, which informs `fsharpi` not to print information about its version etc., thus making the output shorter.

To solve the problem, we made program consisting of several lines, where each line was a *expressions*. The first expression `let a = 357` in line 1 used the *"`let`" keyword* to *bind* the value 357 to the name `a`. This is called a *let-binding*. Likewise, we bound the value 864 to the name `b` in line 2, but to the name `c`, we bound the result of evaluating the sum `a + b` in line 3. That is, first the value `a + b` was calculated by substituting the names of `a` and `b` with their values to give the expression `357 + 864`, then this expression was evaluated by adding the values to give `1221`, and this value was finally bound to the name `c`. Line 4 is a *do-binding*, as noted by the keyword *"`do`"*. Do-bindings are also sometimes called *statements*, and the "`do`" keyword is optional in F#. Here the value of `c` was printed to the console followed by a newline (LF possibly preceded by CR, see Appendix C.1) with the ***`printfn`*** *function*. A function in F# is an entity that takes zero or more arguments and returns a value. The function `printfn` is very special, since it can take any number of arguments and returns the value "***`unit`***", and the "`do`" tells F# to ignore this value. Here `printfn` has been used with 2 arguments: `"%A"` and `c`. Notice that in contrast to many other languages, F# does not use parentheses to frame the list of arguments, nor does it use commas to separate them. In general, the `printfn` function always has 1 or more arguments, and the first is a *format string*. A *string* is a sequence of characters starting

- expressions
- "`let`"
- keyword
- binding
- let-binding
- do-binding
- "`do`"
- statements
- `printfn`
- function
- "`unit`"
- format string
- string

and ending with double quotation marks. E.g., `let s = "this is a string of characters"` binds the string `"this is..."` to the name `s`. For the `printfn` function, the format string may be any string, but if it contains format character sequences, such as `%A`, then the values following the format string are substituted. The format string must match the value *type*, that is, here `c` is of type integer, whereas the format string `%A` matches many types.

· type

Types are a central concept in F#. In the script 4.1 we bound values of integer type to names. There are several different integer types in F#, here we used the one called `int`. The values were not *declared* to have these types, instead the types were *inferred* by F#. Typing these bindings line by line in an interactive session, then we see the inferred types as shown in Listing 4.2.

· type declaration
· type inference

Listing 4.2: Inferred types are given as part of the response from the interpreter.

```
> let a = 357;;
val a : int = 357

> let b = 864;;
val b : int = 864

> let c = a + b;;
val c : int = 1221

> do printfn "%A" c;;
1221
val it : unit = ()
```

The interactive session displays the type using the *"val"* keyword followed by the name used in the binding, its type, and its value. Since the value is also responded, then the last `printfn` statement is superfluous. However, **it is ill advised to design programs to be run in an interactive session, since the scripts needs to be manually copied every time it is to be run, and since the starting state may be unclear.** Notice that `printfn` is automatically bound to the name *it* of type "`unit`" and value *"()"*. F# insists on binding all statements to values, and in lack of an explicit name, then it will use `it`. Rumor has it that `it` is an abbreviation for "irrelevant". Also, the value "`()`" is the only possible value of type "`unit`".

· "val"
Advice
· it
· "()"

Were we to solve a slightly different problem,

Problem 4.2

What is the sum of 357.6 and 863.4?

then we would have to use floating point arithmetic instead of integers, and the program would look like Listing 4.3.

Listing 4.3 quickStartSumFloat.fsx: Floating point types and arithmetic.

```
let a = 357.6
let b = 863.4
let c = a + b
do printfn "%A" c
```

```
$ fsharpc --nologo quickStartSumFloat.fsx && mono quickStartSumFloat.exe
1221.0
```

On the surface, this could appear as an almost negligible change, but the set of integers and the set of real numbers (floats) require quite different representations, in order to be effective on a computer, and as a consequence, the implementation of their operations such as addition are very different. Thus, although the response is an integer, it has type `float`, which is indicated by `1221.0`, and which is not the same as `1221`. F# is very picky about types, and generally does not allow types to be mixed, as demonstrated in the interactive session in Listing 4.4.

Listing 4.4: Mixing types is often not allowed.

```
> let a = 357;;
val a : int = 357

> let b = 863.4;;
val b : float = 863.4

> let c = a + b;;
  let c = a + b;;
  ------------^

stdin(4,13): error FS0001: The type 'float' does not match the type 'int'
```

we see that binding a name to a number without a decimal point is inferred to be integer, while when binding to a number with a decimal point, then the type is inferred to be a float, and when trying to add values of integer and floating point, then we get an error. The *error message* contains much information. First it states that the error is in `stdin(4,13)`, which means that the error was found on standard-input at line 4 and column 13. Since the program was executed using `fsharpi quickStartSumFloat.fsx`, then here standard input means the file `quickStartSumFloat.fsx` shown in Listing 4.3. The corresponding line and column is also shown in Listing 4.4. After the file, line, and column number, then F# informs us that the error number, and a verbal description of the error. Error numbers are an under developed feature in Mono, and should be ignored. However, the verbal description often contains useful information for *debugging*. E.g., here we are informed that there is a type mismatch in the expression, i.e., F# since `a` is an integer, then it had expected `b` to be one too. Debugging is the process of solving errors in programs, and here we can solve the error by either making `a` into a float or `b` into an int. The right solution depends on the application.

· error message

· debugging

F# is a functional first programming language, and one implication is that names have a *lexical scope*. A scope is an area in a program, where a binding is valid, and lexical scope means that when a binding is used, then its value is substituted at the place of binding regardless of whether its value is rebound later in the text. Further, at the outer most level, rebinding is not allowed. If attempted, then F# will return an error as shown in Listing 4.5.

· lexical scope

Listing 4.5 quickStartRebindError.fsx:
A name cannot be rebound.

```
let a = 357
let a = 864
```

```
$ fsharpc --nologo -a quickStartRebindError.fsx

quickStartRebindError.fsx(2,5): error FS0037: Duplicate definition of
   value 'a'
```

However, if the same was performed in an interactive session, then rebinding did not cause an error as shown in Listing 4.6.

Listing 4.6: Names may be reused when separated by the lexeme ";;".

```
> let a = 357;;
val a : int = 357

> let a = 864;;
val a : int = 864
```

The difference is that the *";;" lexeme* is used to specifies the end of a *script-fragment*. A lexeme is a letter or a word, which the F# considers as an atomic unit. Script-fragments may be defined both in scripts and in interactive mode, and rebinding is not allowed at the outermost level in script-fragments. Even with the ";;" lexeme, rebinding is not allowed in compile-mode. In general, **avoid rebinding of names.**

· ";;"
· lexeme
· script-fragment
Advice

In F# *functions* are also values, and we may define a function `sum` as part of the solution to the above program as shown in Listing 4.7.

· function

Listing 4.7 quickStartSumFct.fsx:
A script to add 2 numbers using a user defined function.

```
let sum x y = x + y
let c = sum 357 864
do printfn "%A" c
```

```
$ fsharpc --nologo quickStartSumFct.fsx && mono quickStartSumFct.exe
1221
```

Functions are useful to *encapsulate* code, such that we can focus on the transformation of data by a function while ignore the details on how this is done. Functions are also useful for code reuse, i.e., instead of repeating a piece of code in several places, then such code can be encapsulated in a function and replaced with function calls. This makes debugging and maintenance considerably simpler. Entering the function into an interactive session will illustrate the inferred type, the function `sum` has: `val sum : x:int -> y:int -> int`. The "->" is the mapping operator in the sense that functions are mappings between sets. The type of the function `sum`, should be read as `val sum : x:int -> (y:int -> int)`, that is, `sum` takes an integer and returns a function, which takes an integer and returns an integer. This is an example of a higher-order function.

· encapsulate

Type inference in F# may cause problems, since the type of a function is inferred in the context, in which it is defined. E.g., in an interactive session, defining the sum in one scope on a single line will default the types to integers, F#'s favorite type. Thus, if the next script-fragment uses the function with floats, then we will get an error message as shown in Listing 4.8.

Listing 4.8: Types are inferred in blocks, and F# tends to prefer integers.

```
val sum : x:int -> y:int -> int

> let c = sum 357.6 863.4;;
  let c = sum 357.6 863.4;;
  ------------^^^^^

stdin(3,13): error FS0001: This expression was expected to have type
    'int'
but here has type
    'float'
```

A remedy is to define the function in the same script-fragment as it is used such as shown in Listing 4.9.

Listing 4.9: Type inference is per script-fragment.

```
> let sum x y = x + y
- let c = sum 357.6 863.4;;
val sum : x:float -> y:float -> float
val c : float = 1221.0
```

Alternatively, the types may be explicitly stated as shown in Listing 4.10.

Listing 4.10: Function argument and return types may be stated explicitly.

```
> let sum (x : float) (y : float) : float = x + y;;
val sum : x:float -> y:float -> float

> let c = sum 357.6 863.4;;
val c : float = 1221.0
```

The function `sum` has two arguments and a return type, and in Listing 4.10 we have specified all three. This is done using the "`:`" lexeme, and to resolve confusion, we must use parentheses around the arguments such as `(y : float)`, otherwise F# would not be able to understand, whether the type annotation was for the argument or the return value. Often it is sufficient to specify some of the types, since type inference will enforce the remaining types. E.g., in this example, the "+" operator is defined for identical types, so specifying the return value of `sum` to be a float, implies that the result of the "+" operator is a float, and therefore its arguments must be floats, and finally then the arguments for `sum` must be floats. However, in this book we advocate the following advice: **specify types unless explicitly working with generic functions.** Advice

In this chapter, we have scratched the surface of learning how to program by concentrating on a number of key programming concepts and how they are expressed in the F# language. In the following chapters, we will expand the description of F# with features used in all programming approaches.

5 | Using F# as a calculator

In this chapter, we will exclusively use the interactive mode to illustrate basic types and operations in F#.

5.1 Literals and basic types

All programs rely on processing of data, and an essential property of data is its *type*. A *literal* is a fixed value like the number 3, and if we type the number `3` in an interactive session at the input prompt, then F# responds as shown in Listing 5.1.

· type
· literal

Listing 5.1: Typing the number 3.

```
> 3;;
val it : int = 3
```

What this means is that F# has inferred the type to be ***int*** and bound it to the identifier ***it***. Rumor has it, that the identifier `it` is an abbreviation for 'irrelevant'. For more on binding and identifiers see Chapter 6. Types matter, since the operations that can be performed on integers are quite different from those that can be performed on, e.g., strings. E.g., the number 3 has many different representations as shown in Listing 5.2.

· `int`
· `it`

Listing 5.2: Many representations of the number 3 but using different types.

```
> 3;;
val it : int = 3

> 3.0;;
val it : float = 3.0

> '3';;
val it : char = '3'

> "3";;
val it : string = "3"
```

Each literal represents the number 3, but their types are different, and hence they are quite different values. The types `int` for integer numbers, ***`float`*** for floating point numbers, ***`bool`*** for Boolean values, ***`char`*** for characters, and ***`string`*** for strings of characters are the most common types of literals. A table of all *basic types* predefined in F# is given in Table 5.1. Besides these built-in types, F# is

· `float`
· `bool`
· `char`
· `string`
· basic types

Metatype	Type name	Description
Boolean	"`bool`"	Boolean values true or false
Integer	"`int`"	Integer values from -2,147,483,648 to 2,147,483,647
	"`byte`"	Integer values from 0 to 255
	"`sbyte`"	Integer values from -128 to 127
	"`int8`"	Synonymous with byte
	"`uint8`"	Synonymous with sbyte
	"`int16`"	Integer values from -32768 to 32767
	"`uint16`"	Integer values from 0 to 65535
	"`int32`"	Synonymous with int
	"`uint32`"	Integer values from 0 to 4,294,967,295
	"`int64`"	Integer values from -9,223,372,036,854,775,808 to 9,223,372,036,854,775,807
	"`uint64`"	Integer values from 0 to 18,446,744,073,709,551,615
Real	"`float`"	64-bit IEEE 754 floating point value from $-\infty$ to ∞
	"`double`"	Synonymous with float
	"`single`"	A 32-bit floating point type
	"`float32`"	Synonymous with single
	"`decimal`"	A floating point data type that has at least 28 significant digits
Character	"`char`"	Unicode character
	"`string`"	Unicode sequence of characters
None	"`unit`"	The value ()
Object	"`obj`"	An object
Exception	"`exn`"	An exception

Table 5.1: List of some of the basic types. The most commonly used types are underlined. For at description of integer see Appendix B.1, for floating point numbers see Appendix B.2, for ASCII and Unicode characters see Appendix C, for objects see Chapter 20, and for exceptions see Chapter 18.

designed such that it is easy to define new types.

Humans like to use the *decimal number* system for representing numbers. Decimal numbers are *base* 10, which means that a value is represented as two sequences of decimal digits separated by a *decimal point*, where each *digit* can have values $d \in \{0, 1, 2, \ldots, 9\}$, and the value, which each digit represents is proportional to its position. The part before the decimal point is called the *whole part* and the part after is called the *fractional part* of the number. The whole part with neither a decimal point nor a fractional part is called an *integer*. As an example `35.7` is a decimal number, whose value is $3 \cdot 10^1 + 5 \cdot 10^0 + 7 \cdot 10^{-1}$, and `128` is an integer, whose value is $1 \cdot 10^2 + 2 \cdot 10^1 + 8 \cdot 10^0$. In F# a decimal number is called a *floating point number*. Floating point numbers may alternatively be given using *scientific notation*, such as `3.5e-4` and `4e2`, where the `e`-notation is translated to a value as `3.5e-4` $= 3.5 \cdot 10^{-4} = 0.00035$, and `4e2` $= 4 \cdot 10^2 = 400$.

· decimal number
· base
· decimal point
· digit
· whole part
· fractional part
· integer
· floating point number
· scientific notation

The basic unit of information in almost all computers is the binary digit or *bit* for short. Internally, programs and data is all represented as bits, hence F# has a strong support for binary numbers. A *binary number* consists of a sequence of binary digits separated by a decimal point, where each digit can have values $b \in \{0, 1\}$, and the base is 2. E.g., the binary number $101.01_2 = 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 + 0 \cdot 2^{-1} + 1 \cdot 2^{-2} = 5.25$. Subscripts are often used to indicate the base of a number, e.g., 101.01_2 and 101.01_{10} are different numbers. Since base 10 is so common, the subscript for base 10 numbers is often omitted.

· bit
· binary number

Binary numbers are closely related to *octal* and *hexadecimal numbers*, where octals uses 8 as basis, and where each octal digit can be represented by exactly 3 bits, while hexadecimal numbers uses 16 as basis, and where each hexadecimal digit can be written in binary using exactly 4 bits. The hexadecimal digits uses `0`–`9` to represent the values 0–9 and `a`–`f` in lower or alternatively upper case to represent the values 10-15. Thus, Octals and hexadecimals conveniently serve as shorthand for the much longer

· octal number
· hexadecimal number

Character	Escape sequence	Description
BS	`\b`	Backspace
LF	`\n`	Line feed
CR	`\r`	Carriage return
HT	`\t`	Horizontal tabulation
\	`\\`	Backslash
"	`\"`	Quotation mark
’	`\'`	Apostrophe
BEL	`\a`	Bell
FF	`\f`	Form feed
VT	`\v`	Vertical tabulation
	`\uXXXX`, `\UXXXXXXXX`, `\DDD`	Unicode character

Table 5.2: Escape characters. For the unicode characters ’X’ are hexadecimal digits, while for tricode characters ’D’ is a decimal character.

binary representation. As examples, the octal number 37_8 is $3 \cdot 8^1 + 7 \cdot 8^0 = 31$, and the hexadecimal number $f3_{16}$ is $15 \cdot 16^1 + 3 \cdot 16^0 = 243$.

To denote integers on bases different than 10, F# uses the prefix ’0b’ for binary, ’0o’ for octal, and ’0x’ for hexadecimal numbers. For example the value 367_10 may be written as an integer `367`, as a binary number `0b101101111`, as a octal number `0o557`, and as a hexadecimal number `0x16f`. The character sequences `0b12` and `ff` are not numbers.

A *character* is a *Unicode code point*, and character literals are enclosed in single quotation marks. Appendix C.3 contains more details on code points. The character type in F# is denoted *“`char`”*. Examples of characters are ’a’, ’D’, ’3’. However ’23’ and ’abc’ are not characters. Some characters do not have a visual representation such as the tabulation character. These can still be represented as a character using *escape sequences*. A character escape sequence starts with “\” followed by either a letter for simple escapes such as `\t` for tabulation and `\n` for newline. Escape sequences can also be a numerical representation of a code point, and three versions exist: The trigraph `\DDD`, where `D` is a decimal digit, is used to specify the first 256 code points, the hexadecimal escape codes `\uXXXX`, where `X` is a hexadecimal digit, is used to specify the first 65536 code points, and `\UXXXXXXXX` is used to specify any of the approximately $4.3 \cdot 10^9$ possible code points. All escape sequences are shown in Table 5.2. Examples of “`char`” representations of the letter ’a’ are: `'a'`, `'\097'`, `'\u0061'`, `'\U00000061'`.

· character
· Unicode
· code point
· “`char`”
· escape sequences

A *string* is a sequence of characters enclosed in double quotation marks. Examples are `"a"`, `"this is a string"`, and `"-&#\@"`. Note that the string `"a"` and the character `'a'` are not the same. Some strings are so common that they are given special names: One or more spaces, `" "` is called *whitespace*, and both `"\n"` and `"\r\n"` are called *newline*. The escape-character “\” may be used to break a line in two. This and other examples are shown in Listing 5.3.

· string
· whitespace
· newline

Type	syntax	Examples	Value
"int", "int32"	<int or hex>	3, 0x3	3
	<int or hex>l	3l, 0x3l	
"uint32"	<int or hex>u	3u	3
	<int or hex>ul	3ul	
"byte", "uint8"	<int or hex>uy	97uy	97
	'<char>'B	'a'B	
"byte[]"	"<string>"B	"a\n"B	[|97uy; 10uy|]
	@"<string>"B	@"a\n"B	[|97uy; 92uy; 110uy|]
"sbyte", "int8"	<int or hex>y	3y	3
"int16"	<int or hex>s	3s	3
"uint16"	<int or hex>us	3us	3
"int64"	<int or hex>L	3L	3
"uint64"	<int or hex>UL	3UL	3
	<int or hex>uL	3uL	
"float", "double"	<float>	3.0	3.0
	<hex>LF	0x013fLF	9.387247271e-323
"single", "float32"	<float>F	3.0F	3.0
	<float>f	3.0f	3.0
	<hex>lf	0x013flf	4.4701421e-43f
"decimal"	<float or int>M	3.0M,3M	3.0
	<float or int>m	3.0m,3m	
"string"	"<string>"	"a \"quote\".\n"	a "quote".<newline>
	@"<string>"	@"a ""quote"".\n"	a "quote".\n.
	"""<string>"""	"""a "quote".\n"""	a "quote".\n

Table 5.3: List of literal type. Syntax notation is used such that, e.g., <> means that the programmer replaces the brackets and content with a value on appropriate form. The [||] notation means that the value is an array, see Section 11.3 for details.

Listing 5.3: Examples of string literals.

```
> "abcde";;
val it : string = "abcde"

> "abc
-   de";;
val it : string = "abc
  de"

> "abc\
-   de";;
val it : string = "abcde"

> "abc\nde";;
val it : string = "abc
de"
```

Note that the response from `fsharpi` is shown in double quotation marks, but this is not part of the string.

F# supports *literal types*, where the type of a literal is indicated as a prefix og suffix as shown in the Table 5.3. The table uses a simple syntax notation such that <integer or hexadecimal>UL means that the user supplies an integer or a hexadecimal number followed by the characters 'UL'.

· literal type

The literalt type is closely connected to how the values are represented internally. E.g., a value of type "`int32`" use 32 bits and can be both positive and negative, while a "`uint32`" value also use 32 bits, but is unsigned. A "`byte`" is an 8-bit number, and "`sbyte`" is a signed 8-bit number. Values of type "`float`" uses 64 bits, while "`float32`" only uses 32 bits. The number of bits used to represent numbers directly relates to the range and precession these types can represent. This is summarized in Table 5.1 and discussed in more detail in Appendix B. Strings literals may be *verbatim* by the @-notation or triple double quotation marks, meaning that the escape sequences are not converted to their code point. The two types of string verbatim treat quotation marks differently as illustrated in the table. Further examples are shown in Listing 5.4.

· verbatim

Listing 5.4: Named and implied literals.

```
> 3;;
val it : int = 3

> 4u;;
val it : uint32 = 4u

> 5.6;;
val it : float = 5.6

> 7.9f;;
val it : float32 = 7.9000001f

> 'A';;
val it : char = 'A'

> 'B'B;;
val it : byte = 66uy

> "ABC";;
val it : string = "ABC"

> @"abc\nde";;
val it : string = "abc\nde"
```

Many basic types are compatible, and the type of a literal may be changed by *typecasting*. An example of casting to a "`float`" is shown in Listing 5.5.

· typecasting

Listing 5.5: Casting an integer to a floating point number.

```
> float 3;;
val it : float = 3.0
```

which is a `float`, since when `float` is given an argument, then it acts as a function rather than a type, and for the integer `3` it returns the floating point number `3.0`. For more on functions see Chapter 6. Boolean values are often treated as the integer values 0 and 1, but no short-hand function names exists for their conversions. Instead use functions from the `System.Convert` family of functions, as demonstrated in Listing 5.6.

Listing 5.6: Casting booleans.

```
> System.Convert.ToBoolean 1;;
val it : bool = true

> System.Convert.ToBoolean 0;;
val it : bool = false

> System.Convert.ToInt32 true;;
val it : int = 1

> System.Convert.ToInt32 false;;
val it : int = 0
```

Here `System.Convert.ToBoolean` is the identifier of a function `ToBoolean`, which is a *member* of the *class* `Convert` that is included in the *namespace* `System`. Namespaces, classes, and members will be discussed in Chapter 9.

· member
· class
· namespace

Typecasting is often a destructive operation, e.g., typecasting a `float` to `int` removes the fractional part without rounding as shown in Listing 5.7.

Listing 5.7: Fractional part is removed by downcasting.

```
> int 357.6;;
val it : int = 357
```

Here we typecasted to a lesser type, in the sense that the set of integers is a subset of floating point numbers, and this is called *downcasting*. The opposite is called *upcasting* and is often non-destructive, as Listing 5.5 showed, where an integer was casted to a float while retaining its value. As a side note, *rounding* a number $y.x$, where y is the *whole part* and x is the *fractional part*, is the operation of mapping numbers in the interval $y.x \in [y.0, y.5)$ to y and $y.x \in [y.5, y+1)$ to $y+1$. This can be performed by downcasting as shown in Listing 5.8.

· downcasting
· upcasting
· rounding
· whole part
· fractional part

Listing 5.8: Fractional part is removed by downcasting.

```
> int (357.6 + 0.5);;
val it : int = 358
```

Rounding is achieved in this way by downcasting, since if $y.x \in [y.0, y.5)$, then $y.x + 0.5 \in [y.5, y+1)$, from which downcasting removes the fractional part resulting in y. And if $y.x \in [y.5, y+1)$, then $y.x + 0.5 \in [y+1, y+1.5)$, from which downcasting removes the fractional part resulting in $y+1$.

5.2 Operators on basic types

Listing 5.8 is an example of an arithmetic *expression* using an *binary operator* on written using *infix* notation. The "+" operator is binary, since it takes two arguments, and since it is written between its arguments, then it uses infix notation. Expressions is the basic building block of all F# programs and this section will discuss operator expressions on basic types.

· expression
· binary operator
· infix

The syntax of basic binary operators is shown in Listing 5.9.

Listing 5.9 Syntax for a binary expression.

```
<expr><op><expr>
```

Here `<expr>` is any expression supplied by the programmer, and `<op>` is a binary, infix operator. F# supports a range of arithmetic binary infix operators on its built-in types such as addition, subtraction, multiplication, division, and exponentiation using the "`+`", "`-`", "`*`", "`/`", "`**`" lexemes. Not all operators are defined for all types, e.g., addition is defined for integer and float types as well as for characters and strings, but multiplication is only defined for integer and floating-point types. A complete list of built-in operators on basic types is shown in Table E.1 and E.2 and a range of mathematical functions shown in Table E.3. An example is Examples are `3+4`. Note that expressions can themselves be arguments to expressions , and thus, `4+5+6` is also a legal statement. This is called *recursion*, which means that a rule or a function is used by the rule or function itself in its definition. See Chapter 13 for more on recursive functions.

· recursion

Unary operators takes only one argument and have the syntax shown in Listing 5.10

Listing 5.10 A unary expressions.

```
<op><expr>
```

An example of a unary operator is `-3`, where `-` here is used to negate a positive integer. Since the operator appears before the operand it is a *prefix operator*.

· prefix operator

The concept of *precedence* is an important concept in arithmetic expressions.[1] If parentheses are omitted in Listing 5.8, then F# will interpret the expression as `(int 357.6) + 0.5`, which is erroneous, since addition of an integer with a float is undefined. This is an example of precedence, i.e., function evaluation takes precedence over addition meaning that it is performed before addition. Consider the arithmetic expression,

· precedence

Listing 5.11: A simple arithmetic expression.

```
> 3 + 4 * 5;;
val it : int = 23
```

Here, the addition and multiplication functions are shown in *infix notation* with the *operator* lexemes "`+`" and "`*`". To arrive at the resulting value 23, F# has to decide in which order to perform the calculation. There are 2 possible orders, `3 + (4 * 5)` or `(3 + 4) * 5`, which gives different results. For integer arithmetic, the correct order is of course to multiply before addition, and we say that multiplication takes *precedence* over addition. Every atomic operation that F# can perform is ordered in terms of its precedence, and for some common built-in operators shown in Table E.5, the precedence is shown by the order they are given in the table.

· infix notation
· operator
· precedence

Associativity implies the order in which calculations are performed for operators of same precedence. For some operators and type combinations association matters little, e.g., multiplication associates to the left and exponentiation associates to the right, as demonstrated in Listing 5.12.

[1] Jon: **minor comment on indexing and slice-ranges.**

Operator	Associativity	Description
+<expr>, -<expr>, ~~~<expr>	Left	Unary identity, negation, and bitwise negation operator
f <expr>	Left	Function application
<expr> ** <expr>	Right	Exponent
<expr> * <expr>, <expr> / <expr>, <expr> % <expr>	Left	Multiplication, division and remainder
<expr> + <expr>, <expr> - <expr>	Left	Addition and subtraction binary operators
<expr> ^^^ <expr>	Right	bitwise exclusive or
<expr> < <expr>, <expr> <= <expr>, <expr> > <expr>, <expr> >= <expr>, <expr> = <expr>, <expr> <> <expr>, <expr> <<< <expr>, <expr> >>> <expr>, <expr> &&& <expr>, <expr> ||| <expr>,	Left	Comparison operators, bitwise shift, and bitwise 'and' and 'or'.
<expr> && <expr>	Left	Boolean and
<expr> || <expr>	Left	Boolean or

Table 5.4: Some common operators, their precedence, and their associativity. Rows are ordered from highest to lowest precedences, such that `<*expr*> * <*expr*>` has higher precedence than `<*expr*> + <*expr*>`. Operators in the same row has same precedence. Full table is given in Table E.6.

Listing 5.12: Precedence rules define implicit parentheses.

```
> 3.0*4.0*5.0;;
val it : float = 60.0

> (3.0*4.0)*5.0;;
val it : float = 60.0

> 3.0*(4.0*5.0);;
val it : float = 60.0

> 4.0 ** 3.0 ** 2.0;;
val it : float = 262144.0

> (4.0 ** 3.0) ** 2.0;;
val it : float = 4096.0

> 4.0 ** (3.0 ** 2.0);;
val it : float = 262144.0
```

the expression for `3.0 * 4.0 * 5.0` associates to the left, and thus is interpreted as `(3.0 * 4.0) * 5.0`, but gives the same results as `3.0 * (4.0 * 5.0)`, since association does not matter for multiplication of numbers. However, the expression for `4.0 ** 3.0 ** 2.0` associates to the right, and thus is interpreted as `4.0 ** (3.0 ** 2.0)`, which is quite different from `(4.0 ** 3.0) ** 2.0`. **Whenever in doubt of association or any other basic semantic rules, it is a good idea to use parentheses as here. It is also a good idea to test your understanding of the syntax and semantic**

Advice

a	b	a && b	a || b	not a
false	false	false	false	true
false	true	false	true	true
true	false	false	true	false
true	true	true	true	false

Table 5.5: Truth table for boolean 'and', 'or', and 'not' operators. Value 0 is false and 1 is true.

rules by making a simple script.

5.3 Boolean arithmetic

Boolean arithmetic is the basis of almost all computers and particularly important for controlling program flow, which will be discussed in Chapter 8. Boolean values are one of 2 possible values, true or false, which is also sometimes written as 1 and 0. Basic operations on Boolean values are '*and*', '*or*', and '*not*', which in F# is written as the binary operators `&&`, `||`, and the function `not`. Since the domain of Boolean values is so small, then all possible combination of input on these values can be written on tabular form, known as a *truth table*, and the truth tables for the basic Boolean operators and functions is shown in Table 5.5. A good mnemonic for remembering the result of the 'and' and 'or' operators is to use 1 for true, 0 for false, multiplication for the Boolean 'and' operator, and addition for Boolean 'or' operator, e.g., true and false in this mnemonic translates to $1 \cdot 0 = 0$, and the results translates back to the Boolean value false. In F# the truth table for the basic Boolean operators can be produced by a simple program as shown in Listing 5.13.

- and
- or
- not
- truth table

Listing 5.13: Boolean operators and truth tables.

```
> printfn "a b a*b a+b not a"
- printfn "%A %A %A %A %A"
-   false false (false && false) (false || false) (not false)
- printfn "%A %A %A %A %A"
-   false true (false && true) (false || true) (not false)
- printfn "%A %A %A %A %A"
-   true false (true && false) (true || false) (not true)
- printfn "%A %A %A %A %A"
-   true true (true && true) (true || true) (not true);;
a b a*b a+b not a
false false false false true
false true false true true
true false false true false
true true true true false
val it : unit = ()
```

Here, we used the `printfn` function, to present the results of many expressions on something that resembles a tabular form. The spacing produced using the `printfn` function is not elegant, and in Section 6.5 we will discuss better options for producing more beautiful output. Notice, that the arguments for `printfn` was given on the next line with indentation. The indentation is an important part of telling F#, which part of what you write belongs together. This is an example of the so-called lightweight syntax. Generally, F# ignores newlines and whitespaces except when using the lightweight syntax, and the examples of the difference between regular and lightweight syntax is discussed in Chapter 6.

5.4 Integer arithmetic

The set of integers is infinitely large, but since all computers have limited resources, it is not possible to represent it in their entirety. The various integer types listed in Table 5.1 are finite subsets reduced by limiting their ranges. An in-depth description of integer implementation can be found in Appendix B. The type "`int`" is the most common type.

Table E.1, E.2, and E.3 gives examples operators and functions pre-defined for integer types. Notice that fewer functions are available for integers than for floating point numbers. For most addition, subtraction, multiplication, and negation the result straight forward. However, performing arithmetic operations on integers requires extra care, since the result may cause *overflow* and *underflow*. E.g., the range of the integer type `sbyte` is $[-128 \ldots 127]$, which causes problems in the example in Listing 5.14.

· overflow
· underflow

Listing 5.14: Adding integers may cause overflow.

```
> 100y;;
val it : sbyte = 100y

> 30y;;
val it : sbyte = 30y

> 100y + 30y;;
val it : sbyte = -126y
```

Here $100 + 30 = 130$, which is larger than the biggest `sbyte`, and the result is an overflow. Similarly, we get an underflow, when the arithmetic result falls below the smallest value storable in an `sbyte` as demonstrated in Listing 5.15.

Listing 5.15: Subtracting integers may cause underflow.

```
> -100y - 30y;;
val it : sbyte = 126y
```

I.e., we were expecting a negative number, but got a positive number instead.

The overflow error in Listing 5.14 can be understood in terms of the binary representation of integers: In binary, $130 = 10000010_2$, and this binary pattern is interpreted differently as `byte` and `sbyte`, see Listing 5.16.

Listing 5.16: The left most bit is interpreted differently for signed and unsigned integers, which gives rise to potential overflow errors.

```
> 0b10000010uy;;
val it : byte = 130uy

> 0b10000010y;;
val it : sbyte = -126y
```

That is, for signed bytes, the left-most bit is used to represent the sign, and since the addition of $100 = 01100100_2$ and $30 = 00011110_b$ is $130 = 10000010_2$ causes the left-most bit to be used, then this is wrongly interpreted as a negative number, when stored in an `sbyte`. Similar arguments can be made explaining underflows.

The division and remainder operators, which discards the fractional part after division, and the *remainder* operator calculates the remainder after integer division, as demonstrated in Listing 5.17.

· integer division
· remainder

Listing 5.17: Integer division and remainder operators.

```
> 7 / 3;;
val it : int = 2

> 7 % 3;;
val it : int = 1
```

Together integer division and remainder is a lossless representation of the original number, see Listing 5.18.

Listing 5.18: Integer division and remainder is a lossless representation of an integer, compare with Listing 5.17.

```
> (7 / 3) * 3;;
val it : int = 6

> (7 / 3) * 3 + (7 % 3);;
val it : int = 7
```

And we see that integer division of 7 by 3 followed by multiplication by 3 is less that 7, and the difference is `7 % 3`.

Notice that neither overflow nor underflow error gave rise to an error message, which is why such bugs are difficult to find. Dividing any non-zero number with 0 is infinite, which is also outside the domain of any of the integer types, but in this case, F# casts an *exception* as shown in Listing 5.19.

· exception

Listing 5.19: Integer division by zero causes an exception runtime error.

```
> 3/0;;
System.DivideByZeroException: Attempted to divide by zero.
  at <StartupCode$FSI_0002>.$FSI_0002.main@ () [0x00001] in
   <da48886c466e4b80be4566752c596fa8>:0
  at (wrapper managed-to-native)
   System.Reflection.MonoMethod:InternalInvoke
   (System.Reflection.MonoMethod,object,object[],System.Exception&)
  at System.Reflection.MonoMethod.Invoke (System.Object obj,
   System.Reflection.BindingFlags invokeAttr, System.Reflection.Binder
   binder, System.Object[] parameters, System.Globalization.CultureInfo
   culture) [0x00032] in <c9f8153c41de4f8cbafd0e32f9bf6b28>:0
Stopped due to error
```

The output looks daunting at first sight, but the first and last line of the error message are the most important parts, which tells us what exception was cast and why the program stopped. The middle are technical details concerning which part of the program caused this, and can be ignored for the time being. Exceptions are a type of *runtime error*, and are treated in Chapter 18

· runtime error

Integer exponentiation is not defined as an operator, but this is available the built-in function `pown`. This function is demonstrated in Listing 5.20.

a	b	a ~~~ b
false	false	false
false	true	true
true	false	true
true	true	false

Table 5.6: Boolean exclusive or truth table.

Listing 5.20: Integer exponent function.

```
> pown 2 5;;
val it : int = 32
```

which is equal to 2^5.

For binary arithmetic on integers, the following operators are available: `<leftExpr> <<< <rightExpr>`, which shifts the bit pattern of `<leftExpr> <rightExpr>` positions to the left while inserting 0's to right; `<leftExpr> >>> <rightExpr>`, which shifts the bit pattern of `<leftExpr> <rightExpr>` positions to the right while inserting 0's to left; `<expr> &&& <expr>`, bitwise 'and', returns the result of taking the Boolean 'and' operator position-wise; `<expr> ||| <expr>`, bitwise 'or', as 'and' but using the Boolean 'or' operator; and `<expr> ~~~ <expr>`, bitwise xor, which is returns the result of the Boolean 'xor' operator defined by the truth table in Table 5.6.

· xor
· exclusive or

5.5 Floating point arithmetic

The set of reals is infinitely large, and since all computers have limited resources, it is not possible to represent it in their entirety. Floating point types are finite subsets reduced by sampling the space of reals. An in-depth description of floating point implementations can be found in Appendix B. The type "`float`" is the most common type.

Table E.1, E.2, and E.3 gives examples operators and functions pre-defined for floating point types. For most addition, subtraction, multiplication, divisions, and negation the result straight forward. The remainder operator for floats calculates the remainder after division and discarding the fractional part, see Listing 5.21.

Listing 5.21: Floating point division and remainder operators.

```
> 7.0 / 2.5;;
val it : float = 2.8

> 7.0 % 2.5;;
val it : float = 2.0
```

The remainder for floating point numbers can be fractional, but division, discarding fractional part, and remainder is still a lossless representation of the original number as demonstrated in Listing 5.22.

Listing 5.22: Floating point division, truncation, and remainder is a lossless representation of a number.

```
> float (int (7.0 / 2.5));;
val it : float = 2.0

> (float (int (7.0 / 2.5))) * 2.5;;
val it : float = 5.0

> (float (int (7.0 / 2.5))) * 2.5 + 7.0 % 2.5;;
val it : float = 7.0
```

Arithmetic using `float` will not cause over- and underflow problems, since the IEEE 754 standard includes the special numbers $\pm\infty$ and NaN. As shown in Listing 5.23, no exception is thrown.

Listing 5.23: Floating point numbers include infinity and Not-a-Number.

```
> 1.0/0.0;;
val it : float = infinity

> 0.0/0.0;;
val it : float = nan
```

However, the `float` type has limited precision, since there is only a finite number of numbers that can be stored in a float. E.g., addition and subtraction can give surprising results as demonstrated in Listing 5.24.

Listing 5.24: Floating point arithmetic has finite precision.

```
> 357.8 + 0.1 - 357.9;;
val it : float = 5.684341886e-14
```

That is, addition and subtraction associates to the left, hence the expression is interpreted as `(357.8 + 0.1) - 357.9`, and we see that we do not get the expected 0, since only a limited number of floating point values are available, and the numbers `357.8 + 0.1` and `357.9` do not result in the same floating-point representation. Such errors tend to accumulate and comparing the result of expressions of floating point values should therefore be treated with care. Thus, **equivalence of two floating point expressions should only be considered up to sufficient precision, e.g., comparing `357.8 + 0.1` and `357.9` up to `1e-10` precision should be tested as, `abs ((357.8 + 0.1) - 357.9) < 1e-10`.** Advice

5.6 Char and string arithmetic

Addition is the only operator defined for characters, nevertheless, character arithmetic is often done by casting to integer. A typical example is conversion of case, e.g., to convert the lowercase character 'z' to uppercase, we use the *ASCIIbetical order* and add the difference between any Basic Latin Block letters in upper- and lowercase as `integers` and cast back to `char`, see Listing 5.25. · ASCIIbetical order

Listing 5.25: Converting case by casting and integer arithmetic.

```
> char (int 'z' - int 'a' + int 'A');;
val it : char = 'Z'
```

I.e., the code point difference between upper and lower case for any alphabetical character 'a' to 'z' is constant, hence we can change case by adding or subtracting the difference between any corresponding character. Unfortunately, this does not generalize to characters from other languages.

A large collection of operators and functions exist for `string`. The simplest is concatenation using the "+" operator as demonstrated in Listing 5.26.

Listing 5.26: Example of string concatenation.

```
> "hello" + " " + "world";;
val it : string = "hello world"
```

Characters and strings cannot be concatenated, which is why the above example used the string of a space `" "` instead of the space character `' '`. The characters of a string may be indexed as using the *.[]* notation. This is demonstrated in Listing 5.27.

Listing 5.27: String indexing using square brackets.

```
> "abcdefg".[0];;
val it : char = 'a'

> "abcdefg".[3];;
val it : char = 'd'

> "abcdefg".[3..];;
val it : string = "defg"

> "abcdefg".[..3];;
val it : string = "abcd"

> "abcdefg".[1..3];;
val it : string = "bcd"

> "abcdefg".[*];;
val it : string = "abcdefg"
```

Notice, that the first character has index 0, and to get the last character in a string, we use the string's `length` property. This is done as shown in Listing 5.28.

Listing 5.28: String length attribute and string indexing.

```
> "abcdefg".Length;;
val it : int = 7

> "abcdefg".[7-1];;
val it : char = 'g'
```

Since index counting starts at 0, and the string length is 7, then the index of the last character is 6. The is a long list of built-in functions in `System.String` for working with strings, some of which will be discussed in Appendix F.1.

The *dot notation* is an example of Structured programming, where technically speaking, the string `"abcdefg"` is an immutable *object* of *class* `string`, `[]` is an object *method*, and `Length` is a property. For more on object, classes, and methods see Chapter 20.

· dot notation
· object
· class
· method

Strings are compared letter by letter. For two strings to be equal, they must have the same length and all the letters must be identical. E.g., `"abs" = "absalon"` is false, while `"abs"\space = "abs"` is true. The "<>" operator is the boolean negation of the "=" operator, e.g., `"abs" <> "absalon"` is true, while `"abs" <> "abs"` is false. For the "<" , "<=", ">", and ">=" operators, the strings are ordered alphabetically, such that `"abs" < "absalon" && "absalon" < "milk"` is true, that is, the "<" operator on two strings is true, if the left operand should come before the right, when sorting alphabetically. The algorithm for deciding the boolean value of `leftOp < rightOp` is as follows: we start by examining the first character, and if `leftOp.[0]` and `rightOp.[0]` are different, then the `leftOp < rightOp` is equal to `leftOp.[0] < rightOp.[0]`. E.g., `"milk" < "abs"` is the same as `'m' < 'a'`, which is false, since the letter 'm' does not come before the letter 'a' in the alphabet, or more precisely, the codepoint of 'm' is not less than the codepoint of 'a'. If `leftOp.[0]` and `rightOp.[0]` are equal, then we move onto the next letter and repeat the investigation, e.g., `"abe" < "abs"` is true, since `"ab" = "ab"` is true and `'e' < 's'` is true. If we reach the end of either of the two strings, then the short is smaller than the larger, e.g., `"abs" < "absalon"` is true, while `"abs" < "abs"` is false. The "<=", ">", and ">=" operators are defined similarly.

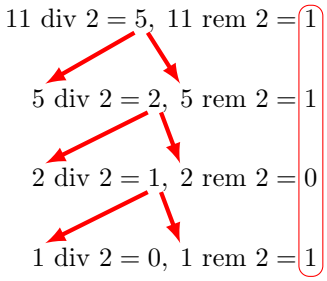
5.7 Programming intermezzo: Hand conversion between decimal and binary numbers

Conversion of integers between decimal and binary form is a key concept in order to understand some of the basic properties of calculations on the computer. From binary to decimal is straight forward using the power-of-two algorithm, i.e., given a sequence of $n+1$ bits that represent an integer $b_n b_{n-1} \ldots b_0$, where b_n and b_0 are the most and least significant bits, then the decimal value is calculated as,

$$v = \sum_{i=0}^{n} b_i 2^i \tag{5.1}$$

For example, $10011_2 = 1 + 2 + 16 = 19$. From decimal to binary is a little more complex, but a simple divide-by-two algorithm exists. The key to understanding the divide-by-two algorithm is to realize that when you divide a number by two, then that is equivalent to shifting its binary representation 1 to the right. E.g., $10 = 1010_2$ and $10/2 = 5 = 101_2$. Odd numbers have $b_0 = 1$, e.g., $11_{10} = 1011_2$ and $11_{10}/2 = 5.5 = 101.1_2$. Hence, if we divide any number by two and get a non-integer number, then its least significant bit was 1. Another way to express this is that the least significant bit is the remainder after integer division by two. Sequential application of this idea leads directly to the divide-by-two algorithm. E.g., if we were to convert the number 11_{10} on decimal form to binary form we would

perform the following steps:



Here we used div and rem to signify the integer division and remainder operators. The algorithm stops, when the result of integer division is zero. Reading off the remainder from below and up we find the sequence 1011_2, which is the binary form of the decimal number 11_{10}. Using interactive mode, we can calculate the same as shown in Listing 5.29.

Listing 5.29: Converting the number 11_{10} to binary form.

```
> printfn "(%d, %d)" (11 / 2) (11 % 2);;
(5, 1)
val it : unit = ()
> printfn "(%d, %d)" (5 / 2) (5 % 2);;
(2, 1)
val it : unit = ()
> printfn "(%d, %d)" (2 / 2) (2 % 2);;
(1, 0)
val it : unit = ()
> printfn "(%d, %d)" (1 / 2) (1 % 2);;
(0, 1)
val it : unit = ()
```

Thus, but reading the second integer-response from `printfn` from below and up, we again obtain the binary form of 11_{10} to be 1011_2. For integers with a fractional part, the divide-by-two may be used on the whole part, while multiply may be used in a similar manner on the fractional part.

Bibliography

[1] Alonzo Church. An unsolvable problem of elementary number theory. *American Journal of Mathematics*, 58:345—-363, 1936.

[2] Ole-Johan Dahl and Kristen Nygaard. SIMULA a language for programming and description of discrete event systems. introduction and user's manual. Technical report, Norwegian Computing Center, 1967.

[3] European Computer Manufacturers Association (ECMA). Standard ecma-335, common language infrastructure (cli). `http://www.ecma-international.org/publications/standards/Ecma-335.htm`.

[4] International Organization for Standardization. Iso/iec 23271:2012, common language infrastructure (cli). `https://www.iso.org/standard/58046.html`.

[5] Object Management Group. Uml version 2.0. `http://www.omg.org/spec/UML/2.0/`.

[6] Programming Research Group. Specifications for the ibm mathematical formula translating system, fortran. Technical report, Applied Science Division, International Business Machines Corporation, 1954.

[7] John McCarthy. Recursive functions of symbolic expressions and their computation by machine, part i. *Communications of the ACM*, 3(4):184–195, 1960.

[8] X3: ASA Sectional Committee on Computers and Information Processing. American standard code for information interchange. Technical Report ASA X3.4-1963, American Standards Association (ASA), 1963. `http://worldpowersystems.com/projects/codes/X3.4-1963/`.

[9] George Pólya. *How to solve it*. Princeton University Press, 1945.

[10] Allan M. Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2-42(1):230–265, 1936.

Index