

Learning to program with F#

Jon Spurring

September 20, 2016

Contents

1	Preface	5
2	Introduction	6
2.1	How to learn to program	6
2.2	How to solve problems	7
2.3	Approaches to programming	7
2.4	Why use F#	8
2.5	How to read this book	9
I	F# basics	10
3	Executing F# code	11
3.1	Source code	11
3.2	Executing programs	11
4	Quick-start guide	14
5	Using F# as a calculator	19
5.1	Literals and basic types	19
5.2	Operators on basic types	24
5.3	Boolean arithmetic	26
5.4	Integer arithmetic	27
5.5	Floating point arithmetic	29
5.6	Char and string arithmetic	31
5.7	Programming intermezzo	32

6	Constants, functions, and variables	34
6.1	Values	37
6.2	Non-recursive functions	42
6.3	User-defined operators	46
6.4	The Printf function	48
6.5	Variables	51
7	In-code documentation	57
8	Controlling program flow	62
8.1	For and while loops	62
8.2	Conditional expressions	66
8.3	Recursive functions	68
8.4	Programming intermezzo	71
9	Ordered series of data	71
9.1	Tuples	72
9.2	Lists	75
9.3	Arrays	79
10	Testing programs	84
10.1	White-box testing	87
10.2	Black-box testing	90
10.3	Debugging by tracing	93
11	Exceptions	100
12	Input and Output	108
12.1	Interacting with the console	109
12.2	Storing and retrieving data from a file	110
12.3	Working with files and directories.	113
12.4	Programming intermezzo	114

II	Imperative programming	121
13	Graphical User Interfaces	123
14	Imperative programming	124
14.1	Introduction	124
14.2	Generating random texts	125
14.2.1	0'th order statistics	125
14.2.2	1'th order statistics	125
III	Declarative programming	126
15	Sequences and computation expressions	127
15.1	Sequences	127
16	Patterns	133
16.1	Pattern matching	133
17	Types and measures	136
17.1	Unit of Measure	136
18	Functional programming	140
IV	Structured programming	143
19	Namespaces and Modules	144
20	Object-oriented programming	146
V	Appendix	147
A	Number systems on the computer	148
A.1	Binary numbers	150
A.2	IEEE 754 floating point standard	150

B	Commonly used character sets	151
B.1	ASCII	151
B.2	ISO/IEC 8859	152
B.3	Unicode	152
C	A brief introduction to Extended Backus-Naur Form	156
D	F_b	160
E	Language Details	165
E.1	Arithmetic operators on basic types	165
E.2	Basic arithmetic functions	168
E.3	Precedence and associativity	169
E.4	Lightweight Syntax	171
F	The Some Basic Libraries	172
F.1	System.String	173
F.2	List, arrays, and sequences	173
F.3	Mutable Collections	176
F.3.1	Mutable lists	176
F.3.2	Stacks	176
F.3.3	Queues	176
F.3.4	Sets and dictionaries	176
	Bibliography	177
	Index	178

Chapter 6

Constants, functions, and variables

In the previous chapter, we saw how to use F# as a calculator working with literals, operators and built-in functions. To save time and make programs easier to read and debug, it is useful to bind expressions to identifiers either as new constants, functions, or operators. As an example, consider the problem,

Problem 6.1:

For given set constants a , b , and c , solve for x in

$$ax^2 + bx + c = 0 \tag{6.1}$$

To solve for x we use the quadratic formula from elementary algebra,

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \tag{6.2}$$

which gives the general solution for any values of the coefficients. Here, we will assume a positive discriminant, $b^2 - 4ac > 0$. In order to write a program, where the code may be reused later, we define a function `discriminant : float -> float -> float -> float`, that is, a function that takes 3 arguments, `a`, `b`, and `c`, and calculates the discriminant. Details on function definition is given in Section 6.2. Likewise, we will define functions `positiveSolution : float -> float -> float -> float` and `negativeSolution : float -> float -> float -> float`, that also takes the polynomial's coefficients as arguments and calculates the solution corresponding to choosing the positive and negative sign for \pm in the equation. Our solution thus looks like Listing 6.1.

Listing 6.1, identifiersExample.fsx:

Finding roots for quadratic equations using function name binding.

```
let discriminant a b c = b ** 2.0 - 4.0 * a * c
let positiveSolution a b c = (-b + sqrt (discriminant a b c)) / (2.0 * a)
let negativeSolution a b c = (-b - sqrt (discriminant a b c)) / (2.0 * a)

let a = 1.0
let b = 0.0
let c = -1.0
let d = discriminant a b c
let xp = positiveSolution a b c
let xn = negativeSolution a b c
printfn "0 = %A * x ** 2.0 + %A * x + %A" a b c
printfn "    has discriminant %A and solutions %A and %A" d xn xp

-----

0 = 1.0 * x ** 2.0 + 0.0 * x + -1.0
    has discriminant 4.0 and solutions -1.0 and 1.0
```

Here, we have further defined names of values `a`, `b`, and `c` used as input to our functions, and the results of function application is bound to the names `d`, `xn`, and `xp`. The names of functions and values given here are examples of identifiers, and with these, we may reuse the quadratic formulas and calculated values later, while avoiding possible typing mistakes and reducing amount of code, which needs to be debugged.

Before we begin a deeper discussion note that F# adheres to two different syntax: regular and *lightweight*. In the regular syntax, newlines and whitespaces are generally ignored, while in lightweight syntax, certain keywords and lexemes may be replaced by specific use of newlines and whitespaces. Lightweight syntax is the most common, but the syntaxes may be mixed, and we will highlight the options, when relevant.

· lightweight
syntax

The use of identifiers is central in programming. For F# not to be confused by built-in functionality, identifiers must follow a specific grammar: An identifier must start with a letter, but can be followed by zero or more of letters, digits, and a range of special characters except SP, LF, and CR (space, line feed, and carriage return). An identifier must not be a keyword or a reserved-keyword listed in Figures 6.1. An identifier is a name for a constant, an expression, or a type, and it is defined by the following EBNF:

Keywords:

abstract, and, as, assert, base, begin, class, default, delegate, do, done, downcast, downto, elif, else, end, exception, extern, false, finally, for, fun, function, global, if, in, inherit, inline, interface, internal, lazy, let, match, member, module, mutable, namespace, new, null, of, open, or, override, private, public, rec, return, sig, static, struct, then, to, true, try, type, upcast, use, val, void, when, while, with, and yield.

Reserved keywords for possible future use:

atomic, break, checked, component, const, constraint, constructor, continue, eager, fixed, fori, functor, include, measure, method, mixin, object, parallel, params, process, protected, pure, recursive, sealed, tailcall, trait, virtual, and volatile.

Symbolic keywords:

let!, use!, do!, yield!, return!, |, ->, <-, ., :, (,), [,], [<, >], [|, |], {, }, ', #, :?>, :?, :>, .., ::, :=, ;;, ;, =, -, ?, ??, (*), <@, @>, <@@, and @@>.

Reserved symbolic keywords for possible future:

~ and `.

Figure 6.1: List of (possibly future) keywords and symbolic keywords in F#.

Listing 6.2: Identifiers

```
ident = (letter | "_" ) {letter | dDigit | specialChar};
longIdent = ident | ident "." longIdent; (*no space around "."*)

longIdentOrOp = [longIdent "."] identOrOp; (*no space around "."*)
identOrOp =
    ident
    | "(" infixOp | prefixOp ")"
    | "(*)";

dDigit = "0" | "1" | "2" | "3" | "4" | "5" | "6" | "7" | "8" | "9";
letter = Lu | Ll | Lt | Lm | Lo | Nl; (*e.g. "A", "B" ... and "a", "b", ...*)
specialChar = Pc | Mn | Mc | Cf; (*e.g., "_"*)

codePoint = ?Any unicode codepoint?;
Lu = ?Upper case letters?;
Ll = ?Lower case letters?;
Lt = ?Digraphic letters, with first part uppercase?;
Lm = ?Modifier letters?;
Lo = ?Gender ordinal indicators?;
Nl = ?Letterlike numeric characters?;
Pc = ?Low lines?;
Mn = ?Nonspacing combining marks?;
Mc = ?Spacing combining marks?;
Cf = ?Soft Hyphens?;
```

Thus, examples of identifiers are `a`, `theCharacter9`, `Next_Word`, `_tok`. Typically, only letters from the english alphabet are used as `letter`, and only `_` is used for `specialChar`, but the full definition refers to the Unicode general categories described in Appendix B.3, and there are currently 19.345 possible Unicode code points in the `letter` category and 2.245 possible Unicode code points in the `specialChar` category.

Expressions are a central concept in F#. An expression can be a mathematical expression, such as `3*5`,

a function application, such as $f3$, and many other things. Central in this chapter is the binding of values and functions to identifiers, which is done with the keyword “`let`”, using the following simplified syntax, e.g., `let a = 1.0`.

Expressions has an enormous variety in how they may be written, we will in this book gradually work through some of the more important facets. For this we will extend the EBNF notation with ellipses: `...`, to denote that what is shown is part of the complete EBNF production rule. E.g., the part of expressions, we will discuss in this chapter is specified in EBNF by,

Listing 6.3: Simple expressions.

```
expr = ...
  | expr ":" type (*type annotation*)
  | expr ";" expr (*sequence of expressions*)
  | "let" valueDefn "in" expr (*binding a value or variable*)
  | "let" ["rec"] functionDefn "in" expr (*binding a function or operator*)
  | "fun" argumentPats "->" expr (*anonymous function*)
  | expr "<-" expr (*assignment*)

type = ...
  | longIdent (*named such as "int"*)

valueDefn = ["mutable"] pat "=" expr;

pat = ...
  | "_" (*wildcard*)
  | ident (*named*)
  | pat ":" type (*type constraint*)
  | "(" pat ")" (*parenthesized*)

functionDefn = identOrOp argumentPats [":" type] "=" expr;
argumentPats = pat | pat argumentPats;
```

In the following sections, we will work through this bit by bit.

6.1 Values

Binding identifiers to literals or expressions that are evaluated to be values, is called value-binding, and examples are `let a = 3.0` and `let b = cos 0.9`. On EBNF the simplified syntax,

Listing 6.4: Value binding expression.

```
expr = ...
  | "let" valueDefn "in" expr (*binding a value or variable*)
```

The “`let`” bindings defines relations between patterns `pat` and expressions `expr` for many different purposes. Most often the pattern is an identifier `ident`, which “`let`” defines to be an alias of the expression `expr`. The pattern may also be defined to have specific type using the “`:`” lexeme and a named type. The “`_`” pattern is called the *wild card* pattern and, when it is in the value-binding, then the expression is evaluated but the result is discarded. The binding may be mutable as indicated by the keyword “`mutable`”, which will be discussed in Section 6.5, and the binding holds *lexically* for the last expression as indicated by the “`in`” keyword. For example, letting the identifier `p` be bound

- “`let`”
- “`:`”
- wild card
- “`mutable`”
- lexically
- “`in`”

to the value 2.0 and using it in an expression is done as follows,

Listing 6.5, letValue.fsx:

The identifier `p` is used in the expression following the “`in`” keyword.

```
let p = 2.0 in printfn "%A" (3.0 ** p)
```

```
9.0
```

F# will ignore most newlines between lexemes, i.e., the above is equivalent to writing,

Listing 6.6, letValueLF.fsx:

Newlines after “`in`” make the program easier to read.

```
let p = 2.0 in
printfn "%A" (3.0 ** p)
```

```
9.0
```

F# also allows for an alternative notation called *lightweight syntax*, where e.g., the “`in`” keyword is replaced with a newline, and the expression starts on the next line at the same column as “`let`” starts in, i.e., the above is equivalent to

· lightweight
syntax

Listing 6.7, letValueLightWeight.fsx:

Lightweight syntax does not require the “`in`” keyword, but expression must be aligned with the “`let`” keyword.

```
let p = 2.0
printfn "%A" (3.0 ** p)
```

```
9.0
```

The same expression in interactive mode will also respond the inferred types, e.g.,

Listing 6.8: Interactive mode also responds inferred types.

```
> let p = 2.0
- printfn "%A" (3.0 ** p);;
9.0
```

```
val p : float = 2.0
val it : unit = ()
```

By the “`val`” keyword in the line `val p : float = 2.0` we see that `p` is inferred to be of type `float` and bound to the value 2.0. The inference is based on the type of the right-hand-side, which is of type `float`. Identifiers may be defined to have a type using the “`:`” lexeme, but the types on the left-hand-side and right-hand-side of the “`=`” lexeme must be identical. I.e., mixing types gives an

error,

Listing 6.9, letValueTypeError.fsx:
Binding error due to type mismatch.

```
let p : float = 3
printfn "%A" (3.0 ** p)
```

```
/Users/sporring/repositories/fsharpNotes/src/letValueTypeError.fsx(1,17):
error FS0001: This expression was expected to have type
float
but here has type
int
```

Here, the left-hand-side is defined to be an identifier of type float, while the right-hand-side is a literal of type integer.

An expression can be a sequence of expressions separated by the lexeme “;”, e.g.,

Listing 6.10, letValueSequence.fsx:
A value-binding for a sequence of expressions.

```
let p = 2.0 in printfn "%A" p; printfn "%A" (3.0 ** p)
```

```
2.0
9.0
```

The lightweight syntax automatically inserts the “;” lexeme at newlines, hence using the lightweight syntax the above is the same as,

Listing 6.11, letValueSequenceLightWeight.fsx:
A value-binding for a sequence using lightweight syntax.

```
let p = 2.0
printfn "%A" p
printfn "%A" (3.0 ** p)
```

```
2.0
9.0
```

A key concept of programming is *scope*. In F#, the scope of a value-binding is lexically meaning that when F# determines the value bound to a name, it looks left and upward in the program text for the “let” statement defining it, e.g.,

Listing 6.12, letValueScopeLower.fsx:
Redefining identifiers is allowed in lower scopes.

```
let p = 3 in let p = 4 in printfn " %A" p;
```

4

F# also has to option of using dynamic scope, where the value of a binding is defined by when it is used, and this will be discussed in Section 6.5.

Scopes are given levels, and scopes may be nested, where the nested scope has a level one lower than its parent.¹ F# distinguishes between the top and lower levels, and at the top level in the lightweight syntax, redefining values is not allowed, e.g.,

Listing 6.13, letValueScopeLowerError.fsx:
Redefining identifiers is not allowed in lightweight syntax at top level.

```
let p = 3
let p = 4
printfn "%A" p;
```

```
/Users/sporring/repositories/fsharpNotes/src/letValueScopeLowerError.fsx
(2,5): error FS0037: Duplicate definition of value 'p'
```

But using parentheses, we create a *block*, i.e., a *nested scope*, and then redefining is allowed, e.g.,

· block
· nested scope

Listing 6.14, letValueScopeBlockAlternative3.fsx:
A block may be created using parentheses.

```
(
  let p = 3
  let p = 4
  printfn "%A" p
)
```

4

In both cases we used indentation, which is good practice, but not required here. Bindings inside are not available outside the nested scope, e.g.,

¹Todo: Drawings would be good to describe scope

Listing 6.15, letValueScopeNestedScope.fsx:
Bindings inside a scope are not available outside.

```
let p = 3
(
    let q = 4
    printfn "%A" q
)
printfn "%A %A" p q
```

```
/Users/sporring/repositories/fsharpNotes/src/letValueScopeNestedScope.fsx
(6,19): error FS0039: The value or constructor 'q' is not defined
```

Nesting is a natural part of structuring code, e.g., through function definitions to be discussed in Section 6.2 and flow control structures to be discussed in Chapter 8. Blocking code by nesting is a key concept for making robust code that is easy to use by others without the user necessarily needing to know the details of the inner workings of a block of code.

Defining blocks is useful for controlling the extend of a lexical scope of bindings. For example, adding a second `printfn` statement,

Listing 6.16, letValueScopeBlockProblem.fsx:
Overshadowing hides the first binding.

```
let p = 3 in let p = 4 in printfn "%A" p; printfn "%A" p
```

```
4
4
```

will print the value 4 last bound to the identifier `p`, since F# interprets the above as `let p = 3 in let p = 4 in (printfn "%A" p; printfn "%A" p)`. Had we intended to print the two different values of `p`, then we should have created a block as,

Listing 6.17, letValueScopeBlock.fsx:
Blocks allow for the return to the previous scope.

```
let p = 3 in (let p = 4 in printfn " %A" p); printfn " %A" p;
```

```
4
3
```

Here, the lexical scope of `let p = 4 in ...` is for the nested scope, which ends at “`)`”, returning to the lexical scope of `let p = 3 in`

6.2 Non-recursive functions

A function is a mapping between an input and output domain. A key advantage of using functions, when programming, is that they *encapsulate code* into smaller units, that are easier to debug and may be reused. F# is a functional first programming language, and offers a number of alternative methods for specifying parameters, which will be discussed in this section. Binding identifiers to functions follows a syntax similar to value-binding,

· encapsulate
code

Listing 6.18: Function binding expression

```
expr = ...  
| "let" functionDefn "in" expr (*binding a function or operator*)
```

Functions may also be recursive, which will be discussed in Chapter 8. An example in interactive mode is,

Listing 6.19: An example of a binding of an identifier and a function.

```
> let sum (x : float) (y : float) : float = x + y in  
- let c = sum 357.6 863.4 in  
- printfn "%A" c;;  
1221.0  
  
val sum : x:float -> y:float -> float  
val c : float = 1221.0  
val it : unit = ()
```

and we see that the function is interpreted to have the type `val sum : x:float -> y:float -> float`. The “->” lexeme means a mapping between sets, in this case floats. The function is also a higher order function, to be discussed in detail below, and here it suffices to think of `sum` as a function that takes 2 floats as argument and returns a float.

Not all types need to be declared, just sufficient for F# to be able to infer the types for the full statement. In the example, one specification is sufficient, and we could just have specified the type of the result,

Listing 6.20, letFunctionAlternative.fsx:
All types need most often not be specified.

```
let sum x y : float = x + y
```

or even just one of the arguments,

Listing 6.21, letFunctionAlternative2.fsx:
Just one type is often enough for F# to infer the rest.

```
let sum (x : float) y = x + y
```

In both cases, since the `+` operator is only defined for *operands* of the same type, then when the type of either the result, any or both operands are declared, then the type of the remaining follows directly.

· operator
· operand

As for values, lightweight syntax automatically inserts the keyword “`in`” and the lexeme “`;`”,

Listing 6.22, letFunctionLightWeight.fsx:
Lightweight syntax for function definitions.

```
let sum x y : float = x + y
let c = sum 357.6 863.4
printfn "%A" c
```

1221.0

Arguments need not always be inferred to types, but may be of generic type, which F# prefers, when *type safety* is ensured, e.g.,

· type safety

Listing 6.23: Typesafety implies that a function will work for any type, and hence it is generic.

```
> let second x y = y
- let a = second 3 5
- printfn "%A" a
- let b = second "horse" 5.0
- printfn "%A" b;;
5
5.0

val second : x:'a -> y:'b -> 'b
val a : int = 5
val b : float = 5.0
val it : unit = ()
```

Here, the function `second` does not use the first argument `x`, which therefore can be of any type, and which F# therefore calls `'a`, and the type of the second element, `y`, can also be of any type and not necessarily the same as `x`, so it is called `'b`. Finally the result is the same type as `y`, whatever it is. This is an example of a *generic function*, since it will work on any type.

· generic function

A function may contain a sequence of expressions, but must return a value. E.g., the quadratic formula may be written as,

Listing 6.24, identifiersExampleAdvance.fsx:
A function may contain sequences of expressions.

```
let solution a b c sgn =
  let discriminant a b c =
    b ** 2.0 - 2.0 * a * c
  let d = discriminant a b c
  (-b + sgn * sqrt d) / (2.0 * a)

let a = 1.0
let b = 0.0
let c = -1.0
let xp = solution a b c +1.0
let xn = solution a b c -1.0
printfn "0 = %A * x ** 2.0 + %A * x + %A" a b c
printfn "  has solutions %A and %A" xn xp
```

```
0 = 1.0 * x ** 2.0 + 0.0 * x + -1.0
  has solutions -0.7071067812 and 0.7071067812
```

Here, we used the lightweight syntax, where the “=” identifies the start of a nested scope, and F# identifies the scope by indentation. The amount of space used for indentation does not matter, but all lines following the first must use the same. The scope ends before the first line with the previous indentation or none. Notice how the last expression is not bound to an identifier, but is the result of the function, i.e., in contrast to many other languages, F# does not have an explicit keyword for returning values, but requires a final expression, which will be returned to the caller of the function. Note also that since the function `discriminant` is defined in the nested scope of `solution`, then `discriminant` cannot be called outside `solution`, since the scope ends before `let a = 1.0`.

Lexical scope and function definitions can be a cause of confusion as the following example shows,²

· lexical scope

Listing 6.25, lexicalScopeNFunction.fsx:
Lexical scope means that $f(z) = 3x$ and not $4x$ at the time of calling.

```
let testScope x =
  let a = 3.0
  let f z = a * z
  let a = 4.0
  f x
printfn "%A" (testScope 2.0)
```

```
6.0
```

Here, the value-binding for `a` is redefined, after it has been used to define a helper function `f`. So which value of `a` is used, when we later apply `f` to an argument? To resolve the confusion, remember that value-binding is lexically defined, i.e., the binding `let f z = a * x` uses the value of `a`, it has by the ordering of the lines in the script, not dynamically by when `f` was called. Hence, **think of lexical scope as substitution of an identifier with its value or function immediately at the place of definition**. I.e., since `a` and `3.0` are synonymous in the first lines of the program, then the function `f` is really defined as, `let f z = 3.0 * x`.

Advice

²Todo: Add a drawing or possibly a spell-out of lexical scope here.

Functions do not need a name, but may be declared as an *anonymous function* using the “`fun`” keyword and the “`->`” lexeme,

· anonymous function

Listing 6.26, functionDeclarationAnonymous.fsx:
Anonymous functions are functions as values.

```
let first = fun x y -> x
printfn "%d" (first 5 3)
```

5

Here, a name is bound to an anonymous function, which returns the first of two arguments. The difference to `let first x y = x` is that anonymous functions may be treated as values, meaning that they may be used as arguments to other functions, and new values may be reassigned to their identifiers, when mutable, as will be discussed in Section 6.5. A common use of anonymous functions is as arguments to other functions, e.g.,

Listing 6.27, functionDeclarationAnonymousAdvanced.fsx:
Anonymous functions are often used as arguments for other functions.

```
let apply f x y = f x y
let mul = fun a b -> a * b
printfn "%d" (apply mul 3 6)
```

18

Note that here `apply` is given 3 arguments, the function `mul` and 2 integers. It is not given the result of `mul 3 6`, since that would not match the definition of `apply`. **Anonymous functions and functions as arguments are powerful concepts, but tend to make programs harder to read, and their use should be limited.**

Advice

Functions may be declared from other functions

Listing 6.28, functionDeclarationCurrying.fsx:
A function can be defined as a subset of another by Currying.

```
let mul x y = x*y
let timesTwo = mul 2.0
printfn "%g" (mul 5.0 3.0)
printfn "%g" (timesTwo 3.0)
```

15

6

Here, `mul 2.0` is a partial specification of the function `mul x y`, where the first argument is fixed, and hence, `timesTwo` is a function of 1 argument being the second argument of `mul`. This notation is called *currying* in tribute of Haskell Curry, and Currying is often used in functional programming, but generally **currying should be used carefully, since currying may seriously reduce readability**

· currying
Advice

of code.

A *procedure* is a generalisation of the concept of functions, and in contrast to functions procedures need not return values, · procedure

Listing 6.29, procedure.fsx:

A procedure is a function that has no return value, and in F# returns “()”.

```
let printIt a = printfn "This is '%A'" a
printIt 3
printIt 3.0
```

```
This is '3'
This is '3.0'
```

In F# this is automatically given the unit type as return value. Procedural thinking is useful for *encapsulation* of scripts, but is prone to *side-effects* and should be minimized by being replaced by functional thinking. More on side-effects in Section 6.5. **Procedural thinking is useful for encapsulation, but is prone to side-effects and should be minimized by being replaced by functional thinking.**

· encapsulation
· side-effects
Advice

6.3 User-defined operators

Operators are functions, and in F#, the infix multiplication operator `+` is equivalent to the function `(+)`, e.g.,

Listing 6.30, addOperatorNFunction.fsx:

Operators have function equivalents.

```
let a = 3.0
let b = 4.0
let c = a + b
let d = (+) a b
printfn "%A plus %A is %A and %A" a b c d
```

```
3.0 plus 4.0 is 7.0 and 7.0
```

All operator has this option, and you may redefine them and define your own operators, but in F# names of user-defined operators are limited by the following simplified EBNF:

Listing 6.31: Grammar for infix and prefix lexemes

```
infixOrPrefixOp = "+" | "-" | "+." | "-." | "%" | "&" | "&&";
prefixOp = infixOrPrefixOp | "~" {"~"} | "!" {opChar} - "!=";
infixOp =
  {"."} (
    infixOrPrefixOp
    | "-" {opChar}
    | "+" {opChar}
    | "||"
    | "<" {opChar}
    | ">" {opChar}
    | "="
    | " |" {opChar}
    | "&" {opChar}
    | "^" {opChar}
    | "*" {opChar}
    | "/" {opChar}
    | "%" {opChar}
    | "!=" )
  | ":@" | "::" | "$" | "?";
opChar =
  "!" | "%" | "&" | "*" | "+" | "-" | "." | "/"
  | "<" | "=" | ">" | "@" | "^" | "|" | "~";
```

The precedence rules and associativity of user-defined operators follows the rules for which they share prefixes with built-in rules, see Table E.6. E.g., `.*`, `+++`, and `<+` are valid operator names for infix operators, they have precedence as ordered, and their associativity are all left. Using `~` as the first character in the definition of an operator makes the operator unary and will not be part of the name. Examples of definitions and use of operators are,

Listing 6.32, operatorDefinitions.fsx:

Operators may be (re)defined by their function equivalent.

```
let (.*) x y = x * y + 1
printfn "%A" (3 .* 4)
let (+++) x y = x * y + y
printfn "%A" (3 +++ 4)
let (<+) x y = x < y + 2.0
printfn "%A" (3.0 <+ 4.0)
let (~+.) x = x+1
printfn "%A" (+.1)
```

```
13
16
true
2
```

Operators beginning with `*` must use a space in its definition, (`*` in order for it not to be confused with the beginning of a comment (`*` , see Chapter 7 for more on comments in code).

Beware, redefining existing operators lexically redefines all future uses of the operators for all types, hence **it is not a good idea to redefine operators, but better to define new**. In Chapter 20

Advice

we will discuss how to define type specific operators including prefix operators.

6.4 The Printf function

A common way to output information to the console is to use one of the family of *printf* commands. These functions are special, since they take a variable number of arguments, and the number is decided by the first - the format string,

· *printf*

Listing 6.33: *printf* statement.

```
"printf" formatString {ident}
```

where a *formatString* is a string (simple or verbatim) with placeholders. The function *printf* prints *formatString* to the console, where all placeholder has been replaced by the value of the corresponding argument formatted as specified, e.g., in *printfn "1 2 %d" 3* the *formatString* is *"1 2 %d"*, and the placeholder is *%d*, and the *printf* replaced the placeholder with the value of the corresponding argument, and the result is printed to the console, in this case *1 2 3*. Possible formats for the placeholder are,

Listing 6.34: Placeholders in *formatString* for *printf* functions.

```
placeholder = "%%" | ("% [flags] [width] [ "." precision] specifier) (* No  
spaces between rules *)  
flags = ["0"] ["+"] [SP] (* No spaces between rules *)  
width = ["-"] ("*" | [dInt]) (* No spaces between rules *)  
specifier = "b" | "d" | "i" | "u" | "x" | "X" | "o" | "e" | "E" | "f" | "F" |  
"g" | "G" | "M" | "O" | "A" | "a" | "t"
```

There are specifiers for all the basic types and more as elaborated in Table 6.1. The placeholder can be given a specified with, either by setting a specific integer, or using the *** character, indicating that the with is given as an argument prior to the replacement value. Default is for the value to be right justified in the field, but left justification can be specified by the *-* character. For number types, you can specify their format by: *"O"* for padding the number with zeros to the left, when right justifying the number; *"+"* to explicitly show a plus sign for positive numbers; *SP* to enforce a space, where there otherwise would be a plus sign for positive numbers. For floating point numbers, the precision integer specifies the number of digits displayed of the fractional part. Examples of some of these combinations are,

Specifier	Type	Description
%b	bool	Replaces with boolean value
%s	string	
%c	char	
%d, %i	basic integer	
%u	basic unsigned integers	
%x	basic integer	formatted as unsigned hexadecimal with lower case letters
%X	basic integer	formatted as unsigned hexadecimal with upper case letters
%o	basic integer	formatted as unsigned octal integer
%f, %F,	basic floats	formatted on decimal form
%e, %E,	basic floats	formatted on scientific form. Lower case uses "e" while upper case uses "E" in the formatting.
%g, %G,	basic floats	formatted on the shortest of the corresponding decimal or scientific form.
%M	decimal	
%O	Objects ToString method	
%A	any built-in types	Formatted as a literal type
%a	Printf.TextWriterFormat -> 'a -> ()	
%t	(Printf.TextWriterFormat -> ()	

Table 6.1: Printf placeholder string

Listing 6.35, printfExample.fsx:
Examples of printf and some of its formatting options.

```

let pi = 3.1415192
let hello = "hello"
printf "An integer: %d\n" (int pi)
printf "A float %f on decimal form and on %e scientific form, and a char
      '%c'\n" pi pi
printf "A char '%c' and a string \"%s\"\n" hello.[0] hello
printf "Float using width 8 and 1 number after the decimal:\n"
printf "  \"%8.1f\" \"\n\" pi -pi
printf "  \"%08.1f\" \"%08.1f\"\n\" pi -pi
printf "  \"% 8.1f\" \"% 8.1f\"\n\" pi -pi
printf "  \"%-8.1f\" \"%-8.1f\"\n\" pi -pi
printf "  \"%+8.1f\" \"%+8.1f\"\n\" pi -pi
printf "  \"%8s\" \"\n\" \"%-8s\" \"\n\" "hello" "hello"

```

```

An integer: 3
A char 'h' and a string "hello"
Float using width 8 and 1 number after the decimal:
  "    3.1" "    -3.1"
  "000003.1" "-00003.1"
  "    3.1" "    -3.1"
  "3.1" "    -3.1"
  "    +3.1" "    -3.1"
  "    hello"
"hello"

```

Function	Example	Description
<code>printf</code> <code>printfn</code>	<code>printf "%d apples" 3</code>	Prints to the console, i.e., <code>stdout</code> as <code>printf</code> and adds a newline.
<code>fprintf</code> <code>fprintfn</code>	<code>fprintf stream "%d apples" 3</code>	Prints to a stream, e.g., <code>stderr</code> and <code>stdout</code> , which would be the same as <code>printf</code> and <code>eprintf</code> . as <code>fprintf</code> but with added newline.
<code>eprintf</code> <code>eprintfn</code>	<code>eprintf "%d apples" 3</code>	Print to <code>stderr</code> as <code>eprintf</code> but with added newline.
<code>sprintf</code>	<code>printf "%d apples" 3</code>	Return printed string
<code>failwithf</code>	<code>failwithf "%d failed apples" 3</code>	prints to a string and used for raising an exception.

Table 6.2: The family of printf functions.

Not all combinations of flags and identifier types are supported, e.g., strings cannot have number of integers after the decimal specified. The placeholder types `"%A"`, `"%a"`, and `"%t"` are special for F#, examples of their use are,

Listing 6.36, `printfExampleAdvance.fsx`:

Custom format functions may be used to specialize output.

```
let noArgument writer = printf "I will not print anything"
let customFormatter writer arg = printf "Custom formatter got: \"%A\"" arg
printf "Print examples: %A, %A, %A\n" 3.0m 3uy "a string"
printf "Print function with no arguments: %t\n" noArgument
printf "Print function with 1 argument: %a\n" customFormatter 3.0
```

```
Print examples: 3.0M, 3uy, "a string"
Print function with no arguments: I will not print anything
Print function with 1 argument: Custom formatter got: "3.0"
```

The `%A` is special in that all built-in types including tuples, lists, and arrays to be discussed in Chapter 9 can be printed using this formatting string, but notice that the formatting performed includes the named literal string. The two formatting strings `%t` and `%a` are options for user-customizing the formatting, and will not be discussed further.

Beware, `formatString` is not a `string` but a `Printf.TextWriterFormat`, so to predefine a `formatString` as, e.g., `let str = "hello %s" in printf str "world"` will be a type error.

The family of `printf` is shown in Table 6.2. The function `fprintf` prints to a stream, e.g., `stderr` and `stdout`, of type `System.IO.TextWriter`. Streams will be discussed in further detail in Chapter 12. The function `failwithf` is used with exceptions, see Chapter 11 for more details. The function has a number of possible return value types, and for testing the `ignore` function ignores it all, e.g., `ignore (failwithf "%d failed apples" 3)`

6.5 Variables

The “`mutable`” in “`let`” bindings means that the identifier may be rebound to a new value using the “`<-`” lexeme with the following syntax,³

· “`<-`”

Listing 6.37: Value reassignment for mutable variables.

```
expr = ...  
| expr "<-" expr (*assignment*)
```

Mutable data is synonymous with the term *variable*. A variable is an area in the computers working memory associated with an identifier and a type, and this area may be read from and written to during program execution. For example,

· Mutable data
· variable

Listing 6.38, mutableAssignReassingShort.fsx:
A variable is defined and later reassigned a new value.

```
let mutable x = 5  
printfn "%d" x  
x <- -3  
printfn "%d" x
```

```
5  
-3
```

Here, an area in memory was denoted `x`, initially assigned the integer value 5, hence the type was inferred to be `int`. Later, this value of `x` was replaced with another integer using the “<-” lexeme. The “<-” lexeme is used to distinguish the assignment from the comparison operator, i.e., if we by mistake had written,

· “<-”

Listing 6.39: It is a common error to mistake “=” and “<-” lexemes for mutable variables.

```
> let mutable a = 0  
- a = 3;;  
  
val mutable a : int = 0  
val it : bool = false
```

then we instead would have obtained the default assignment of the result of the comparison of the content of `a` with the integer 3, which is false. However, it is important to note, that when the variable is initially defined, then the “=” operator must be used, while later reassignments must use the “<-” expression.

Assignment type mismatches will result in an error,

³Todo: Discussion on heap and stack should be added here.

Listing 6.40, mutableAssignReassignTypeError.fsx:
Assignment type mismatching causes a compile time error.

```
let mutable x = 5
printfn "%d" x
x <- -3.0
printfn "%d" x
```

```
/Users/sporring/repositories/fsharpNotes/src/
  mutableAssignReassignTypeError.fsx(3,6): error FS0001: This expression
    was expected to have type
      int
but here has type
      float
```

I.e., once the type of an identifier has been declared or inferred, then it cannot be changed.

A typical variable is a counter of type integer, and a typical use of counters is to increment them, for example,

Listing 6.41, mutableAssignIncrement.fsx:
Variable increment is a common use of variables.

```
let mutable x = 5 // Declare a variable x and assign the value 5 to it
printfn "%d" x
x <- x + 1 // Assign a new value -3 to x
printfn "%d" x
```

```
5
6
```

Using variables in expressions as opposed to the left-hand-side of an assignment operation, reads the value of the variable. Thus, when using a variable as the return value of a function, then the value is copied from the local scope of the function to the scope from which it is called. E.g.,

Listing 6.42: Returning a mutable variable returns its value.

```
> let g () =
-   let mutable y = 0
-   y
-   printfn "%d" (g ());;
0

val g : unit -> int
val it : unit = ()
```

In the example, we see that the type is a value, and not mutable.

Variables implement dynamic scope, e.g., in comparison with the lexical scope, where the value of an

identifier depends on *where* it is defined, dynamic scope depends on, *when* it is used. E.g., the script in Listing 6.25 defines a function using lexical scope and returns the number 6.0, however, if `a` is made `mutable`, then the behaviour is different:

Listing 6.43, dynamicScopeNFunction.fsx:

Mutual variables implement dynamics scope rules. Compare with Listing 6.25.

```
let testScope x =  
  let mutable a = 3.0  
  let f z = a * x  
  a <- 4.0  
  f x  
printfn "%A" (testScope 2.0)
```

8.0

Here, the response is 8.0, since the value of `a` changed before the function `f` was called.

It is possible to work with mutable variables but through a special technique called *encapsulation*. E.g., in the following example we create a counter as an encapsulated mutable variable,

· encapsulation

Listing 6.44, mutableAssignIncrementEncapsulation.fsx:

Local mutable content can be indirectly accessed outside its scope.

```
let incr =  
  let mutable counter = 0  
  fun () ->  
    counter <- counter + 1  
    counter  
printfn "%d" (incr ())  
printfn "%d" (incr ())  
printfn "%d" (incr ())
```

1
2
3

This works because the line `let mutable counter = 0` is only executed once, when the function `incr` is defined. This is also an example of a side-effect. **Encapsulation of mutable data is good programming practice, but avoiding mutable data all together is better practice.**

Advice

F# has a variation of mutable variables called *reference cells*. Reference cells have built-in function `ref` and operators “!” and “:=”,

· reference cells

Listing 6.45, refCell.fsx:
Reference cells are variants of mutable variables.

```
let x = ref 0
printfn "%d" !x
x := !x + 1
printfn "%d" !x
```

0
1

That is, the `ref` function creates a reference variable, the “!” and the “:=” operators reads and writes its value. Reference cells are in some language called pointers, and their use is strongly discouraged, since they may cause *side-effects*, where variable changes are performed across independent scopes. The `incr` example in Listing 6.44 is an example of a side-effect. Another example is,

· side-effects

Listing 6.46, mutableAssignReturnSideEffect.fsx:
Intertwining independent scopes is typically a bad idea.

```
let updateFactor factor =
    factor := 2

let multiplyWithFactor x =
    let a = ref 1
    updateFactor a
    !a * x

printfn "%d" (multiplyWithFactor 3)
```

6

In the example, the function `updateFactor` changes a variable in the scope of `multiplyWithFactor`, which is prone to errors, since the computations are not local at the place of writing, i.e., in `multiplyWithFactor`, and if `updateFactor` were defined in a library, then the source code may not be available. Better style of programming is,

Listing 6.47, mutableAssignReturnWithoutSideEffect.fsx:
A solution of Listing 6.46 avoiding side-effects.

```
let updateFactor () =  
    2  
  
let multiplyWithFactor x =  
    let a = ref 1  
    a := updateFactor ()  
    !a * x  
  
printfn "%d" (multiplyWithFactor 3)
```

6

Here, there can be no doubt in `multiplyWithFactor` that the value of `a` is changing. Side-effects do have their use, but should in general be avoided at almost all costs, and in general it is advised to **minimize the use of side effects**.

Advice

Reference cells gives rise to an effect called *aliasing*, where two or more identifiers refer to the same data as illustrated by the following example:

· aliasing

Listing 6.48, refCellAliasing.fsx:
Aliasing can cause surprising results and should be avoided.

```
let a = ref 1  
let b = a  
printfn "%d, %d" !a !b  
b := 2  
printfn "%d, %d" !a !b
```

1, 1
2, 2

Here, `a` is defined as a reference cell, and by defining `b` to be equal to `a`, we have created an alias. This can be very confusing, since as the example shows, changing the value of `b` causes `a` to change as well. Aliasing is a variant of side-effects, and **aliasing should be avoided at all costs**.

Advice

Bibliography

- [1] Alonzo Church. A set of postulates for the foundation of logic. *Annals of Mathematics*, 33(2):346–366, 1932.
- [2] Programming Research Group. Specifications for the ibm mathematical formula translating system, fortran. Technical report, Applied Science Division, International Business Machines Corporation, 1954.
- [3] John McCarthy. Recursive functions of symbolic expressions and their computation by machine, part i. *Communications of the ACM*, 3(4):184–195, 1960.
- [4] 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/>.
- [5] George Pólya. *How to solve it*. Princeton University Press, 1945.

Index

. [], 31
ReadKey, 112
ReadLine, 112
Read, 112
System.Console.ReadKey, 112
System.Console.ReadLine, 112
System.Console.Read, 112
System.Console.WriteLine, 112
System.Console.Write, 112
WriteLine, 112
Write, 112
abs, 168
acos, 168
asin, 168
atan2, 168
atan, 168
bignum, 23
bool, 19
byte[], 23
byte, 23
ceil, 168
char, 19
cosh, 168
cos, 168
decimal, 23
double, 23
eprintfn, 51
eprintf, 51
exn, 19
exp, 168
failwithf, 51
float32, 23
float, 19
floor, 168
fprintfn, 51
fprintf, 51
ignore, 51
int16, 23
int32, 23
int64, 23
int8, 23
int, 19
it, 19
log10, 168
log, 168

max, 168
min, 168
nativeint, 23
obj, 19
pown, 168
printfn, 51
printf, 48, 51
round, 168
sbyte, 23
sign, 168
single, 23
sinh, 168
sin, 168
sprintf, 51
sqrt, 168
stderr, 51, 112
stdin, 112
stdout, 51, 112
string, 19
tanh, 168
tan, 168
uint16, 23
uint32, 23
uint64, 23
uint8, 23
unativeint, 23
unit, 19

aliasing, 55
American Standard Code for Information Inter-
change, 151
and, 26
anonymous function, 45
array sequence expressions, 131
Array.toArray, 84
Array.toList, 84
ASCII, 151
ASCIIbetical order, 31, 152

base, 19, 150
Basic Latin block, 152
Basic Multilingual plane, 152
basic types, 19
binary, 150
binary number, 21
binary operator, 25

- binary64, 150
- binding, 14
- bit, 21, 150
- black-box testing, 89
- block, 40
- blocks, 152
- boolean and, 169
- boolean or, 169
- branches, 66
- branching coverage, 91
- bug, 88
- byte, 150

- character, 21
- class, 24, 32
- code point, 21, 152
- compiled, 11
- computation expressions, 78, 83
- conditions, 66
- Cons, 80
- console, 11
- coverage, 91
- currying, 46

- debugging, 13, 89, 97
- decimal number, 19, 150
- decimal point, 20, 150
- Declarative programming, 8
- digit, 20, 150
- dot notation, 32
- double, 150
- downcasting, 24

- EBNF, 20, 156
- efficiency, 89
- encapsulate code, 42
- encapsulation, 46, 53
- environment, 98
- exception, 28
- exclusive or, 29
- executable file, 11
- expression, 14, 24
- expressions, 8
- Extended Backus-Naur Form, 20, 156
- Extensible Markup Language, 56

- file, 111
- floating point number, 20
- format string, 14
- fractional part, 20, 24
- function, 17
- function coverage, 91
- Functional programming, 8, 125
- functional programming, 8
- functionality, 88

- functions, 8

- generic function, 43

- hand tracing, 97
- Head, 80
- hexadecimal, 150
- hexadecimal number, 21
- HTML, 59
- Hyper Text Markup Language, 59

- IEEE 754 double precision floating-point format, 150
- Imperativ programming, 124
- Imperative programming, 8
- implementation file, 11
- infix notation, 25
- infix operator, 24
- integer, 20
- integer division, 28
- interactive, 11
- IsEmpty, 80
- Item, 80

- jagged arrays, 85

- keyword, 14

- Latin-1 Supplement block, 152
- Latin1, 152
- least significant bit, 150
- Length, 80
- length, 75
- lexeme, 17
- lexical scope, 16, 44
- lexically, 38
- lightweight syntax, 35, 38
- list, 78
- list sequence expression, 131
- List.Empty, 80
- List.toArray, 80
- List.toList, 80
- literal, 19
- literal type, 23

- machine code, 124
- maintainability, 89
- member, 24, 75
- method, 32
- mockup code, 97
- module elements, 144
- modules, 11
- most significant bit, 150
- Mutable data, 51
- mutually recursive, 69

- namespace, 24
- namespace pollution, 139
- NaN, 150
- nested scope, 40
- newline, 22
- not, 26
- not a number, 150

- obfuscation, 78
- object, 32
- Object oriented programming, 124
- Object-oriented programming, 8
- objects, 8
- octal, 150
- octal number, 21
- operand, 43
- operands, 25
- operator, 25, 43
- or, 26
- overflow, 27

- pattern matching, 133, 140
- portability, 89
- precedence, 25
- prefix operator, 25
- Procedural programming, 124
- procedure, 46
- production rules, 156

- ragged multidimensional list, 83
- raise an exception, 104
- range expression, 79
- reals, 150
- recursive function, 67
- reference cells, 54
- reliability, 88
- remainder, 28
- rounding, 24
- run-time error, 29

- scientific notation, 20
- scope, 40
- script file, 11
- script-fragment, 17
- script-fragments, 11
- Seq.initInfinite, 131
- Seq.item, 128
- Seq.take, 128
- Seq.toArray, 131
- Seq.toList, 131
- side-effect, 84
- side-effects, 46, 54
- signature file, 11
- slicing, 84
- software testing, 89

- state, 8
- statement, 14
- statement coverage, 91
- statements, 8, 124
- states, 124
- stopping criterium, 68
- stream, 112
- string, 14, 22
- Structured programming, 8
- subnormals, 150

- Tail, 80
- tail-recursive, 68
- terminal symbols, 156
- tracing, 97
- truth table, 26
- tuple, 75
- type, 15, 19
- type declaration, 15
- type inference, 13, 15
- type safety, 43
- typecasting, 23

- unary operator, 25
- underflow, 27
- Unicode, 21
- unicode general category, 152
- Unicode Standard, 152
- unit of measure, 136
- unit testing, 89
- unit-less, 137
- unit-testing, 13
- upcasting, 24
- usability, 89
- UTF-16, 154
- UTF-8, 154

- variable, 51
- verbatim, 23

- white-box testing, 89, 91
- whitespace, 22
- whole part, 20, 24
- wild card, 38
- word, 150

- XML, 56
- xor, 29

- yield bang, 128