COSE212: Programming Languages

Lecture 3 — Functional Programming in OCaml (1)

Hakjoo Oh
2016 Fall

Why learn ML?

Learning ML is a good way of experiencing modern language features:

- functional programming: scala, java8, haskell, python, JavaScript, etc
- value-oriented programming: scala, haskell, scheme, etc
- type inference: scala, haskell, etc
- pattern matching: scala, etc
- algebraic data types, module system, etc

Today: Basics of the Language

- Expressions
- Names
- Functions
- Pattern matching
- Type inference

Write and run all examples in the slides by yourself!

An OCaml Program is an Expression

Statement and expressions:

- A statement *does something*.
- An expression *evaluates to a value*.

Programming languages can be classified into

- statement-oriented: C, C++, Java, Python, JavaScript, etc
  - often called "imperative languages"
- expression-oriented: ML, Haskell, Scala, Lisp, etc
  - often called "functional languages"

Arithmetic Expressions

- Arithmetic expressions evaluate to numbers: e.g., `1+2*3`, `1+5`, `7`
- Try to evaluate expressions in the REPL:

```
# 1+2*3;;
- : int = 7
```

- Arithmetic operators on integers:

a `+` b	addition
a `-` b	subtraction
a `*` b	multiplication
a `/` b	divide a by b, returning the whole part
a `mod` b	divide a by b, returning the remaining part

Boolean Expressions

- Boolean expressions evaluate to boolean values (i.e., true, false).
- Try to evaluate boolean expressions:

```
# true;;
- : bool = true
# true;;
- : bool = true
# 1 > 2;;
- : bool = false
```

- Comparison operators produces boolean values:

a = b	true if a and b are equal
a <> b	true if a and b are not equal
a < b	true if a is less than b
a <= b	true if a is less than or equal to b
a > b	true if a is greater than b
a >= b	true if a is greater than or equal to b

Boolean Operators

- Boolean expressions are combined by boolean operators:

```
# true && false;;
- : bool = false
# true || false;;
- : bool = true
# (2 > 1) && (3 > 2);;
- : bool = true
```

ML is a Statically Typed Language

If you try to evaluate an expression that does not make sense, OCaml rejects and does not evaluate the program: e.g.,

```
# 1 + true;;
Error: This expression has type bool but an expression was
expected of type int
```

cf) Static Types and Dynamic Types

Programming languages are classified into:

- *Statically typed languages*: type checking is done at compile-time.
  - type errors are detected before program executions
  - C, C++, Java, ML, Scala, etc
- *Dynamically typed languages*: type checking is done at run-time.
  - type errors are detected during program executions
  - Python, JavaScript, Ruby, Lisp, etc

Statically typed languages are further classified into:

- *Type-safe languages* guarantee that compiled programs do not have type errors at run-time.
  - All type errors are detected at compile time.
  - Compiled programs do not stuck.
  - ML, Haskell, Scala
- *Unsafe languages* do not provide such a guarantee.
  - Some type errors remain at run-time.
  - C, C++

Which one is better?

Statically typed languages:

- (+) Type errors are caught early in the development cycle.
- (+) Program execution is efficient by omitting runtime checks.
- (−) Less flexible than dynamic languages.

Dynamically typed languages:

- (−) Type errors appear at run-time, often unexpectedly.
- (+) Provide more flexible language features.
- (+) Easy and fast prototyping.

Conversion between Different Types

- In OCaml, different types of values are distinguished:

```
# 3 + 2.0;;
Error: This expression has type float but an expression
was expected of type int
```

- Types must be explicitly converted:

```
# 3 + int_of_float 2.0;;
- : int = 5
```

- Operators for floating point numbers:

```
# 1.2 +. 2.3;;
- : float = 3.5
# 1.5 *. 2.0;;
- : float = 3.
# float_of_int 1 +. 2.2;;
- : float = 3.2
```

Other Primitive Values

- OCaml provides six primitive values: integers, booleans, floating point numbers, characters, strings, and unit.

```
# 'c';;
- : char = 'c'
# "cose212";;
- : string = "cose212"
# ();;
- : unit = ()
```

Conditional Expressions

$$\texttt{if } \boldsymbol{be} \texttt{ then } \boldsymbol{e_1} \texttt{ else } \boldsymbol{e_2}$$

- If $\boldsymbol{be}$ is true, the value of the conditional expression is the value of $\boldsymbol{e_1}$.
- If $\boldsymbol{be}$ is false, the value of the expression is the value of $\boldsymbol{e_2}$.
- ```
  # if 2 > 1 then 0 else 1;;
  - : int = 0
  # if 2 < 1 then 0 else 1;;
  - : int = 1
  ```
- $\boldsymbol{be}$ must be a boolean expression.
- types of $\boldsymbol{e_1}$ and $\boldsymbol{e_2}$ must be equivalent.
- ```
  # if 1 then 1 else 2;;
  Error: ...
  # if true then 1 else true;;
  Error: ...
  # if true then true else false;;
  - : bool = true
  ```

Names and Functions

- Create a global variable with the `let` keyword:

```
# let x = 3 + 4;;
val x : int = 7
```

We say a variable `x` is *bound* to value $\mathbf{7}$.

```
# let y = x + x;;
val y : int = 14
```

- Create a local variable with `let ... in ...` construct:

$$\texttt{let } \boldsymbol{x} \texttt{ = } \boldsymbol{e_1} \texttt{ in } \boldsymbol{e_2}$$

  - $\boldsymbol{x}$ is bound to the value of $\boldsymbol{e_1}$
  - the scope of $\boldsymbol{x}$ is $\boldsymbol{e_2}$
  - the value of $\boldsymbol{e_2}$ becomes the value of the entire expression

Examples

- ```
  # let a = 1 in a;;
  - : int = 1
  # let a = 1 in a * 2;;
  - : int = 2
  ```
- ```
  # let a = 1 in
    let b = a + a in
    let c = b + b in
      c + c;;
  - : int = 8
  ```
- ```
  # let d =
      let a = 1 in
      let b = a + a in
      let c = b + b in
        c + c;;
   val d : int = 8
  ```
  ```

Functions

- Define a function with `let`:

```
# let square x = x * x;;
 val square : int -> int = <fun>
```

- Apply the function:

```
# square 2;;
- : int = 4
# square (2 + 5);;
- : int = 49
# square (square 2);;
- : int = 16
```

- The body can be any expression:

```
# let neg x = if x < 0 then true else false;;
val neg : int -> bool = <fun>
# neg 1;;
- : bool = false
# neg (-1);;
- : bool = true
```

Functions

- Functions with multiple arguments:

```
# let sum_of_squares x y = (square x) + (square y);;
val sum_of_squares : int -> int -> int = <fun>
# sum_of_squares 3 4;;
- : int = 25
```

- Recursive functions are defined with `let rec` construct:

```
# let rec factorial a =
    if a = 1 then 1 else a * factorial (a - 1);;
val factorial : int -> int = <fun>
# factorial 5;;
- : int = 120
```

Nameless Functions

- Many modern programming languages provide nameless functions, e.g., ML, Scala, Java8, JavaScript, Python, etc.
- In OCaml, a function can be defined without names:

  ```
  # fun x -> x * x;;
  - : int -> int = <fun>
  ```

  Called *nameless* or *anonymous* functions.
- Apply nameless function as usual:

  ```
  # (fun x -> x * x) 2;;
  - : int = 4
  ```

- A variable can be bound to functions:

  ```
  # let square = fun x -> x * x;;
  val square : int -> int = <fun>
  ```

- The followings are equivalent:

  ```
  let square = fun x -> x * x
  ```

  ```
  let square x = x * x
  ```

Example: Square Roots with Newton's Method

We implement the function

$$\texttt{sqrt} : \texttt{float} -> \texttt{float}$$

using Newton's method. To compute `sqrt(x)`,

- Start with an initial guess y for the value of the square root of x (pick $y = 1$).
- Repeatedly improve the estimate by taking the mean of y and x/y.

ex)

Estimation	Quotient	Mean
$\mathbf{1}$	$\mathbf{2/1 = 2}$	$\mathbf{1.5}$
$\mathbf{1.5}$	$\mathbf{2/1.5 = 1.3333}$	$\mathbf{1.4167}$
$\mathbf{1.4167}$	$\mathbf{2/1.4167 = 1.4118}$	$\mathbf{1.4142}$
$\mathbf{1.4142}$	$\ldots$	$\ldots$

Example: Square Roots with Newton's Method

```
let is_good_enough guess x =
  abs_float (guess *. guess -. x) < 0.001

let improve guess x = (guess +. x /. guess) /. 2.0

let rec sqrt_iter guess x =
  if is_good_enough guess x then guess
  else sqrt_iter (improve guess x) x

let sqrt x = sqrt_iter 1.0 x
```

Example: Square Roots with Newton's Method

```
# #use "sqrt.ml";;
val is_good_enough : float -> float -> bool = <fun>
val improve : float -> float -> float = <fun>
val sqrt_iter : float -> float -> float = <fun>
val sqrt : float -> float = <fun>
# sqrt 9.0;;
- : float = 3.00009155413138
# sqrt (sqrt 2.0 +. sqrt 3.0);;
- : float = 1.77392790232078923
# sqrt 1000.0 *. sqrt 1000.0;;
- : float = 1000.00036992436605
```

Pattern Matching

- An elegant way of doing case analysis.
- E.g., using pattern-matching, the factorial function

  ```
  let rec factorial a =
    if a = 1 then 1 else a * factorial (a - 1)
  ```

  can be written as follows:

  ```
  let factorial a =
    match a with
     1 -> 1
    |_ -> a * factorial (a - 1)
  ```

Pattern Matching

The nested if-then-else expression

```
let isabc c = if c = 'a' then true
              else if c = 'b' then true
              else if c = 'c' then true
              else false
```

can be written using pattern matching:

```
let isabc c =
  match c with
   'a' -> true
 |'b' -> true
 |'c' -> true
 | _  -> false
```

or simply,

```
 let isabc c =
  match c with
   'a' | 'b' | 'c' -> true
   | _ -> false
```

Type Inference

In C or Java, types must be annotated:

```
public static int f(int n)
{
  int a = 2;
  return a * n;
}
```

In OCaml, type annotations are not mandatory:

```
# let f n =
    let a = 2 in
      a * n;;
val f : int -> int = <fun>
```

Type Inference

OCaml can infer types, no matter how complex the program is:

```
# let sum_if_true test first second =
    (if test first then first else 0)
  + (if test second then second else 0);;
val sum_if_true : (int -> bool) -> int -> int -> int = <fun>
```

OCaml compiler infers the type through the following reasoning steps:

1. the types of `first` and `second` must be `int`, because both branches of a conditional expression must have the same type,
2. the type of `test` is a function type $\alpha \to \beta$, because `test` is used as a function,
3. α must be of `int`, because `test` is applied to `first`, a value of `int`,
4. β must be of `bool`, because conditions must be boolean expressions,
5. the return value of the function has type `int`, because the two conditional expressions are of `int` and their addition gives `int`.

Type Annotation

Explicit type annotations are possible:

```
# let sum_if_true (test : int -> bool) (x : int) (y : int) : int =
    (if test x then x else 0) + (if test y then y else 0);;
val sum_if_true : (int -> bool) -> int -> int -> int = <fun>
```

If the annotation is wrong, OCaml finds the error and report it:

```
# let sum_if_true (test : int -> int) (x : int) (y : int) : int =
    (if test x then x else 0) + (if test y then y else 0);;
Error: The expression (test x) has type int but an expression
was expected of type bool
```

Polymorphic Types

- What is the type of the program?

  ```
  let id x = x
  ```

  See how OCaml infers its type:

  ```
  # let id x = x;;
  val id : 'a -> 'a = <fun>
  ```

  The function works for values of any type:

  ```
  # id 1;;
  - : int = 1
  # id "abc";;
  - : string = "abc"
  # id true;;
  - : bool = true
  ```

- Such a function is called *polymorphic* and `'a` is a *type variable*.

Polymorphic Types

Quiz) What is the type of the function?

```
let first_if_true test x y =
  if test x then x else y
```