

Developing Applications With
Objective Caml

Emmanuel Chailloux

Pascal Manoury

Bruno Pagano

2000

Оригинал этой книги (ISBN 2-84177-121-0) был опубликован во Франции издательством O'Reilly & Associates под названием *Developpement d'applications avec Objective Caml*.

©O'REILLY & ASSOCIATES, 2000

Édition: Xavier CAZIN.

Les programmeurs figurant dans ce livre ont pour but d'illustrer les sujets traités. Il n'est donnée aucune garantie quant à leur fonctionnement une fois compilés, assemblés ou interprétés dans le cadre d'une utilisation professionnelle ou commerciale.

©Éditions O'Reilly, Paris, 2000 ISBN

Toute représentation ou reproduction, intégrale ou partielle, faite sans le consentement de l'auteur, de ses ayants droit, ou ayants cause, est illicite (loi du 11 mars 1957, alinéa 1^{er} de l'article 40). Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait une contrefaçon sanctionnée par les articles 425 et suivants du Code pénal. La loi du 11 mars 1957 autorise uniquement, aux termes des alinéas 2 et 3 de l'article 41, les copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective d'une part et, d'autre part, les analyses et les courtes citations dans un but d'exemple et d'illustration.

Оглавление

I	Основы языка	5
1	Функциональное программирование	7
1.1	Функциональное ядро Objective CAML	9
1.1.1	Значения, функции и базовые типы	9
1.1.2	Структуры условного контроля	16
1.1.3	Объявление значений	16
1.1.4	Функциональное выражение, функции	19
1.1.5	Полиморфизм и ограничение типа	28
1.1.6	Примеры	31
1.2	Объявления типов и сопоставление с образцом	34
1.2.1	Сопоставление с образцом	34
1.2.2	Декларация типов	43
1.2.3	Записи	45
1.2.4	Тип сумма	47
1.2.5	Рекурсивный тип	50
1.2.6	Параметризованный тип	51
1.2.7	Видимость описания	52
1.2.8	Функциональные типы	53
1.2.9	Пример: реализация деревьев	53
1.2.10	Не функциональные рекурсивные значения	56
1.3	Типизация, область определения и исключения	57
1.3.1	Частичные функции и исключения	58
1.3.2	Определения исключения	59
1.3.3	Возбуждение исключения	60
1.3.4	Перехват исключения	61
1.4	Полиморфизм и значения возвращаемые функциями	62
1.5	Калькулятор	64
1.6	Резюме	68

2	Императивное программирование	69
2.1	Физически изменяемые структуры данных	71
2.1.1	Векторы	71
2.1.2	Строки	75
2.1.3	Изменяемые поля записей	76
2.1.4	Указатели	77
2.1.5	Полиморфизм и изменяемые значения	78
2.2	Ввод-вывод	80
2.2.1	Каналы	81
2.2.2	Чтение и запись	81
2.2.3	Пример «Больше/Меньше»	82
2.3	Структуры контроля	83
2.3.1	Последовательность	84
2.3.2	Циклы	86
2.3.3	Пример: реализация стека	87
2.3.4	Пример: расчет матриц	90
2.4	Порядок вычисления аргументов	91
2.5	Калькулятор с памятью	92
2.6	Упражнения	95
2.6.1	Двусвязный список	95
2.7	Резюме	95
3	Функциональный и императивный стиль	97
3.1	Сравнение между функциональным и императивным сти- лями	99
3.1.1	С функциональной стороны	99
3.1.2	Императивная сторона	100
3.1.3	Рекурсия или итерация	102
3.2	Какой стиль выбрать?	104
3.2.1	Последовательность или композиция функций	104
3.2.2	Общее использование или копии значений	107
3.2.3	Критерии выбора	110
3.3	Смесь стилей	112
3.3.1	Замыкание и побочные эффекты	113
3.3.2	Физическое изменение и исключения	114
3.3.3	Изменяемые функциональные структуры данных	115
3.3.4	Ленивые изменяемые структуры	117
3.4	Поток данных	121
3.4.1	Конструкция	122
3.4.2	Деструкция и сопоставление потоков	123
3.5	Резюме	127

Часть I

Основы языка

Глава 1

Функциональное программирование

Первый язык функционального программирования LISP появился в конце 1950, в тот же момент, что и Fortran — один из первых императивных языков. Оба эти языка существуют и по сей день, хотя они немало изменились. Область их применения: вычислительные задачи для Фортрана и символьные (symbolic) для Lisp. Интерес к функциональному программированию состоит в простоте написания программ, где под программой подразумевается функция, примененная к аргументам. Она вычисляет результат, который возвращается как вывод программы. Таким образом можно с легкостью комбинировать программы: вывод одной становится входным аргументом для другой.

Функциональное программирование основывается на простой модели вычислений, состоящей из трех конструкций: переменные, определение функции и ее применение к какому-либо аргументу. Эта модель, называемая λ -исчисление, была введена Alonzo Church в 1932, еще до появления первых компьютеров. В λ -исчислении любая функция является переменной, так она может быть использована как входной параметр другой функции или возвращена как результат другой. Теория λ -исчисления утверждает, что все то, что **вычисляемо**, может быть представлено этим формализмом. Однако, синтаксис этой теории слишком ограничен, чтобы его можно было использовать как язык программирования. В связи с этим были добавлены базовые типы (например, целые числа или строки символов), операторы для этих типов, управляющие структуры и объявления позволяющие именовать переменные или функции, и, в частности, рекурсивные функции.

Существует разные классификации языков функционального программирования. Мы будем различать их по двум характеристикам, которые нам кажутся наиболее важными:

- без побочных эффектов (чистые, или правильные) или с побочным эффектом (соответственно, не чистые): чистый язык — это тот, в котором присвоения не существует. Все есть вычисление, и как оно происходит, нас не интересует. Не чистые языки, такие как Caml или ML, имеют императивные особенности такие как присвоение. Они позволяют писать программы в стиле близкому к Фортрану, в котором важен порядок вычисления выражений.
- типизация динамическая или статическая: типизация необходима для проверки соответствия аргумента, переданного функции, типу формального параметра. Эта проверка может быть выполнена во время выполнения программы. В этом случае типизация называется динамической. В случае ошибки программа будет остановлена, как это происходит в Lisp. В случае статической типизации проверка осуществляется во время компиляции, то есть до выполнения программы. Таким образом она (проверка) не замедлит программу во время выполнения. Эта типизация используется в ML и в его диалектах, таких как *Objective CAML*. Только правильно типизированные программы, то есть успешно прошедшие проверку типов, могут быть скомпилированы и затем выполнены.

План главы

В этой главе представлены базовые элементы функциональной части языка *Objective CAML*, а именно: синтаксис, типы и механизм исключений. После этого вводного курса мы сможем реализовать нашу первую программу.

В первом разделе описаны основы языка, начиная с базовых типов и функций. Затем мы рассмотрим структурные и функциональные типы. После этого займемся управляющими структурами, а также локальными и глобальными объявлениями. Во втором разделе мы обсудим определение типов для создания структур и механизм сопоставления с образцом, который используется для доступа к этим структурам. В третьем разделе рассматривается выводимый тип функций и область их применения, после чего следует описание механизма исключений. Четвертый раздел объединяет введенные понятия в простой пример программы-калькулятора.

1.1 Функциональное ядро Objective CAML

Как любой другой язык функционального программирования, Objective CAML — это язык выражений, состоящий в основном из создания функций и их применения. Результатом вычисления одного из таких выражений является значение данного языка (*value in the language*) и выполнение программы заключается в вычислении всех выражений из которых она состоит.

1.1.1 Значения, функции и базовые типы

В Objective CAML определены следующие типы: целые числа, числа с плавающей запятой, символьный, строковый и логический.

Числа

Различают целые¹ `int` и числа с плавающей запятой `float`. Objective CAML следует спецификации IEEE 754 для представления чисел с плавающей запятой двойной точности². Операции с этими числами описаны в таблице 1.1. Если результат целочисленной операции выходит за интервал значений типа `int`, то это не приведет к ошибке. Результат будет находиться в интервале целых чисел, то есть все действия над целыми ограничены операцией `modulo` с границами интервала.

Разница между целыми числами и числами с плавающей запятой. Значения разных типов, таких как `float` и `int`, не могут сравниваться между собой напрямую. Для этого существует функции перевода одного типа в другой (`float_of_int` и `int_of_float`).

```
# 2 = 2.0 ;;
This expression has type float but is here used with type int
# 3.0 = float_of_int 3 ;;
- : bool = true
```

Аналогично, операции над целыми числами и числами с плавающей запятой различны:

¹На 32-х битных процессорах в интервале $[-2^{30}, 2^{30} - 1]$ и на 64-х битных процессорах в интервале $[-2^{62}, 2^{62} - 1]$.

²Числа с плавающей запятой $m * 10^n$ представлены в виде 53-х битной мантиссы m и экспоненты n в интервале $[-1022, 1023]$.

целые	с плавающей запятой
<code>+</code> сложение	<code>+. </code> сложение
<code>-</code> вычитание и унарный минус	<code>-.</code> вычитание и унарный минус
<code>*</code> умножение	<code>*.</code> умножение
<code>/</code> деление	<code>/. </code> деление
<code>mod</code> остаток целочисленного деления	<code>**</code> возведение в степень
<pre># 1;; - : int = 1 # 1 + 2 ;; - : int = 3 # 9 / 2 ;; - : int = 4 # 11 mod 3 ;; - : int = 2 # (* лимит представления *) (* целых чисел *) 2147483650 ;; - : int = 2</pre>	<pre># 2.0 ;; - : float = 2. # 1.1 +. 2.2 ;; - : float = 3.3 # 9.1 /. 2.2 ;; - : float = 4.13636363636 # 1. /. 0. ;; - : float = inf. # (* лимит представления *) (* чисел с *) (* плавающей запятой *) 22222222222.11111 ;; - : float = 22222222222.</pre>

Таблица 1.1: Операции над числами

```
# 3 + 2 ;;
- : int = 5
# 3.0 +. 2.0 ;;
- : float = 5.
# 3.0 + 2.0 ;;
This expression has type float but is here used with type int
# sin 3.14159 ;;
- : float = 2.65358979335e-06
```

Неопределенный результат, например получаемый при делении на ноль, приведет к возникновению исключения (см. 1.3, стр. 57), которое остановит вычисление. Числа с плавающей запятой имеют специальные значения для бесконечных величин (`Inf`) и для не определенного результата (`NaN`³). Основные операции над этими числами приведены в таблице 1.2.

³Not a Number

Ф-ии над числами с плавающей запятой	тригонометрические функции
<code>ceil</code> <code>floor</code> <code>sqrt</code> квадратный корень <code>exp</code> экспонента <code>log</code> натуральный логарифм <code>log10</code> логарифм по базе 10	<code>cos</code> косинус <code>sin</code> синус <code>tan</code> тангенс <code>acos</code> арк косинус <code>asin</code> арк синус <code>atan</code> арк тангенс
<pre># ceil 3.4 ;; - : float = 4. # floor 3.4 ;; - : float = 3. # ceil (-.3.4) ;; - : float = -3 # floor (-.3.4) ;; - : float = -4.</pre>	<pre># sin 1.57078 ;; - : float = 0.999999999867 # sin (asin 0.707) ;; - : float = 0.707 # acos 0.0 ;; - : float = 1.57079632679 # asin 3.14 ;; - : float = nan.</pre>

Таблица 1.2: Функции над числами с плавающей запятой

Символы и строки

Символы, тип `char`, соответствуют целым числам в интервале от 0 до 255, первые 128 значений соответствуют кодам ASCII. Функция `char_of_int` и `int_of_char` преобразуют один тип в другой. Строки, тип `string` — это последовательность символов определенной длины (не длиннее $2^{24} - 6$). Оператором объединения строк (конкатенации) является `^`. Следующие функции необходимы для преобразования типов `int_of_string`, `string_of_int`, `string_of_float` и `float_of_string`.

```
# 'B' ;;
- : char = 'B'
# int_of_char 'B' ;;
- : int = 66
# "is a string" ;;
- : string = "is a string"
# (string_of_int 1987) ^ " is the year Caml was created" ;;
- : string = "1987 is the year Caml was created"
```

Если строка состоит из цифр, то мы не сможем использовать ее в численных операциях, не выполнив явного преобразования.

```
# "1999" + 1 ;;
```

```
This expression has type string but is here used with type int
# (int_of_string "1999") + 1 ;;
- : int = 2000
```

В модуле `String` собрано много функций для работы со строками (стр. ??).

Булевый тип

Значение типа `bool` принадлежит множеству состоящему из двух элементов: `true` и `false`. Основные операторы описаны в таблице 1.3. По историческим причинам операторы «и» и «или» имеют две формы.

<code>not</code>	отрицание		
<code>&&</code>	последовательное и	<code>&</code>	синоним <code>&&</code>
<code> </code>	последовательное или	<code>or</code>	синоним <code> </code>

Таблица 1.3: Булевы операторы

```
# true ;;
- : bool = true
# not true ;;
- : bool = false
# true && false ;;
- : bool = false
```

Операторы `&&` и `||` или их синонимы, вычисляют аргумент слева и в зависимости от его значения, вычисляют правый аргумент. Они могут быть переписаны в виде условной структуры (см. 1.1.2 стр. 16).

<code>=</code>	равенство структурное	<code><</code>	меньше
<code>==</code>	равенство физическое	<code>></code>	больше
<code><></code>	отрицание =	<code><=</code>	меньше или равно
<code>!=</code>	отрицание ==	<code>>=</code>	больше или равно

Таблица 1.4: Операторы сравнения и равенства

Операторы сравнения и равенства описаны в таблице 1.4. Это полиморфные операторы, то есть они применимы как для сравнения двух целых, так и двух строк. Единственным ограничением это то что операнды должны быть одного типа (см. 1.1.5 стр. 28).

```
# 1<=118 && (1=2 || not(1=2)) ;;
- : bool = true
# 1.0 <= 118.0 && (1.0 = 2.0 || not(1.0 = 2.0)) ;;
- : bool = true
# "one" < "two" ;;
- : bool = true
# 0 < '0' ;;
```

This expression has type char but is here used with type int

Структурное равенство, при проверке двух переменных, сравнивает значение полей структуры, тогда как физическое равенство проверяет занимают ли эти переменные одно и то же место в памяти. Оба оператора возвращают одинаковый результат для простых типов: `bool`, `char`, `int` и константные конструкторы (см. 1.2.4 стр. 47).

Внимание

Числа с плавающей запятой и строки рассматриваются как структурные типы.

Unit

Тип `unit` определяет множество из всего одного элемента, значение которого: `()`.

```
# () ;;
- : unit = ()
```

Это значение будет широко использоваться в императивных программах (см. 2 стр. 69) в функциях с побочным эффектом. Функции, результат которых `()`, соответствуют понятию процедуры, которое отсутствует в Objective CAML, так же как и аналог типа `void` в языке C.

Декартово произведение, кортежи

Значения разных типов могут быть сгруппированы в кортежи. Значения из которых состоит кортеж разделяются запятой. Конструктор типа `*` означает кортеж. `int*string` есть кортеж, в котором первый элемент целое число и второй строка.

```
# (12, "October") ;;
- : int * string = (12, "October")
```

Иногда мы можем использовать более простую форму записи.

```
# 12, "October" ;;
- : int * string = (12, "October")
```

Функции *fst* и *snd* дают доступ первому и второму элементу соответственно.

```
# fst (12, "October") ;;
- : int = 12
# snd (12, "October") ;;
- : string = "October"
```

Эти обе функции полиморфные, входной аргумент может быть любого типа.

```
# fst ;;
- : 'a * 'b -> 'a = <fun>
# fst ("October", 12) ;;
- : string = "October"
```

Тип `int*char*string` — это триплет в котором первый элемент типа `int`, второй `char` и третий — `string`.

```
# (65, 'B', "ascii") ;;
- : int * char * string = (65, 'B', "ascii")
```

Внимание

Если аргумент функций *fst* и *snd* не пара, а другой *n*-кортеж, то мы получим ошибку.

```
# snd (65, 'B', "ascii") ;;
This expression has type int * char * string but is here used with type
'a * 'b
```

Существует разница между парой и триплетом: тип `int*int*int` отличается от `(int*int)*int` и `int*(int*int)`. Методы доступа к элементам триплета (и других кортежей) не определены в стандартной библиотеке. В случае необходимости мы используем сопоставление с образцом (см. [1.2.1](#)).

Списки

Значения одного и того же типа могут быть объединены в списки. Список может быть либо пустым, либо содержать однотипные элементы.

```
# [] ;;
- : 'a list = []
# [ 1; 2; 3 ] ;;
- : int list = [1; 2; 3]
# [ 1; "two"; 3 ] ;;
This expression has type string but is here used with type int
```

Для того чтобы добавить элемент в начало списка существует следующая функция в виде инфиксного оператора ‘::’ — аналог `cons` в Lisp.

```
# 1 :: 2 :: 3 :: [] ;;
- : int list = [1; 2; 3]
```

Для конкатенации списков существует инфиксный оператор: ‘@’.

```
# [ 1 ] @ [ 2; 3 ] ;;
- : int list = [1; 2; 3]
# [ 1; 2 ] @ [ 3 ] ;;
- : int list = [1; 2; 3]
```

Остальные функции манипуляции списками определены в библиотеке `List`. Функции `hd` и `tl` из этой библиотеки дают доступ к первому и последнему элементу списка. Эти функции обозначаются `List.hd` и `List.tl` для указания что их нужно искать в модуле `List`⁴.

```
# List.hd [ 1; 2; 3 ] ;;
- : int = 1
# List.hd [] ;;
Exception: Failure "hd".
```

В последнем примере получить последний элемент пустого списка действительно «сложно», поэтому возбуждается исключение (см. 1.3.1).

⁴Модуль `List` описан на стр. ??

1.1.2 Структуры условного контроля

Одна из структур контроля необходимая в каждом языка программирования — условный оператор (или альтернативный).

Синтаксис `if $expr_1$ then $expr_2$ else $expr_3$`

Тип выражения $expr_1$ — `bool`. Выражения $expr_2$ и $expr_3$ должны быть одного и того же типа.

```
# if 3=4 then 0 else 4 ;;
- : int = 4
# if 3=4 then "0" else "4" ;;
- : string = "4"
# if 3=4 then 0 else "4" ;;
This expression has type string but is here used with type int
```

Условное выражение, как и любое другое, возвращает значение.

```
# (if 3=5 then 8 else 10) + 5 ;;
- : int = 15
```

Примечание

Ветка **else** может быть опущена, в этом случае будет подставлено значение по умолчанию равное **else ()**, в соответствии с этим выражение **expr2** должно быть типа `unit` (см. [2.3](#) стр. 83).

1.1.3 Объявление значений

Определение связывает имя со значением. Различают глобальные и локальные определения. В первом случае, объявленные имена видны во всех выражениях, следуют за ним, во втором — имена доступны только в текущем выражении. Мы также можем *одновременно* объявить несколько пар имя-значение.

Глобальные объявления

Синтаксис `let name = expr ;;`

Глобальное объявление определяет связь имени *name* со значением выражения *expr*, которая будет доступно всем следующим выражениям.

```
# let yr = "1999" ;;
val yr : string = "1999"
# let x = int_of_string (yr) ;;
val x : int = 1999
# x ;;
- : int = 1999
# x + 1 ;;
- : int = 2000
# let new_yr = string_of_int (x + 1) ;;
val new_yr : string = "2000"
```

Одновременное глобальное объявление

Синтаксис `let name1 = expr1
and name2 = expr2
:
and namen = exprn ;;`

При одновременном объявлении переменные будут известны только к концу всех объявлений.

```
# let x = 1 and y = 2 ;;
val x : int = 1
val y : int = 2
# x + y ;;
- : int = 3
# let z = 3 and t = z + 2 ;;
Unbound value z
```

Можно сгруппировать несколько глобальных объявлений в одной фразе, вывод типов и значений произойдет к концу фразы, отмеченным двойной точкой с запятой `;;`. В данном случае объявления будут вычислены по порядку.

```
# let x = 2
  let y = x + 3 ;;
val x : int = 2
val y : int = 5
```

Глобальное объявление может быть скрыто локальным с тем же именем (см. 1.1.4 стр. 25).

Локальное объявление

Синтаксис `let name = expr1 in expr2 ;;`

Имя *name* связанное с выражением *expr*₁ известно только для вычисления *expr*₂.

```
# let x1 = 3 in x1 * x1 ;;
- : int = 9
```

Локальное объявление, которое связывает **x1** со значением 3, существует только в ходе вычисления **x1*x1**.

```
# x1 ;;
Unbound value x1
```

Локальное объявление скрывает любые глобальные с тем же именем, но как только мы выходим из блока в котором была оно определено, мы находим старое значение связанное с этим именем.

```
# let x = 2 ;;
val x : int = 2
# let x = 3 in x * x ;;
- : int = 9
# x * x ;;
- : int = 4
```

Локальное объявление — это обычное выражение, соответственно оно быть использовано для построения других выражений.

```
# (let x = 3 in x * x) + 1 ;;
- : int = 10
```

Локальные объявления так же могут быть одновременными.

Синтаксис

```
let name1 = expr1
and name2 = expr2
:
and namen = exprn
in expr ;;
```

```
# let a = 3.0 and b = 4.0 in sqrt (a*.a +. b*.b) ;;
- : float = 5.
# b ;;
Unbound value b
```

1.1.4 Функциональное выражение, функции

Функциональное выражение состоит из параметра и тела. Формальный параметр — это имя переменной, а тело — это выражение. Обычно говорят что формальный параметр является абстрактным, по этой причине функциональное выражение тоже называется абстракцией.

Синтаксис `function p -> expr`

Таким образом функция возведения в квадрат будет выглядеть так:

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

Objective Caml сам определяет тип. Функциональный тип `int->int` это функция с параметром типа `int`, возвращающая значение типа `int`.

Функция с одним аргументом пишется как функция и аргумент следующий за ней.

```
# (function x -> x * x) 5 ;;
- : int = 25
```

Вычисление функции состоит в вычисление ее тела, в данном случае `x*x`, где формальный параметр `x`, заменен значением аргумента (эффективный параметр), здесь он равен 5.

При конструкции функционального выражения `expr` может быть любым выражением, в частности функциональным.

```
# function x -> ( function y -> 3*x + y ) ;;
- : int -> int -> int = <fun>
```

Скобки не обязательны, мы можем писать просто:

```
# function x -> function y -> 3*x + y ;;
- : int -> int -> int = <fun>
```

В простом случае мы скажем что функция ожидает два аргумента целого типа на входе и возвращает значение целого типа. Но когда речь идет о функциональном языке, таком как Objective Caml , скорее это соответствует типу функции с входным аргументом типа `int` и возвращающая функциональное значение типа `int->int`:

```
# (function x -> function y -> 3*x + y) 5 ;;
- : int -> int = <fun>
```

Естественно, мы можем применить это функциональное выражение к двум аргументам. Для этого напишем:

```
# (function x -> function y -> 3*x + y) 4 5 ;;
- : int = 17
```

Когда мы пишем `f a b`, подразумевается применение `(f a)` к `b`. Давайте подробно рассмотрим выражение

```
(function x -> function y -> 3*x + y) 4 5
```

Для того чтобы вычислить это выражение, необходимо сначала вычислить значение

```
(function x -> function y -> 3*x + y) 4
```

что есть функциональное выражение равное

```
function y -> 3*4 + y
```

в котором `x` заменен на `4` в `3*x+y`. Применяя это значение (являющееся функцией) к `5`, мы получаем конечное значение `3*4+5=17`:

```
# (function x -> function y -> 3*x + y) 4 5 ;;
- : int = 17
```

Арность функции

Арностью функции называется число аргументов функции. По правилам, унаследованным из математики, аргументы функции задаются в скобках после имени функции. Мы пишем : `f (4, 5)`. Как было указано ранее, в Objective Caml мы чаще используем следующий синтаксис: `f 4 5`. Естественно, можно написать функциональное выражение применимое к (4,5):

```
# function (x,y) -> 3*x + y ;;
- : int * int -> int = <fun>
```

Но в данном случае, функция ожидает не два аргумента, а один; тип которого пара целых. Попытка применить два аргумента к функции ожидающей пару или наоборот, передать пару функции для двух аргументов приведет к ошибке:

```
# (function (x,y) -> 3*x + y) 4 5 ;;
This function is applied to too many arguments
# (function x -> function y -> 3*x + y) (4, 5) ;;
This expression has type int * int but is here used with type int
```

Альтернативный синтаксис

Существует более компактная форма записи функций с несколькими аргументами, которая дошла к нам из старых версий Caml. Выглядит она так:

Синтаксис	<code>fun p₁ ... p_n -> expr</code>
-----------	---

Это позволяет не повторять слово **function** и стрелки. Данная запись эквивалентна:

`function p1 -> ... -> function pn -> expr`

```
# fun x y -> 3*x + y ;;
- : int -> int -> int = <fun>
# (fun x y -> 3*x + y) 4 5 ;;
- : int = 17
```

Эту форму можно часто встретить в библиотеках идущих в поставку с Objective Caml.

Замыкание

Objective CAML рассматривает функциональное выражение также как любое другое. Значение возвращаемое функциональным выражением называется **замыканием**. Каждое выражение Objective Caml вычисляется в **окружении** состоящем из соответствий имя–значение, которые были объявлены до вычисляемого выражения. Замыкание может быть описано как триплет, состоящий из имени формального параметра, тела функции и окружения выражения. Нам необходимо хранить это окружение, поскольку в теле функции кроме формальных параметров могут использоваться другие переменные. Эти переменные называются «свободными» в функциональном выражении, нам понадобится их значение в момент применения функционального выражения.

```
# let m = 3 ;;
val m : int = 3
# function x -> x + m ;;
- : int -> int = <fun>
# (function x -> x + m) 5 ;;
- : int = 8
```

В случае когда применение замыкания к аргументу возвращает новое замыкание, оно (новое замыкание) получает в свое окружение все необходимые связи для следующего применения. Параграф 1.1.4 подробно рассматривает эти понятия. В главе 3, а так же в главе ?? мы вернемся к тому как замыкание представляется в памяти.

До сих пор рассмотренные функциональные выражения были **анонимными**, однако мы можем дать им имя.

Объявление функциональных значений

Функциональное значение объявляется так же как и другие, при помощи конструктора **let**

```
# let succ = function x -> x + 1 ;;
val succ : int -> int = <fun>
# succ 420 ;;
- : int = 421
# let g = function x -> function y -> 2*x + 3*y ;;
val g : int -> int -> int = <fun>
# g 1 2;;
- : int = 8
```

Для упрощения записи, можно использовать следующий синтаксис:

Синтаксис `let name p1 ... pn = expr`

что эквивалентно:

`let name = function p1 -> ... -> function pn -> expr`

Следующие объявления `succ` и `g` эквивалентны предыдущим:

```
# let succ x = x + 1 ;;
val succ : int -> int = <fun>
# let g x y = 2*x + 3*y ;;
val g : int -> int -> int = <fun>
```

В следующем примере демонстрируется функциональная сторона Objective Caml, где функция `h1` получена применением `g` к аргументу. В данном случае мы имеем дело с **частичным применением**.

```
# let h1 = g 1 ;;
val h1 : int -> int = <fun>
# h1 2 ;;
- : int = 8
```

С помощью `g` мы можем определить другую функцию `h2` фиксируя значение второго параметра (`y`):

```
# let h2 = function x -> g x 2 ;;
val h2 : int -> int = <fun>
# h2 1 ;;
- : int = 8
```

Объявление инфиксных функций

Некоторые бинарные функции могут быть использованы в инфиксной форме. Например при сложении двух целых мы пишем `3 + 5` для применения `+` к 3 и 5. Для того чтобы использовать символ `+` как классическое функциональное значение, необходимо указать это, окружая символ скобками.

Синтаксис `(op)`

В следующем примере определяется функция `succ` используя `(+)`:

```
# (+) ;;
- : int -> int -> int = <fun>
# let succ = (+) 1 ;;
val succ : int -> int = <fun>
# succ 3 ;;
- : int = 4
```

Таким образом мы можем определить новые операторы, что мы и сделаем определив ++ для сложения двух пар целых.

```
# let (++) c1 c2 = (fst c1)+(fst c2), (snd c1)+(snd c2) ;;
val ++ : int * int -> int * int -> int * int = <fun>
# let c = (2,3) ;;
val c : int * int = 2, 3
# c ++ c ;;
- : int * int = 4, 6
```

Существуют однако ограничения на определение новых операторов, они должны содержать только символы (такие как '*', '+', '@', и т.д.), исключая буквы и цифры. Следующие функции являются исключением: *or mod land lor lxor lsr asr*

Функции высшего порядка

Функциональное значение (замыкание) может быть возвращено как результат, а так же передано как аргумент функции. Такие функции, берущие на входе или возвращающие функциональные значения, называются функциями высшего порядка.

```
# let h = function f -> function y -> (f y) + y ;;
val h : (int -> int) -> int -> int = <fun>
```

Примечание

Выражения группируются справа налево, но функциональные типы объединяются слева направо. Таким образом тип функции *h* может быть написан:

```
(int -> int) -> int -> int
```

или

```
(int -> int) -> (int -> int)
```


При помощи функций высшего порядка можно элегантно обрабатывать списки. К примеру функция *List.map* применяет какую-нибудь функцию ко всем элементам списка и возвращает список результатов.

```
# List.map ;;
- : ('a -> 'b) -> 'a list -> 'b list = <fun>
# let square x = string_of_int (x*x) ;;
val square : int -> string = <fun>
# List.map square [1; 2; 3; 4] ;;
- : string list = ["1"; "4"; "9"; "16"]
```

Другой пример — функция *List.for_all* проверяет соответствуют ли элементы списка определенному критерию.

```
# List.for_all ;;
- : ('a -> bool) -> 'a list -> bool = <fun>
# List.for_all (function n -> n<>0) [-3; -2; -1; 1; 2; 3] ;;
- : bool = true
# List.for_all (function n -> n<>0) [-3; -2; 0; 1; 2; 3] ;;
- : bool = false
```

Видимость переменных

Для вычисления выражения, необходимо чтобы все используемые им переменные были определены, как, например, выражение *e* в определении **let** *p=e* переменная *p* не известна в этом выражении, она может быть использована только в случае если *p* была объявлена ранее.

```
# let p = p ^ "-suffix" ;;
Unbound value p
# let p = "prefix" ;;
val p : string = "prefix"
# let p = p ^ "-suffix" ;;
val p : string = "prefix-suffix"
```

В Objective Caml переменные связаны статически. При применении замыкания используется окружение в момент ее (замыкания) объявления (статическая видимость), а не в момент ее применения (динамическая видимость)

```
# let p = 10 ;;
val p : int = 10
```

```
# let k x = (x, p, x+p) ;;
val k : int -> int * int * int = <fun>
# k p ;;
- : int * int * int = 10, 10, 20
# let p = 1000 ;;
val p : int = 1000
# k p ;;
- : int * int * int = 1000, 10, 1010
```

В функции `k` имеется свободная переменная `p`, которая была определена в глобальном окружении, поэтому определение `k` принято. Связь между именем `p` и значением `10` в окружении замыкания `k` статическая, то есть не зависит от последнего определения `p`.

Рекурсивное объявление

Объявление переменной называется рекурсивным, если оно использует свой собственный идентификатор в своем определении. Эта возможность часто используется для определения рекурсивных функций. Как мы видели ранее, **let** не позволяет этого, поэтому необходимо использовать специальный синтаксис.

Синтаксис `let rec name = expr ;;`

Другой способ записи для функции с аргументами:

Синтаксис `let rec name p1 ... pn = expr ;;`

Определим функцию `sigma` вычисляющую сумму целых от 0 до значения указанного аргументом:

```
# let rec sigma x = if x = 0 then 0 else x + sigma (x-1) ;;
val sigma : int -> int = <fun>
# sigma 10 ;;
- : int = 55
```

Как заметил читатель — эта функция рискует «не закончиться» если входной аргумент меньше 0.

Обычно, рекурсивное значение это функция, компилятор не принимает некоторые рекурсивные объявления, значения которых не функциональные.

```
# let rec x = x + 1 ;;
```

This kind of expression is not allowed as right-hand side of ‘let rec’

Как мы увидим позднее, такие определения все таки возможны в некоторых случаях (см. 1.2.10 стр. 56).

Объявление **let rec** может быть скомбинировано с **and**. В этом случае функции определенные на одном и том же уровне, видны всем остальным. Это может быть полезно при декларации взаимно рекурсивных функций.

```
# let rec even n = (n<>1) && ((n=0) or (odd(n-1)))
  and   odd  n = (n<>0) && ((n=1) or (even(n-1))) ;;
val even : int -> bool = <fun>
val odd  : int -> bool = <fun>
# even 4 ;;
- : bool = true
# odd 5 ;;
- : bool = true
```

По тому же принципу, локальные функции могут быть рекурсивными. Новое определение функции **sigma** проверяет корректность входного аргумента, перед тем как посчитать сумму локальной функцией **sigma_rec**.

```
# let sigma x =
  let rec sigma_rec x = if x=0 then 0 else x + sigma_rec(x-1) in
    if (x<0) then "error: negative argument"
    else "sigma = " ^ (string_of_int (sigma_rec x)) ;;
val sigma : int -> string = <fun>
```

Примечание

Мы вынуждены были определить возвращаемый тип как **string**, поскольку необходимо чтобы он был один и тот же, независимо от входного аргумента, отрицательного или положительного. Какое значение должна вернуть **sigma** если аргумент больше нуля? Далее, мы увидим правильный способ решения этой проблемы (см. 1.3 стр. 57).

1.1.5 Полиморфизм и ограничение типа

Некоторые функции выполняют одни и те же инструкции независимо от типа аргументов. К примеру, для создание пары из двух значений нет смысла определять функции для каждого известного типа⁵. Другой пример, доступ к первому полю пары не зависит от того, какого типа это поле.

```
# let make_pair a b = (a,b) ;;
val make_pair : 'a -> 'b -> 'a * 'b = <fun>
# let p = make_pair "paper" 451 ;;
val p : string * int = "paper", 451
# let a = make_pair 'B' 65 ;;
val a : char * int = 'B', 65
# fst p ;;
- : string = "paper"
# fst a ;;
- : char = 'B'
```

Функция, для которой не нужно указывать тип входного аргумента или возвращаемого значения называется полиморфной. Синтезатор типов, включенный в компилятор Objective Caml находит наиболее общий тип для каждого выражения. В этом случае Objective Caml использует переменные, здесь они обозначены как 'a и 'b, для указания общих типов. Эти переменные конкретизируются типом аргумента в момент применения функции.

При помощи полиморфных функций, мы получаем возможность написания универсального кода для любого типа переменных, сохраняя при этом надежность статической типизации. Действительно, несмотря на то что `make_pair` полиморфная, значение созданное (`make_pair 'b' 65`) имеет строго определенный тип, который отличен от (`make_pair "pair" 451`). Проверка типов реализуется в момент компиляции, таким образом универсальность кода никак не сказывается на эффективности программы.

Пример функций и полиморфных значений

В следующем примере приведена полиморфная функция с входным параметром функционального типа.

```
# let app = function f -> function x -> f x ;;
val app : ('a -> 'b) -> 'a -> 'b = <fun>
```

⁵Число типов ограничивается лишь размером доступной памяти.

Мы можем применить ее к функции `odd`, которая была определена ранее.

```
# app odd 2;;
- : bool = false
```

Функция тождества (`id`) возвращает полученный аргумент.

```
# let id x = x ;;
val id : 'a -> 'a = <fun>
# app id 1 ;;
- : int = 1
```

Следующая функция, `compose` принимает на входе две функции и еще один аргумент и возвращает результат применения двух функций к этому аргументу.

```
# let compose f g x = f (g x) ;;
val compose : ('a -> 'b) -> ('c -> 'a) -> 'c -> 'b = <fun>
# let add1 x = x+1 and mul5 x = x*5 in compose mul5 add1 9 ;;
- : int = 50
```

Как мы видим, тип результата возвращаемого `g` должен быть таким же как и тип входного аргумента `f`.

Не только функциональные значения могут быть полиморфными, проиллюстрируем это на примере пустого списка.

```
# let l = [] ;;
val l : 'a list = []
```

Следующий пример иллюстрирует тот факт, что создание типов основывается на разрешении ограничений, накладываемых применением функций, а вовсе не на значениях, полученных в процессе выполнения.

```
# let q = List.tl [2] ;;
val q : int list = []
```

Здесь тип равен `List.tl 'a list -> 'a list`, то есть эта функция применяется к списку целых и возвращает список целых. Тот факт что в момент выполнения возвращен пустой список, не влияет на его тип.

Objective Caml создает параметризованные типы для каждой функции, которые не зависят от ее аргументов. Этот вид полиморфизма называется **параметрическим полиморфизмом**⁶.

⁶Некоторые встроенные функции не подчиняются этому правилу, особенно функция структурного равенства (`=`), которая является полиморфной (ее тип `'a -> 'a -> bool`), но она исследует структуру своих аргументов для проверки равенства.

Ограничение типа

Синтезатор типа Objective Caml образует самый общий тип и иногда бывает необходимо уточнить тип выражения.

Синтаксис ограничения типа следующий:

Синтаксис `(expr : t)`

В этом случае синтезатор типа воспользуется этим ограничением при конструкции типа выражения. Использование ограничения типа позволяет

- сделать видимым тип параметра функции
- запретить использование функции вне своей области применения
- уточнить тип выражения, это окажется необходимым в случае физически изменяемых значений (см. 2.1 стр. 71).

Рассмотрим использование ограничения типа.

```
# let add (x:int) (y:int) = x + y ;;
val add : int -> int -> int = <fun>
# let make_pair_int (x:int) (y:int) = x,y ;;
val make_pair_int : int -> int -> int * int = <fun>
# let compose_fn_int (f: int -> int) (g: int -> int) (x:int) =
  compose f g x ;;
val compose_fn_int : (int -> int) -> (int -> int) -> int -> int ←
  = <fun>
# let nil = ([]: string list) ;;
val nil : string list = []
# 'H'::nil ;;
This expression has type string list but is here used with type ←
  char list
```

Это ограничение полиморфизма позволяет лучше контролировать тип выражений, тем самым ограничивая тип определенный системой. В таких выражениях можно использовать любой определенный тип.

```
# let llnil = ([] : 'a list list) ;;
val llnil : 'a list list = []
# [1;2;3]::llnil ;;
- : int list list = [[1; 2; 3]]
```

`llint` является списком списков любого типа.

В данном случае подразумевается ограничение типа, а не явная типизация заменяющая тип, который определил Objective Caml. В частности, мы не можем обобщить тип, более чем это позволяет вывод типов Objective Caml.

```
# let add_general (x:'a) (y:'b) = add x y ;;
val add_general : int -> int -> int = <fun>
```

Ограничения типа будут использованы в интерфейсах модулей (см. ??), а так же в декларации классов (см. ??).

1.1.6 Примеры

В этом параграфе мы приведем несколько примеров функций. Большинство из них уже определены в Objective Caml, мы делаем это только в «педагогических» целях.

Тест остановки рекурсивных функций реализован при помощи проверки, имеющей стиль более близкий к Lisp. Мы увидим как это сделать в стиле ML (см. 1.2.1 стр. 34).

Размер списка

Начнем с функции проверяющей пустой список или нет.

```
# let null l = (l=[]);;
val null : 'a list -> bool = <fun>
```

Определим функцию `size` вычисления размера списка (т.е. число элементов).

```
# let rec size l =
  if null l then 0
  else 1+(size(List.tl l));;
val size : 'a list -> int = <fun>
# size [] ;;
- : int = 0
# size [1;2;18;22] ;;
- : int = 4
```

Функция `size` проверяет список: если он пуст возвращает 0, иначе прибавляет 1 к длине остатка списка.

Итерация композиций (Iteration of composition)

Выражение `iterate n f` вычисляет f^n , что соответствует применению функции `f` `n` раз.

```
# let rec iterate n f =
  if n = 0 then (function x -> x)
  else compose f (iterate (n-1) f) ;;
val iterate : int -> (('a -> 'a) * 'a -> 'a) -> 'a -> 'a = <fun>
```

Функция `iterate` проверяет аргумент `n` на равенство нулю, если аргумент равен, то возвращаем функцию идентичности, иначе возвращаем композицию `f` с итерацией `f n-1` раз.

Используя `iterate` можно определить операцию возведения в степень, как итерацию умножения.

```
# let rec power i n =
  let i_times = ( * ) i in
  iterate n i_times 1 ;;
val power : int -> int -> int = <fun>
# power 2 8 ;;
- : int = 256
```

Функция `power` повторяет `n` раз функциональное выражение `i_times`, затем применяет этот результат к `1`, таким образом мы получаем `n`-ю степень целого числа.

Таблица умножения

Напишем функцию `multab`, которая вычисляет ряд таблицы умножения соответствующую целому числу переданному в аргументе.

Для начала определим функцию `apply_fun_list`. Пусть `f_list` список функций, тогда вызов `apply_fun_list x f_list` возвращает список результатов применения каждого элемента списка `f_list` к `x`.

```
# let rec apply_fun_list x f_list =
  if null f_list then []
  else ((List.hd f_list) x)::(apply_fun_list x (List.tl f_list)) ;;
val apply_fun_list : 'a -> ('a -> 'b) list -> 'b list = <fun>
# apply_fun_list 1 [( + ) 1; ( + ) 2; ( + ) 3] ;;
- : int list = [2; 3; 4]
```


Функция `mk_mult_fun_list` возвращает список функций умножающих их аргумент на `i`, $0 \leq i \leq n$.

```
# let mk_mult_fun_list n =
  let rec mmfl_aux p =
    if p = n then [ ( * ) n ]
    else (( * ) p) :: (mmfl_aux (p+1))
  in (mmfl_aux 1) ;;
val mk_mult_fun_list : int -> (int -> int) list = <fun>
```

Подсчитаем ряд для 7

```
# let multab n = apply_fun_list n (mk_mult_fun_list 10) ;;
val multab : int -> int list = <fun>
# multab 7 ;;
- : int list = [7; 14; 21; 28; 35; 42; 49; 56; 63; 70]
```

Итерация в списке

Вызов функции `fold_left f a [e1; e2; ...; en]` возвращает `f ... (f (f a e1) e2) ... en`, значит получаем `n` применений.

```
# let rec fold_left f a l =
  if null l then a
  else fold_left f ( f a (List.hd l) ) (List.tl l) ;;
val fold_left : ('a -> 'b -> 'a) -> 'a -> 'b list -> 'a = <fun>
```

С помощью функции `fold_left` можно компактно определить функцию вычисления суммы элементов списка целых чисел.

```
# let sum_list = fold_left (+) 0 ;;
val sum_list : int list -> int = <fun>
# sum_list [2;4;7] ;;
- : int = 13
```

Или, например конкатенация элементов списка строк.

```
# let concat_list = fold_left (^) "" ;;
val concat_list : string list -> string = <fun>
# concat_list ["Hello "; "word"; "!"] ;;
- : string = "Hello word!"
```

1.2 Объявления типов и сопоставление с образцом

С помощью типов определенных в Objective Caml мы можем определять структурные типы, при помощи кортежей или списков. Но в некоторых случаях бывает необходимо определить новые типы для описания специальных структур. В Objective Caml, объявления типов рекурсивны и могут быть параметризованными переменными типа, как в случае `'a list` который мы уже обсуждали. Вывод типов принимает во внимание новые объявления для определения типов выражений. Конструкция значений новых типов использует конструктор описанный определением типа. Особой особенностью языков семейства ML является сопоставление с образцом, которое обеспечивает простой метод доступа к компонентам (полям) структур данных. Определение функции соответствует чаще всего сопоставлению с образцом по одному из ее параметров, что позволяет определять функции для различных случаев.

Для начала, продемонстрируем механизм сопоставления на существующих типах и затем опишем различные объявления структурных типов, конструкций таких переменных и доступ к компонентам по сопоставлению с образцом.

1.2.1 Сопоставление с образцом

Образец (по английски *pattern*), строго говоря не совсем выражение Objective Caml. Речь скорее идет о корректной компоновке (синтаксической и с точки зрения типов) элементов, таких как константы базовых типов (`int`, `bool`, `char`, ...), переменные, конструкторы и символ `'_'`, называемый универсальным образцом⁷. Другие символы служат для записи шаблона, которые мы опишем по ходу.

Сопоставление с образцом применяется к значениям, оно служит для распознавания формы этих значений и позволяет определять порядок вычислений. С каждым образцом связано выражение для вычисления.

Синтаксис	match <i>expr</i> with
	<i>p</i> ₁ -> <i>expr</i> ₁
	:
	<i>p</i> _{<i>n</i>} -> <i>expr</i> _{<i>n</i>}

⁷Wildcard pattern

1.2. ОБЪЯВЛЕНИЯ ТИПОВ И СОПОСТАВЛЕНИЕ С ОБРАЗЦОМ 35

Выражение *expr* последовательно сопоставляется (фильтруется) с разными образцами p_1, \dots, p_n . Если один из образцов (например p_i) соответствует значению *expr*, то соответствующее ответвление будет вычислено (*expr_i*). Образцы p_i одинакового типа, так же как и *expr_i*. Вертикальная черта перед первым образцом не обязательна.

Примеры

Приведем два способа, при помощи сопоставления с образцом, определения функции `imply` типа `(bool * bool) -> bool`, реализующую логическую импликацию. Образец для пар — (,).

Первая версия перечисляет все возможности, как таблица истинности.

```
# let imply v = match v with
  (true ,true)   -> true
  |(true, false) -> false
  |(false, true)  -> true
  |(false, false) -> true ;;
val imply : bool * bool -> bool = <fun>
```

Используя переменные группирующие несколько случаев, мы получаем более компактное определение.

```
# let imply v = match v with
  (true, x)   -> x
  |(false, x) -> true ;;
val imply : bool * bool -> bool = <fun>
```

Обе версии `imply` выполняют одну и ту же функцию, то есть они возвращают одинаковые значения, для одинаковых входных аргументов.

Линейный образец

Образец обязательно должен быть линейным, то есть определенная переменная не может быть использована дважды. Мы могли бы написать:

```
# let equal c = match c with
  (x, x) -> true
  |(x, y) -> false ;;
This variable is bound several times in this matching
```

Но для этого компилятор должен уметь реализовывать тесты на равенство, что приводит к множеству проблем. Если мы ограничимся физическим значением переменных, то мы получим слишком «слабую» систему, не в состоянии, например, распознать равенство между двумя списками [1; 2]. Если же мы будем проверять на структурное равенство, то рискуем бесконечно проверять циклические структуры. Рекурсивные функции, например, это циклические структуры, но мы так же можем определить рекурсивные значения не являющиеся функциями (см. 1.2.10 стр. 56).

Универсальный образец

Символ ‘_’ совпадает со всеми возможными значениями — он называется универсальным. Этот образец может быть использован для сопоставления сложных типов. Воспользуемся им для определения еще одной версии функции `imply`:

```
# let imply v = match v with
  (true, false) -> false
  | _           -> true ;;
val imply : bool * bool -> bool = <fun>
```

Сопоставление должно обрабатывать все случаи, иначе компилятор выводит сообщение:

```
# let in_zero n = match n with 0 -> true ;;
Warning: this pattern-matching is not exhaustive.
Here is an example of a value that is not matched:
1
val in_zero : int -> bool = <fun>
```

В случае если аргумент не равен 0, функция не знает какое значение должно быть возвращено. Для исправления добавим универсальный образец:

```
# let is_zero n = match n with
  0 -> true
  | _ -> false ;;
val is_zero : int -> bool = <fun>
```

Если во время выполнения ни один образец не выбран, возбуждается исключение:

1.2. ОБЪЯВЛЕНИЯ ТИПОВ И СОПОСТАВЛЕНИЕ С ОБРАЗЦОМ 37

```
# let f x = match x with 1 -> 3 ;;
Warning: this pattern-matching is not exhaustive.
Here is an example of a value that is not matched:
0
val f : int -> int = <fun>
# f 1 ;;
- : int = 3
# f 4 ;;
Exception: Match_failure ("", 46, -32).
```

Исключение `Match_Failure` возбуждено при вызове `f 4` и если оно не обработано, то вычисление останавливается (см. 1.3 стр. 57).

Комбинация образцов

Комбинация образцов позволяет получить новый образец, который может разложить значение в соответствии с одним или другим из начальных шаблонов.

Синтаксис $p_1 \mid \dots \mid p_n$

Эта форма создает новый образец из комбинации мотивов p_1, \dots, p_n , с единственным ограничением — каждый из этих образцов должен содержать константные значения либо универсальный образец.

Следующий пример показывает как проверить является ли входной символ гласной буквой.

```
# let is_a_vowel c = match c with
    'a' | 'e' | 'i' | 'o' | 'u' | 'y' -> true
    | _ -> false ;;
val is_a_vowel : char -> bool = <fun>
# is_a_vowel 'i' ;;
- : bool = true
# is_a_vowel 'j' ;;
- : bool = false
```

Параметризованное сопоставление

Параметризованное сопоставление используется в основном для определения функций выбора. Для облегчения записи подобных определений, конструктор **function** разрешает следующее сопоставление одного параметра.

Синтаксис	function		$p_1 \rightarrow expr_1$
			$p_2 \rightarrow expr_2$
			\vdots
			$p_n \rightarrow expr_n$

Вертикальная черта перед первым образцом не обязательна. Каждый раз при определении функции, мы используем сопоставление с образцом. Действительно, конструкция **function** $x \rightarrow expression$ является определением сопоставления по уникальному образцу одной переменной. Можно использовать эту особенность в простых шаблонах:

```
# let f = function (x, y) -> 2*x + 3*y + 4 ;;
val f : int * int -> int = <fun>
```

Форма

$$\mathbf{function} \ p_1 \rightarrow expr_1 \mid \dots \mid p_n \rightarrow expr_n$$

эквивалентна

$$\mathbf{function} \ expr \rightarrow \mathbf{match} \ expr \ \mathbf{with} \ p_1 \rightarrow expr_1 \mid \dots \mid p_n \rightarrow expr_n$$

Используя схожесть определения на (см. 1.1.4 стр. 23), мы бы написали:

```
# let f (x, y) = 2*x + 3*y + 4 ;;
val f : int * int -> int = <fun>
```

Но подобная запись возможна лишь в случае если фильтруемое значение принадлежит к типу с единственным конструктором, иначе сопоставление не является исчерпывающим:

```
# let is_zero 0 = true ;;
Warning: this pattern-matching is not exhaustive.
Here is an example of a value that is not matched:
1
val is_zero : int -> bool = <fun>
```

Присвоение имени фильтруемому значению

Иногда при сопоставлении с образцом, бывает необходимо дать имя всему образцу или лишь его части. Следующая синтаксическая форма вводит ключевое слово **as**, которое ассоциирует имя образцу.

Синтаксис `(p as name)`

Это бывает полезно когда необходимо разбить значение на части, сохраняя при этом его целостность. В следующем примере функция возвращает наименьшее рациональное число из пары. Каждое рациональное число представлено числителем и знаменателем.

```
# let min_rat pr = match pr with
  ((_, 0), p2) -> p2
| (p1, (_, 0)) -> p1
| (((n1, d1) as r1), ((n2, d2) as r2)) ->
  if (n1 * d2) < (n2 * d1) then r1 else r2 ;;
val min_rat : (int * int) * (int * int) -> int * int = <fun>
```

Для сравнения двух рациональных чисел, необходимо разбить их для именованного числителя и знаменателя (`n1`, `n2`, `d1` и `d2`), но так же для собирания в одно целое начальные пары (`r1` или `r2`). Конструкция позволяет дать имена частям значения — это избавляет от реконструкции рационального числа при возвращении результата.

Сопоставление с ограничением

Сопоставление с ограничением соответствует вычислению условного выражения сразу после сопоставления с образцом. Если это выражение возвращает значение **true** то выражение соответствующее этому образцу будет вычислено, иначе сопоставление будет продолжаться дальше.

Синтаксис `match expr with`
`:`
`| pi when condi -> expri`
`:`

В следующем примере используется два ограничения для проверки равенства двух рациональных чисел.

```
# let eq_rat cr = match cr with
  ((_, 0), (_, 0)) -> true
  | ((_, 0), _) -> false
  | (_, (_, 0)) -> false
  | ((n1, 1), (n2, 1)) when n1 = n2 -> true
  | ((n1, d1), (n2, d2)) when ((n1 * d2) = (n2 * d1)) -> true
  | _ -> false ;;
val eq_rat : (int * int) * (int * int) -> bool = <fun>
```

Если при фильтрации четвертого образца ограничение не выполнится, то сопоставление продолжится по пятому образцу.

Внимание

При проверке исчерпываемости сопоставления Objective Caml предполагает что условное выражение может быть ложным. В следствии, верификатор не учитывает этот образец, так как не возможно знать до выполнения сработает ограничение или нет. Исчерпываемость следующего фильтра не может быть определена.

```
# let f = function x when x = x -> true ;;
Warning: Bad style, all clauses in this pattern-matching are guarded.
val f : 'a -> bool = <fun>
```

Образец интервала символов

При сопоставлении символов, неудобно описывать комбинацию всех образцов, которая соответствует интервалу символов. Действительно, для того чтобы проверить является ли символ буквой, необходимо написать как минимум 26 образцов и скомбинировать их. Для символьного типа разрешается запись следующего шаблона:

Синтаксис `'c1' .. 'cn'`

что соответствует `'c1' | 'c2' | ... | 'cn'`

К примеру образец `'0' .. '9'` соответствует образу `'0' | '1' | '2' | '3' | '4' | '5' | '6' | '7' | '8' | '9'`. Первая форма быстрее и проще воспринимается при чтении.

1.2. ОБЪЯВЛЕНИЯ ТИПОВ И СОПОСТАВЛЕНИЕ С ОБРАЗЦОМ 41

Внимание

Это особенность является расширением языка и может измениться в следующих версиях.

Используя комбинированные образцы и интервалы, напомним функцию по определению символа по нескольким критериям.

```
# let char_discriminate c = match c with
  'a' | 'e' | 'i' | 'o' | 'u' | 'y'
  | 'A' | 'E' | 'I' | 'O' | 'U' | 'Y' -> "Vowel"
  | 'a'..'z' | 'A'..'Z' -> "Consonant"
  | '0'..'9' -> "Digit"
  | _ -> "Other" ;;
val char_discriminate : char -> string = <fun>
```

Заметим, что порядок образцов важен, второе множество содержит первое, но оно проверяется только после неудачи первого теста.

Образцы списков

Как мы уже видели список может быть:

- либо пуст (в форме `[]`)
- либо состоять из первого элемента (заголовок) и под-списка (хвост).
В этом случае он имеет форму `t::q`.

Обе эти записи могут быть использованы в образце при фильтрации списка.

```
# let rec size x = match x with
  [] -> 0
  | _::tail_x -> 1 + (size tail_x) ;;
val size : 'a list -> int = <fun>
# size [] ;;
- : int = 0
# size [7; 9; 2; 6] ;;
- : int = 4
```

Перепишем ранее показанный пример (см. [1.1.6 стр. 31](#)) используя сопоставление с образцом для итерации списков.

```
# let rec fold_left f a = function
  [] -> a
  | head::tail -> fold_left f (f a head) tail ;;
val fold_left : ('a -> 'b -> 'a) -> 'a -> 'b list -> 'a = <fun>
# fold_left (+) 0 [8; 4; 10] ;;
- : int = 22
```

Объявление значения через сопоставлением с образцом

Объявление значений само по себе использует сопоставление. `let x = 18` сопоставляет образцы `x` со значением 18. Любой образец принимается как фильтр объявления, переменные образца ассоциируются со значениями которые они сопоставляют.

```
# let (a, b, c) = (1, true, 'A') ;;
val a : int = 1
val b : bool = true
val c : char = 'A'
# let (d, c) = 8, 3 in d + c ;;
- : int = 11
```

Видимость переменных фильтра та же самая что у локальных переменных. Здесь с ассоциировано с `'A'`.

```
# a + (int_of_char c) ;;
- : int = 66
```

Как и любой фильтр, определение значения может быть не исчерпывающим.

```
# let [x; y; z] = [1; 2; 3] ;;
Warning: this pattern-matching is not exhaustive.
Here is an example of a value that is not matched:
[]
val x : int = 1
val y : int = 2
val z : int = 3
# let [x; y; z] = [1; 2; 3; 4] ;;
Warning: this pattern-matching is not exhaustive.
Here is an example of a value that is not matched:
[]
Exception: Match_failure ("", 72, -41).
```

1.2. ОБЪЯВЛЕНИЯ ТИПОВ И СОПОСТАВЛЕНИЕ С ОБРАЗЦОМ⁴³

Принимается любой образец, включая конструктор, универсальный и комбинированный.

```
# let head :: 2 :: _ = [1; 2; 3] ;;
Warning: this pattern-matching is not exhaustive.
Here is an example of a value that is not matched:
[]
val head : int = 1
# let _ = 3. +. 0.14 in "PI" ;;
- : string = "PI"
```

Последний пример мало интересен для функционального программирования, поскольку значение `3.14` не именовано, и соответственно будет потерянно.

1.2.2 Декларация типов

Объявление типа — это одна из других возможных фраз синтаксиса Objective Caml. Она позволяет определить новые типы соответствующие обычным структурам данных используемым в программах. Существует два больших семейства типов: тип-произведение для кортежей или записей или тип-сумма для `union`.

Синтаксис `type name = typedef ;;`

В отличии от объявления переменных, объявление типов по умолчанию рекурсивно. То есть объявления типов, когда они скомбинированы, могут объявлять взаимно рекурсивные типы.

Синтаксис

```
type  name1 = typedef1
and   name2 = typedef2
      ⋮
and   namen = typedefn ;;
```

Объявления типов могут быть параметризованы переменной типа. Имя переменной типа начинается всегда с апострофа (символ `'`).

Синтаксис `type 'a name = typedef ;;`

В случае когда таких переменных несколько, параметры типа декларируются как кортеж перед именем типа.

Синтаксис `type ('a1 ... 'an) name = typedef ;;`

Только те параметры, которые определены в левой части декларации, могут появиться в ее правой части.

Примечание

Во время вывода на экран Objective Caml переименует параметры типа, первый в 'a, второй в 'b и так далее.

Мы всегда можем объявить новый тип используя ранее объявленные типы.

Синтаксис `type name = typeexpression`

Это полезно в том случае, когда мы хотим ограничить слишком общий тип:

```
# type 'param paired_with_integer = int * 'param ;;
type 'a paired_with_integer = int * 'a
# type specific_pair = float paired_with_integer ;;
type specific_pair = float paired_with_integer
```

Без ограничения, выводится наиболее общий тип:

```
# let x = (3, 3.14) ;;
val x : int * float = (3, 3.14)
```

Но мы можем использовать ограничитель типа для того, чтобы получить желаемое имя:

```
# let (x: specific_pair) = (3, 3.14) ;;
val x : specific_pair = (3, 3.14)
```

1.2.3 Записи

Запись — это кортеж, в котором каждому полю присваивается имя на подобии `record` в Pascal или `struct` в C. Запись всегда соответствует объявлению нового типа. Для определения записи необходимо указать ее имя, а так же имя и тип для каждого поля записи.

Синтаксис `type name = { name1: t1; ... ; namen: tn } ;;`

Определим тип комплексного числа следующим образом.

```
# type complex = { re: float; im: float } ;;
type complex = { re : float; im : float; }
```

Для того чтобы создать значение типа запись, нужно присвоить каждому полю значение (в каком угодно порядке).

Синтаксис `{ namei1 = expri1; ... ; namein = exprin } ;;`

В следующем примере мы создадим комплексное число, в котором действительная часть равна 2 и мнимая 3:

```
# let c = { re = 2.; im = 3. } ;;
val c : complex = {re = 2.; im = 3.}
# c = { im = 3.; re = 2. } ;;
- : bool = true
```

В случае если не хватает некоторых полей, произойдет следующая ошибка:

```
# let d = { im = 4. } ;;
Some record field labels are undefined: re
```

Доступ к полям возможен двумя способами: синтаксис с точкой, либо сопоставлением некоторых полей.

Синтаксис с точкой заключается в следующем:

Синтаксис `expr.name`

Выражение `expr` должно иметь тип «запись с полем `name`».

Сопоставление записи позволяет получить значение связанное с определенным полем. Синтаксис образца для записи следующий.

Синтаксис `{ namei = pi ; ... ; namej = pj }`

Образцы находятся справа от символа '=' (`pi, ..., pj`). Необязательно перечислять все поля записи в образце.

Функция `add_complex` обращается к полям с помощью «точки», тогда как функция `mult_complex` обращается при помощи фильтра.

```
# let add_complex c1 c2 = {re = c1.re +. c2.re; im = c1.im +.  ◀
    c2.im} ;;
val add_complex : complex -> complex -> complex = <fun>
# add_complex c c ;;
- : complex = {re = 4.; im = 6.}
# let mult_complex c1 c2 = match (c1, c2) with
    ({re=x1; im=y1}, {re=x2; im=y2}) -> {re = x1 *. x2 -. y1 *.  ◀
        y2; im = x1 *. y2 +. x2 *. y1} ;;
val mult_complex : complex -> complex -> complex = <fun>
# mult_complex c c ;;
- : complex = {re = -5.; im = 12.}
```

Преимущество записей, по сравнению с кортежами как минимум двойное:

- более информативное описание, благодаря именам полей — это в частности позволяет облегчить образцы фильтра;
- идентичное обращение по имени, порядок значения не имеет — главное имя.

```
# let a = (1, 2, 3) ;;
val a : int * int * int = (1, 2, 3)
# let f tr = match tr with x, _, _ -> x ;;
val f : 'a * 'b * 'c -> 'a = <fun>
# f a ;;
- : int = 1
# type triplet = { x1: int; x2: int; x3: int } ;;
type triplet = { x1 : int; x2 : int; x3 : int; }
# let b = { x1=1; x2=2; x3=3 } ;;
val b : triplet = {x1 = 1; x2 = 2; x3 = 3}
# let g tr = tr.x1 ;;
val g : triplet -> int = <fun>
# g b ;;
- : int = 1
```

При сопоставлении с образцом не обязательно перечислять все поля записи, тип будет вычислен по последней описанной записи, которая содержит указанные поля.

```
# let h tr = match tr with { x1 = x } -> x ;;
val h : triplet -> int = <fun>
# h b ;;
- : int = 1
```

1.2. ОБЪЯВЛЕНИЯ ТИПОВ И СОПОСТАВЛЕНИЕ С ОБРАЗЦОМ⁴⁷

Существует конструкция позволяющая создать идентичную запись, с разницей в несколько полей, что удобно для записей с большим числом полей.

Синтаксис $\{name \textbf{ with } name_i = expr_i ; \dots ; name_j = expr_j\}$

```
# let c = { b with x1=0 } ;;  
val c : triplet = {x1 = 0; x2 = 2; x3 = 3}
```

Новая копия значения **b** отличается только лишь значением поля **x1**.

Внимание

Это особенность является расширением языка и может измениться в следующих версиях.

1.2.4 Тип сумма

В отличие от кортежей или записей, которые соответствуют декартовому произведению, тип сумма⁸ соответствует объединению множеств (**union**). В один тип мы группируем несколько разных типов (например, целые числа и строки). Различные члены суммы определяются конструкторами, которые с одной стороны позволяют сконструировать значение этого типа и с другой получить доступ к компонентам этих значений при помощи сопоставления с образцом. Применить конструктор к аргументу, значит указать что возвращаемое значение будет этого нового типа.

Тип сумма описывается указыванием имени конструкторов и типа их возможного аргумента.

Синтаксис

```
type name = ...  
| Namei ...  
| Namej of tj ...  
| Namek of tk * ... * tl ... ;;
```

Имя конструктора это специальный идентификатор.

Внимание

Имя конструктора должна всегда начинаться с заглавной буквы.

⁸Sum type

Константный конструктор

Мы называем константным, конструктор без аргументов. Такие конструкторы могут затем использоваться как значения языка, в качестве констант.

(Орел Решка)

```
# type coin = Heads | Tails ;;
type coin = Heads | Tails
# Tails ;;
- : coin = Tails
```

Тип `bool` может быть определен подобным образом.

Конструктор с аргументами

Конструктора могут иметь аргументы, ключевое слово **of** указывает тип аргумента. Это позволяет сгруппировать под одним типом объекты разных типов, имеющих разные конструкторы.

Этот классический пример определяет тип для карты в игре Tarot⁹. Определим типы `suit` и `card` следующим образом.

- тип `suit` определяет масть карты — пика, черва, бубна и трефа
- тип `card` определяет достоинство карты (король, дама...)

```
# type suit = Spades | Hearts | Diamonds | Clubs ;;

# type card =
    King of suit
  | Queen of suit
  | Knight of suit
  | Knave of suit
  | Minor_card of suit * int
  | Trump of int
  | Joker ;;
```

Создание значения типа `card` получается применением конструктора к значению с нужным типом.

⁹Правила игры Tarot можно найти например на <http://www.pagat.com/tarot/frtarot.html>.

1.2. ОБЪЯВЛЕНИЯ ТИПОВ И СОПОСТАВЛЕНИЕ С ОБРАЗЦОМ 49

```
# King Spades ;;
- : card = King Spades
# Minor_card(Hearts, 10) ;;
- : card = Minor_card (Hearts, 10)
# Trump 21 ;;
- : card = Trump 21
```

Следующая функция, `all_cards` создает список всех карт цвета указанного аргументом.

```
# let rec interval a b = if a = b then [b] else a:: (interval (a+1) b) ;;
val interval : int -> int -> int list = <fun>
# let all_cards s =
    let face_cards = [ Knave s; Knight s; Queen s; King s ]
    and other_cards = List.map (function n -> Minor_card(s, n)) (interval 1 10)
    in face_cards @ other_cards ;;
val all_cards : suit -> card list = <fun>
# all_cards Hearts ;;
- : card list =
[Knave Hearts; Knight Hearts; Queen Hearts; King Hearts;
 Minor_card (Hearts, 1); Minor_card (Hearts, 2); Minor_card (Hearts, 3);
 Minor_card (Hearts, 4); Minor_card (Hearts, 5); Minor_card (Hearts, 6);
 Minor_card (Hearts, 7); Minor_card (Hearts, 8); Minor_card (Hearts, 9);
 Minor_card (Hearts, 10)]
```

Для манипуляции значений типа `сумма`, мы используем сопоставление с образцом. В следующем примере опишем функцию преобразующую значения типа `suit` и типа `card` в строку (тип `string`):

```
# let string_of_suit = function
    Spades -> "spades"
  | Diamonds -> "diamonds"
  | Hearts -> "hearts"
  | Clubs -> "clubs" ;;
val string_of_suit : suit -> string = <fun>
# let string_of_card = function
```

```

    King c          -> "king of " ^ (string_of_suit c)
  | Queen c         -> "queen of " ^ (string_of_suit c)
  | Knave c          -> "knave of " ^ (string_of_suit c)
  | Knight c         -> "knight of " ^ (string_of_suit c)
  | Minor_card (c, n) -> (string_of_int n) ^ " of " ^ (
(string_of_suit c)
  | Trump n          -> (string_of_int n) ^ " of trumps"
  | Joker            -> "joker" ;;
val string_of_card : card -> string = <fun>

```

Конструктор `Minor_card` бинарный, для сопоставления такого значения мы должны дать имя обоим компонентам.

```

# let is_minor_card c = match c with
    Minor_card v -> true
  | _             -> false ;;
The constructor Minor_card expects 2 argument(s),
but is here applied to 1 argument(s)

```

Чтобы не давать имя каждой компоненте конструктора, объявим его с одним аргументом, заключив в скобки тип ассоциированного кортежа. Сопоставление двух следующих конструкторов отличается.

```

# type t =
    C of int * bool
  | D of (int * bool) ;;
type t = C of int * bool | D of (int * bool)
# let access v = match v with
    C (i, b) -> i, b
  | D x      -> x ;;
val access : t -> int * bool = <fun>

```

1.2.5 Рекурсивный тип

Определение рекурсивного типа необходимо в любом языке программирования, с его помощью мы можем определить такие типы как список, стек, дерево и так далее. В отличие от объявления **let**, **type** всегда рекурсивный в Objective Caml.

Типу список, определенному в Objective Caml, необходим один единственный аргумент. Для того чтобы хранить значения двух разных типов, как например `int` или `char` мы напомним:

1.2. ОБЪЯВЛЕНИЯ ТИПОВ И СОПОСТАВЛЕНИЕ С ОБРАЗЦОМ 51

```
# type int_of_char_list =
  Nil
  | Int_cons of int * int_of_char_list
  | Char_cons of char * int_of_char_list ;;

# let l1 = Char_cons ('=', Int_cons(5, Nil)) in
  Int_cons (2, Char_cons('+', Int_cons(3, l1))) ;;
- : int_of_char_list =
Int_cons (2,
  Char_cons ('+', Int_cons (3, Char_cons ('=', Int_cons (5,
    Nil))))))
```

1.2.6 Параметризованный тип

Мы также можем определить тип с параметрами, что позволит нам обобщить предыдущий список для двух любых типов.

```
# type ('a, 'b) list2 =
  Nil
  | Acons of 'a * ('a, 'b) list2
  | Bcons of 'b * ('a, 'b) list2 ;;

# Acons(2, Bcons('+', Acons(3, Bcons('=', Acons(5, Nil))))) ;;
- : (int, char) list2 =
Acons (2, Bcons ('+', Acons (3, Bcons ('=', Acons (5, Nil)))))
```

Естественно, оба параметра 'a и 'b могут быть одного типа:

```
# Acons(1, Bcons(2, Acons(3, Bcons(4, Nil)))) ;;
- : (int, int) list2 = Acons (1, Bcons (2, Acons (3, Bcons (4,
  Nil))))
```

Как в предыдущем примере, мы можем использовать тип list2 чтобы пометить четные и нечетные числа. Что бы создать обычный список, достаточно извлечь подсписок четных чисел.

```
# let rec extract_odd = function
  Nil -> []
  | Acons(_, x) -> extract_odd x
  | Bcons(n, x) -> n::(extract_odd x) ;;
val extract_odd : ('a, 'b) list2 -> 'b list = <fun>
```

Определение этой функции ничего не говорит о свойстве значений хранящихся в структуре, по этой причине ее тип параметризован.

1.2.7 Видимость описания

На имена конструкторов распространяются те же правила, что и на глобальные объявления. Переопределение скрывает своего предшественника. Скрытые значения всегда существуют, они никуда не делись. Однако цикл не сможет различить эти оба типа, поэтому мы получим не совсем понятное сообщение об ошибке.

В следующем примере, константный конструктор `Nil` типа `int_of_char` скрывается объявлением конструктора `('a, 'b) list2`.

```
# Int_cons(0, Nil) ;;
This expression has type ('a, 'b) list2 but is here used with  ←
type int_of_char_list
```

Во втором примере мы получим совсем бестолковое сообщение, по крайней мере на первый взгляд. Пусть мы имеем следующую программу:

```
# type t1 = Empty | Full ;;
type t1 = Empty | Full
# let empty_t1 x = match x with Empty -> true | Full -> false ;;
val empty_t1 : t1 -> bool = <fun>
# empty_t1 Empty ;;
- : bool = true
```

Затем переопределим тип `t1`:

```
# type t1 = u : int; v : int ;;
type t1 = u : int; v : int;
# let y = u=2; v=3 ;;
val y : t1 = u = 2; v = 3
```

Теперь если мы применим функцию `empty_t1` к значению с новым типом `t1`, то получим следующее сообщение.

```
# empty_t1 y ;;
This expression has type t1 but is here used with type t1
```

Первое упоминание типа `t1` соответствует ранее определенному типу, тогда как второе — последнему.

1.2.8 Функциональные типы

Тип аргумента конструктора может быть каким угодно, и в частности он может быть функциональным. Следующий тип создает список состоящий из функциональных значений, кроме последнего элемента.

```
# type 'a listf =
  Val of 'a
  | Fun of ('a -> 'a) * 'a listf ;;
type 'a listf = Val of 'a | Fun of ('a -> 'a) * 'a listf
```

Поскольку функциональные значения входят во множество значений которые обрабатываются языком, то мы можем создать значения типа `listf`:

```
# let eight_div = (/) 8 ;;
val eight_div : int -> int = <fun>
# let gl = Fun (succ, (Fun (eight_div, Val 4))) ;;
val gl : int listf = Fun (<fun>, Fun (<fun>, Val 4))
```

функция сопоставляющая подобные значения:

```
# let rec compute = function
  Val v -> v
  | Fun(f, x) -> f(compute x) ;;
val compute : 'a listf -> 'a = <fun>
# compute gl ;;
- : int = 3
```

1.2.9 Пример: реализация деревьев

Деревья — одна из часто встречающихся структур в программировании. Рекурсивные типы позволяют с легкостью определять подобные вещи. В этой части мы приведем два примера с такими структурами.

Бинарные деревья Определим дерево, в котором узлы обозначены значениями одного и того же типа:

```
# type 'a bin_tree =
  Empty
  | Node of 'a bin_tree * 'a * 'a bin_tree ;;
type 'a bin_tree = Empty | Node of 'a bin_tree * 'a * 'a bin_tree
```

Этой структурой мы воспользуемся в программе сортировки бинарных деревьев. Бинарное дерево имеет следующее свойство: значение левого дочернего узла меньше корневого и всех значений правых дочерних узлов. На рисунке 1.1 изображен пример бинарного дерева для целых чисел. Пустой узел (конструктор `Empty`) обозначен прямоугольником, а остальные узлы (конструктор `Node`) — кружком с хранимым значением.

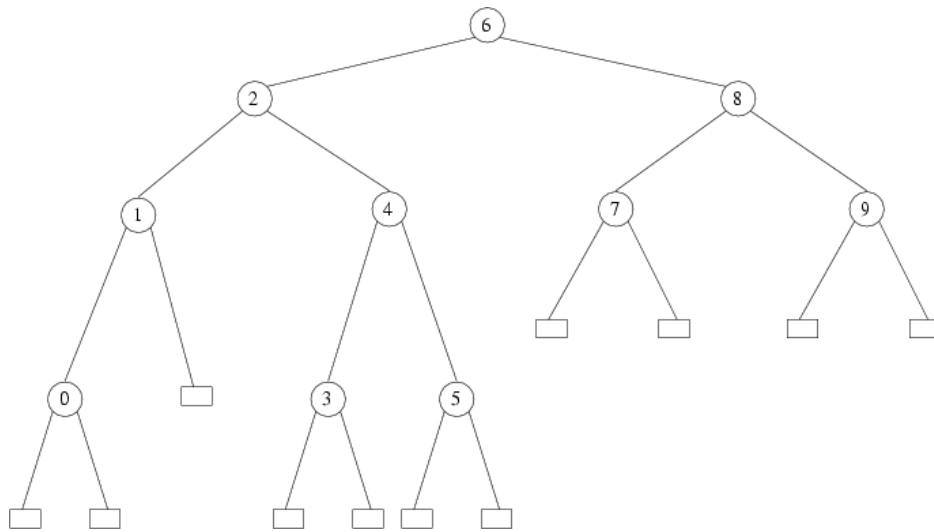


Рис. 1.1: Бинарное дерево поиска

Извлечем значения узлов в виде отсортированного списка при помощи инфиксного просмотра следующей функцией:

```
# let rec list_of_tree = function
  Empty -> []
  | Node(lb, r, rb) -> (list_of_tree lb) @ (r :: (list_of_tree rb))
val list_of_tree : 'a bin_tree -> 'a list = <fun>
```

Для того чтобы получить из списка бинарное дерево, определим функцию:

```
# let rec insert x = function
  Empty -> Node(Empty, x, Empty)
  | Node(lb, r, rb) -> if x < r then Node(insert x lb, r, rb)
                        else Node(lb, r, insert x rb)
val insert : 'a -> 'a bin_tree -> 'a bin_tree = <fun>
```

1.2. ОБЪЯВЛЕНИЯ ТИПОВ И СОПОСТАВЛЕНИЕ С ОБРАЗЦОМ 55

Функция трансформирующая список в бинарное дерево может быть получена использованием функции `insert`

```
# let rec tree_of_list = function
  [] -> Empty
  | h::t -> insert h (tree_of_list t) ;;
val tree_of_list : 'a list -> 'a bin_tree = <fun>
```

Тогда функция сортировки это всего-навсего композиция функций `tree_of_list` и `list_of_tree`.

```
# let sort x = list_of_tree (tree_of_list x) ;;
val sort : 'a list -> 'a list = <fun>
# sort [5; 8; 2; 7; 1; 0; 3; 6; 9; 4] ;;
- : int list = [0; 1; 2; 3; 4; 5; 6; 7; 8; 9]
```

Общие планарные деревья Воспользуемся функцией определенной в модуле `List` (см. ?? стр. ??):

- `List.map`: которая применяет одну функцию ко всем элементам списка и возвращает список результатов.
- `List.fold_left`: эквивалент функции `fold_left` определенной на стр. 33
- `List.exists` применяет функцию возвращающую булево значение ко всем элементам и если одно из применений возвращает `true`, то результат `true`, иначе `false`

Общее планарное дерево — это дерево в котором нет ограничение на число дочерних узлов: каждому узлу ассоциируем список дочерних узлов и длина этого списка не ограничена.

```
# type 'a tree = Empty
  | Node of 'a * 'a tree list ;;
type 'a tree = Empty | Node of 'a * 'a tree list
```

Пустое дерево представлено значением `Empty`, лист — это узел без дочерних узлов формы `Node(x, [])` или `Node(x, [Empty; Empty; ...])`. Таким образом достаточно просто написать функции, манипулирующие подобными деревьями такие как принадлежность элемента дереву или вычисление высоты дерева.

Для проверки принадлежности элемента `e` воспользуемся следующим алгоритмом: если дерево пустое, тогда элемент `e` не принадлежит этому дереву, в противном случае `e` принадлежит дереву только если он равен значению корня дерева или принадлежит дочерним узлам.

```
# let rec belongs e = function
  Empty -> false
  | Node(v, bs) -> (e=v) or (List.exists (belongs e) bs) ;;
val belongs : 'a -> 'a tree -> bool = <fun>
```

Для вычисления высоты дерева, напомним следующую функцию: высота пустого дерева равна 0, в противном случае его высота равна высоте самого большого поддерева плюс 1.

```
# let rec height =
  let max_list l = List.fold_left max 0 l in
  function
    Empty -> 0
    | Node (_, bs) -> 1 + (max_list (List.map height bs)) ;;
val height : 'a tree -> int = <fun>
```

1.2.10 Не функциональные рекурсивные значения

Рекурсивное объявление не функциональных значений позволяет конструировать циклические структуры данных.

Следующее объявление создает циклический список одного элемента.

```
# let rec l = 1 :: l ;;
val l : int list =
  [1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; ...]
```

Применение рекурсивной функции к подобному списку приведет к переполнению памяти.

```
# size l ;;
Stack overflow during evaluation (looping recursion?).
```

Структурное равенство таких списков может быть проверенно лишь в случае, когда физическое равенство верно.

```
# l = l ;;
- : bool = true
```

В случае, если мы определим такой же новый список, не стоит использовать проверку на структурное равенство, под угрозой заикливания программы. Таким образом следующая инструкция не рекомендуется.

```
let rec l2 = 1::l2 in l=l2 ;;
```


1.3. ТИПИЗАЦИЯ, ОБЛАСТЬ ОПРЕДЕЛЕНИЯ И ИСКЛЮЧЕНИЯ 57

Однако, физическое равенство остается возможным.

```
# let rec l2 = 1::l2 in l==l2 ;;
- : bool = false
```

Предикат `==` сравнивает непосредственно значение или разделения структурного объекта (одним словом равенство по адресу). Мы воспользуемся этим тестом для того чтобы при просмотре списка, не проверять уже просмотренные подсписки. Сначала, определим функцию `memq` которая проверяет присутствие элемента в списке при помощи физического равенства. В это же время функция `mem` проверяет равенство структурное. Обе функции принадлежат модулю `List`.

```
# let rec memq a l = match l with
  [] -> false
  | b :: l -> (a==b) or (memq a l) ;;
val memq : 'a -> 'a list -> bool = <fun>
```

Функция вычисляющая размера списка определена при помощи списка уже проверенных списков и останавливается при встрече списка второй раз.

```
# let special_size l =
  let rec size_aux previous l = match l with
    [] -> 0
    | _::l1 -> if memq l previous then 0
                else 1 + (size_aux (l::previous) l1)
  in size_aux [] l ;;
val special_size : 'a list -> int = <fun>
# special_size [1; 2; 3; 4] ;;
- : int = 4
# special_size l ;;
- : int = 1
# let rec l1 = 1 :: 2 :: l2 and l2 = 1 :: 2 :: l1 in
  special_size l1 ;;
- : int = 4
```

1.3 Типизация, область определения и исключения

Вычисленный тип функции принадлежит подмножеству множества в котором она определена. Если входной аргумент функции типа `int`,

это не значит что она сможет быть выполнена для любого переданного ей целого числа. Таких случаях мы используем механизм исключений Objective Caml. Запуск исключения провоцирует остановку вычислений, которое может быть выявлено и обработано. Для этого необходимо установить обработчик исключения, до того как исключение может быть возбуждено.

1.3.1 Частичные функции и исключения

Область определения функции соответствует множеству значений которыми функция манипулирует. Существует множество математических частичных функций, таких как, например, деление или натуральный логарифм. Проблемы возникают в том случае когда функция манипулирует более сложными структурами данных. Действительно, каков будет результат функции определяющей первый элемент списка, если список пуст. Аналогично, вычисление факториала для отрицательного числа приводит к бесконечному вычислению.

Определенное число исключительных ситуаций может произойти во время выполнения программы, как например деление на ноль. Попытка деления на ноль может привести в лучшем случае к остановке программы и в худшем к неправильному (некоэрентному) состоянию машины. Безопасность языка программирования заключается в гарантии того, что такие случаи не возникнут. Исключение есть один из способов такой гарантии.

Деление 1 на 0 провоцирует запуск специального исключения:

```
# 1/0 ;;
Exception: Division_by_zero.
```

Сообщение `Exception: Division_by_zero.` указывает во первых на то, что исключение `Division_by_zero` возбуждено и во вторых что оно не было обработано.

Часто, тип функции не соответствует области ее определения в том случае когда сопоставление с образцом не является полным, то есть когда он не отслеживает все возможные случаи. Чтобы избежать такой ситуации, Objective Caml выводит следующее сообщение.

```
# let head l = match l with h :: t -> h ;;
Warning: this pattern-matching is not exhaustive.
Here is an example of a value that is not matched:
[]
val head : 'a list -> 'a = <fun>
```

1.3. ТИПИЗАЦИЯ, ОБЛАСТЬ ОПРЕДЕЛЕНИЯ И ИСКЛЮЧЕНИЯ 59

Однако, если программист все таки настаивает на своей версии, Objective Caml воспользуется механизмом исключения в случае неправильного вызова частичной функции.

```
# head [] ;;
Exception: Match_failure ("", 8, -164).
```

Мы уже встречали другое исключение определенное в Objective Caml: **Failure**, с входным аргументом типа **string**. Запустить это исключение можно функцией *failwith*. Воспользуемся этим и дополним определение функции **head**.

```
# let head = function
  [] -> failwith "Empty list"
  | h :: t -> h ;;
val head : 'a list -> 'a = <fun>
# head [] ;;
Exception: Failure "Empty list".
```

1.3.2 Определения исключения

В Objective Caml, тип исключений — **exn**. Это особенный тип — расширяемая сумма: мы можем увеличить множество значений этого типа объявляя новые конструкторы¹⁰. Такая особенность позволяет программисту определять свои собственные исключения.

Синтаксис объявления следующий:

Синтаксис `exception Name ;;`

или

Синтаксис `exception Name of t ;;`

Вот несколько примеров объявления исключений.

```
# exception MY_EXN ;;
exception MY_EXN
# MY_EXN ;;
- : exn = MY_EXN
# exception Depth of int ;;
exception Depth of int
# Depth 4 ;;
- : exn = Depth 4
```

¹⁰Благодаря новой возможности Objective Caml 3.0 «полиморфные варианты» некоторые типы суммы теперь могут расширяться.

Внимание

Имена исключений являются конструкторами — это значит они должны начинаться с заглавной буквы.

```
# exception lowercase ;;
Syntax error
```

Внимание

Исключения мономорфны, при объявлении типа аргумента мы не можем указать параметризующий тип.

```
# exception Value of 'a' ;;
Unbound type parameter 'a'
```

Полиморфное исключение позволило бы определение функций, возвращающих результат любого типа. Подобный пример мы рассмотрим далее на стр. [62](#).

1.3.3 Возбуждение исключения

raise — функциональный примитив языка Objective Caml, ее входной аргумент есть исключение, а возвращает он полиморфный тип.

```
# raise ;;
- : exn -> 'a = <fun>
# raise MY_EXN ;;
Exception: MY_EXN.
# 1 + (raise MY_EXN) ;;
Exception: MY_EXN.
# raise (Depth 4) ;;
Exception: Depth 4.
```

В Objective Caml нет возможности написать функцию *raise*, она должна быть определена системой.

1.3.4 Перехват исключения

Весь интерес возбуждения исключений содержится в возможности их обработать и изменить выполнение программы в зависимости от этого исключения. В этом случае, порядок вычисления выражения имеет значение для того чтобы определить какое исключение было запущено. Таким образом мы выходим из рамок чисто функционального программирования, потому как порядок вычисления аргументов может изменить результат. Об этом мы поговорим в следующей главе (стр. 91)

Следующая конструкция, вычисляющая значение какого-то выражения, отлавливает исключение, возбужденного во время этого вычисления.

Синтаксис	<pre> try <i>expr</i> with <i>p</i>₁ -> <i>expr</i>₁ : <i>p</i>_{<i>n</i>} -> <i>expr</i>_{<i>n</i>} </pre>
-----------	---

Если вычисление *expr* не возбудит исключения, то тип результата будет типом вычисленным этим выражением. Иначе, в приведенном выше сопоставлении будет искаться ответвление соответствующее этому исключению и будет возвращено значение следующего за ним выражения.

Если ни одно ответвление не соответствует исключению, то оно будет переправлено до следующего обработчика **try-with** установленного в программе. Таким образом сопоставление исключения всегда исчерпывающее. Подразумевается что последний фильтр | *e* -> *raise e*. Если ни одного обработчика не найдено в программе, система сама займется обработкой этого исключения и остановит программу с выводом сообщения об ошибке.

Важно понимать разницу между вычислением исключения (то есть значение типа **exn**) и его обработкой, которое приостанавливает вычисления. Исключение, как и любой другой тип, может быть возвращено функцией.

```

# let return x = Failure x ;;
val return : string -> exn = <fun>
# return "test" ;;
- : exn = Failure "test"
# let my_raise x = raise (Failure x) ;;
val my_raise : string -> 'a = <fun>
# my_raise "test" ;;
Exception: Failure "test".

```

Как вы наверно, заметили, функция `my_raise` не возвращает никакого значение, тогда как применение `return` возвращает один из типов `exn`.

Вычисления с исключениями

Кроме обработки не желаемых значений, исключения можно использовать для оптимизации. Следующий пример реализует вычисление произведения всех элементов списка целых чисел. Мы воспользуемся исключением для остановки просмотра списка если обнаружим 0, который мы вернем как результат функции.

```
# exception Found_zero ;;
exception Found_zero
# let rec mult_rec l = match l with
  [] -> 1
  | 0 :: _ -> raise Found_zero
  | n :: x -> n * (mult_rec x) ;;
val mult_rec : int list -> int = <fun>
# let mult_list l =
  try mult_rec l with Found_zero -> 0 ;;
val mult_list : int list -> int = <fun>
# mult_list [1; 2; 3; 0; 5; 6] ;;
- : int = 0
```

Оставшиеся вычисления, то есть произведения элементов после встреченного нуля, не выполняются. После встречи `raise`, вычисление вновь продолжится с `with`.

1.4 Полиморфизм и значения возвращаемые функциями

Полиморфизм Objective Caml позволяет определить функции, тип возвращаемого выражения которых конкретно не определен. Например:

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

Однако, возвращаемый тип `id` зависит от типа аргумента. Таким образом если мы применим функцию `id` к какому-нибудь аргументу, то механизм вывода (определения) типа сможет реализовать переменную

1.4. ПОЛИМОРФИЗМ И ЗНАЧЕНИЯ ВОЗВРАЩАЕМЫЕ ФУНКЦИЯМИ ⁶³

типа 'a. Для каждого частного использования тип функции id может быть определен.

Иначе, использование жесткой статической типизации, которая обеспечивает правильность выполнения, не будет иметь смысла. Функция с полностью неопределенным типом как 'a->'b, разрешит какое попало преобразование типа, что привело бы к ошибке выполнения, так как физическое представление значений не одно и то же.

Очевидное противоречие

Мы можем определить функции, которые возвращают значение, тип которого не соответствует типу аргументов. Рассмотрим несколько примеров и проанализируем почему это не противоречит жесткой статической типизации.

Вот первый пример:

```
# let f x = [] ;;
val f : 'a -> 'b list = <fun>
```

Эта функция конструирует полиморфные значения из любого типа.

```
# f () ;;
- : 'a list = []
# f "anything at all" ;;
- : 'a list = []
```

Однако, полученное значение не совсем является неопределенным — подразумевается список. Таким образом она не может использоваться где попало.

Вот три примера функций типа 'a -> 'b :

```
# let rec f1 x = f1 x ;;
val f1 : 'a -> 'b = <fun>
# let f2 x = failwith "anything at all" ;;
val f2 : 'a -> 'b = <fun>
# let f3 x = List.hd [] ;;
val f3 : 'a -> 'b = <fun>
```

Эти функции не являются «опасными» в отношении правильности выполнения программы, так как их невозможно использовать для конструкции значений: первая заикливется, две последние запускают исключение, которые останавливают выполнение программы.

Чтобы не было возможности определить функции типа `'a -> 'b`, новые конструкторы исключений не должны иметь аргументы типа, который содержит переменную.

Если бы могли объявить полиморфное исключение `Poly_exn` типа `'a -> exn`, то тогда следующая функция была бы возможна:

```
let f = function
  0 -> raise (Poly_exn false)
  | n -> n+1 ;;
```

Таким образом, тип функции `f` это `int -> int` и тип `Poly_exn` это `'a -> exn`, теперь напомним следующее:

```
let g n = try f n with Poly_exn x -> x+1 ;;
```

Эта правильно типизированная функция (поскольку аргумент `Poly_exn` может быть каким угодно) и вызов `g 0` попытается сложить целое с левым значением!

1.5 Калькулятор

Для того чтобы понять как организуются программы на Objective Caml, необходимо самим реализовать такую программу. Напишем калькулятор манипулирующий целыми числами при помощи 4 стандартных операций.

Для начала, определим тип `key`, для представления кнопки калькулятора. Всего таких кнопок будет 15: по одной для каждой операции, для каждой цифры и для кнопки равно.

```
# type key = Plus | Minus | Times | Div | Equals | Digit of
  int ;;
```

Заметим, что цифровые кнопки сгруппированы под одним конструктором `Digit` с аргументом целого типа. Таким образом, некоторые значения типа `key` не являются на самом деле кнопкой. Например, `(Digit 32)` является значением типа `key`, но не представляет ни одну кнопку калькулятора.

Напишем функцию `valid` которая проверяет входной аргумент на принадлежность к типу `key`. Тип функции `key -> bool`.

На первый этапе определим функцию проверяющую что число принадлежит интервалу 0-9. Объявим ее с именем `is_digit`.


```
# let is_digit = function x -> (x >= 0) && (x <= 9) ;;
val is_digit : int -> bool = <fun>
```

Теперь определим функцию `valid` проверяющую свой аргумент тип которого `key`:

```
# let valid ky = match ky with
    Digit n -> is_digit n
  | _       -> true ;;
val valid : key -> bool = <fun>
```

Первый фильтр применяется в случае если входной аргумент построен конструктором `Digit`, в этом случае он проверяется функцией `is_digit`. Второй фильтр применяется в любом другом случае. Напомним, что благодаря фильтру, тип проверяемого значения обязательно будет `key`.

Перед тем как начать реализовывать алгоритм калькулятора, определим модель формально описывающую реакцию на нажатие кнопок. Мы подразумеваем что калькулятор располагает памятью в которой хранится результат последней операции, последняя нажатая кнопка, последний использованный оператор и число выведенное на экран. Все эти 4 значения назовем состоянием калькулятора, оно изменяется каждый раз при нажатии на одну из кнопок. Это изменение называется переходом, а теория управляющая подобным механизмом есть теория автоматов.

Состояние представлено следующим типом:

```
# type state = {
    lcd : int; (* результат последней операции *)
    lka : key; (* последняя нажатая кнопка      *)
    loa : key; (* последний оператор            *)
    vpr : int  (* число выведенное на экран     *)
} ;;
```

В таблице 1.5 приведен пример состояний калькулятора.

Нам нужна функция `evaluate` с тремя аргументами: двумя операндами и оператором, она возвращает результат операции. Для этого функция фильтрует входной аргумент типа `key`:

```
# let evaluate x y ky = match ky with
    Plus      -> x + y
  | Minus     -> x - y
  | Times     -> x * y
  | Div       -> x / y
```

	состояние	кнопка
	(0, =, =, 0)	3
→	(0, 3, =, 3)	+
→	(3, +, +, 3)	2
→	(3, 2, +, 2)	1
→	(3, 1, +, 21)	×
→	(24, *, *, 24)	2
→	(24, 2, *, 2)	=
→	(48, =, =, 48)	

Таблица 1.5: Переходы для $3+21*2=$

```

| Equals  -> y
| Digit _ -> failwith "evaluate : no op" ;;
val evaluate : int -> int -> key -> int = <fun>

```

Дадим определение переходов, перечисляя все возможные случаи. Предположим, что текущее состояние есть квадриплет (a, b, \oplus, d) :

- кнопка с цифрой x нажата, есть два возможных случая:
 - предыдущая нажатая кнопка тоже цифра, значит пользователь продолжает набирать число и необходимо добавить x к значению выводимому на экран то есть изменить его на $d \times 10 + x$. Новое состояние будет:

$$(a, (\text{Digit } x), \oplus, d \times 10 + x)$$

- предыдущая нажатая кнопка не цифра, тогда мы только начинаем писать новое число. Новое состояние:

$$(a, (\text{Digit } x), \oplus, x)$$

- кнопка с оператором \otimes была нажата, значит второй операнд был полностью введен и теперь мы хотим «что-то» сделать с двумя операндами. Для этого сохранена последняя операция (\oplus) . Новое состояние:

$$(\oplus d, \otimes, \otimes, a \oplus d)$$

Для того чтобы написать функцию `transition`, достаточно дословно перевести в Objective Caml предыдущее описание при помощи сопоставления входного аргумента. Кнопку с двумя значениями обработаем локальной функцией `digit_transition` при помощи сопоставления.

```
# let transition st ky =
  let digit_transition n = function
    Digit _ -> { st with lka=ky; vpr=st.vpr*10+n }
    | _      -> { st with lka=ky; vpr=n }
  in
    match ky with
    Digit p -> digit_transition p st.lka
    | _      -> let res = evaluate st.lcd st.vpr st.loa
                  in { lcd=res; lka=ky; loa=ky; vpr=res };;
val transition : state -> key -> state = <fun>
```

Эта функция из `state` и `key` считает новое состояние `state`.

Протестируем теперь программу:

```
# let initial_state = lcd=0; lka=Equals; loa=Equals; vpr=0 ;;
val initial_state : state = lcd = 0; lka = Equals; loa = ←
  Equals; vpr = 0
# let state2 = transition initial_state (Digit 3) ;;
val state2 : state = lcd = 0; lka = Digit 3; loa = Equals; ←
  vpr = 3
# let state3 = transition state2 Plus ;;
val state3 : state = lcd = 3; lka = Plus; loa = Plus; vpr = 3
# let state4 = transition state3 (Digit 2) ;;
val state4 : state = lcd = 3; lka = Digit 2; loa = Plus; ←
  vpr = 2
# let state5 = transition state4 (Digit 1) ;;
val state5 : state = lcd = 3; lka = Digit 1; loa = Plus; ←
  vpr = 21
# let state6 = transition state5 Times ;;
val state6 : state = lcd = 24; lka = Times; loa = Times; ←
  vpr = 24
# let state7 = transition state6 (Digit 2) ;;
val state7 : state = lcd = 24; lka = Digit 2; loa = Times; ←
  vpr = 2
# let state8 = transition state7 Equals ;;
val state8 : state = lcd = 48; lka = Equals; loa = Equals; ←
  vpr = 48
```

Мы можем сделать тоже самое, передав список переходов.

```
# let transition_list st ls = List.fold_left transition st ls ;;
```

```
val transition_list : state -> key list -> state = <fun>
# let example = [ Digit 3; Plus; Digit 2; Digit 1; Times;  ◀
    Digit 2; Equals ]
  in transition_list initial_state example ;;
- : state = lcd = 48; lka = Equals; loa = Equals; vpr = 48
```

1.6 Резюме

В этой главе мы изучили основы функционального программирования и параметризованный полиморфизм, что вместе являются основными характеристиками Objective Caml. Также мы описали синтаксис функциональной части ядра и типов. Это дало нам возможность написать первые программы. Так же мы подчеркнули глубокую разницу между типом функции и областью ее применения. Механизм исключений позволяет решить эту проблему, и в то же время вводит новый стиль программирования, где мы явно указываем манеру вычисления.

Глава 2

Императивное программирование

Введение

В отличие от функционального программирования, где значение вычисляется посредством применения функции к аргументам, не заботясь о том как это происходит, императивное программирование ближе к машинному представлению, так как оно вводит понятие состояния памяти, которое изменяется под воздействием программы. Каждое такое воздействие называется **инструкцией** и императивная программа есть **последовательность** упорядоченных инструкций. Состояние памяти может быть изменено при выполнении каждой инструкции. Операции ввода/вывода можно рассматривать как изменение оперативной памяти, видео памяти или файлов.

Подобный стиль программирования напрямую происходит от программирования на ассемблере. Мы встречаем этот стиль в языках программирования первого поколения (Fortran, C, Pascal, ...). Следующие элементы Objective Caml соответствуют приведенной модели:

- физически изменяемые¹ структуры данных, такие как массив или запись с изменяемыми (mutable) полями
- операции ввода/вывода
- структуры контроля выполнения программы как цикл и исключения.

¹В оригинальном издании (французском) используется выражение *физически изменяемые*, тогда как в английском переводе лишь *изменяемые*

Некоторые алгоритмы реализуются проще таким стилем программирования. Примером может послужить произведение двух матриц. Несмотря на то что эту операцию можно реализовать чисто функциональным способом, при этом списки заменят массивы, это не будет ни эффективнее, ни естественней по отношению к императивному стилю.

Интерес интеграции императивной модели в функциональный язык состоит в написании при необходимости подобных алгоритмов в этом стиле программирования. Два главных недостатка императивного программирования по отношению к функциональному это:

- усложнение системы типов языка и отклонение (rejecting) некоторых программ, которые бы без этого рассматривались бы как «корректные»,
- необходимость учитывать состоянием памяти и порядок вычисления.

Однако, при соблюдении некоторых правил написания программ, выбор стиля программирования предоставляет большие возможности в написании алгоритмов, что является главной целью языков программирования. К тому же, программа написанная в стиле близкому к стилю алгоритма, имеет больше шансов быть корректной (или по крайней мере быстрее реализована).

По этим причинам в Objective Caml имеются типы данных, значения которых физически изменяемые, структуры контроля выполнения программ и библиотека ввода/вывода в императивном стиле.

План главы

В данной главе продолжается представление базовых элементов Objective Caml затронутых в предыдущей главе, но теперь мы заинтересуемся императивными конструкциями. Глава разбита на 5 разделов. Первый, наиболее важный, раскрывает различные физически изменяемые структуры данных и описывает их представление в памяти. Во втором разделе кратко излагаются базовые операции ввода/вывода. Третий знакомит с новыми итеративными управляющими операторами. В четвертом разделе рассказывается об особенностях выполнения императивных программ, в частности о порядке вычисления аргументов функции. В последнем разделе мы переделаем калькулятор предыдущей главы в калькулятор с памятью.

2.1 Физически изменяемые структуры данных

Значения следующих типов: массивы, строки, записи с модифицируемыми полями и указатели есть структурированные значения, компоненты которых могут быть физически изменены.

Мы уже видели, что переменная в Objective Caml привязана к значению и хранит эту связку до конца ее существования. Мы можем изменить связку лишь переопределением, но и в этом случае речь не будет идти о той же самой переменной. Новая переменная с тем же именем скроет предыдущую, которая перестанет быть доступной напрямую и останется неизменной. В случае с изменяемыми переменными, мы можем ассоциировать новое значение переменной без ее переопределения. Значение переменной доступно в режиме чтение/запись.

2.1.1 Векторы

Векторы или одномерные массивы группируют определенное число однотипных элементов. Существует несколько способов создания векторов, первый — перечисление элементов, разделенных запятой и заключенных между символами '[' и ']'.

```
# let v = [| 3.14; 6.28; 9.42 |] ;;
val v : float array = [|3.14; 6.28; 9.42|]
```

Функция *Array.create* с двумя аргументами на входе: размер вектора и начальное значение, возвращает новый вектор.

```
# let v = Array.create 3 3.14 ;;
val v : float array = [|3.14; 3.14; 3.14|]
```

Для того чтобы изменить или просто посмотреть значение элемента необходимо указать его номер.

Синтаксис *expr*₁ . (*expr*₂)

Синтаксис *expr*₁ . (*expr*₂) <- *expr*₃

*expr*₁ должен быть вектором (тип **array**) с элементами типа *expr*₃. Конечно, выражение *expr*₂ должно быть типа **int**. Результат модификации есть выражение типа **unit**. Номер первого элемента вектора 0 и последнего — размер вектора минус 1. Скобки вокруг выражения обязательны.

```
# v . (1) ;;
- : float = 3.14
# v.(0) <- 100.0 ;;
- : unit = ()
# v ;;
- : float array = [|100.; 3.14; 3.14|]
```

Если индекс элемента не принадлежит интервалу индексов вектора, то возбуждается исключение в момент доступа.

```
# v.(-1) +. 4.0 ;;
Exception: Invalid_argument "Array.get".
```

Эта проверка осуществляется в момент выполнения программы, что может сказаться на ее скорости. Однако, это необходимо для избежания заполнения зоны памяти вне вектора, что может привести к серьезным ошибкам.

Функции манипуляции векторами являются частью модуля **Array** стандартной библиотеки, описание которого дано в главе ?? на стр. ?. В следующих примерах мы воспользуемся тремя функциями из этого модуля:

create создающая вектор заданного размера и начальным значением

length возвращающая размер вектора

append для конкатенации двух векторов

Совместное использование значений вектора

Все элементы вектора содержат значение, переданное во время создания вектора. В случае, если это значение структурное, оно будет совместно использоваться (sharing). Создадим, для примера, матрицу как вектор векторов при помощи функции *Array.create*:

```
# let v = Array.create 3 0 ;;
val v : int array = [|0; 0; 0|]
# let m = Array.create 3 v ;;
val m : int array array = [| [|0; 0; 0|]; [|0; 0; 0|]; [|0; 0; 0|] |]
```

Если поменять одно из полей вектора *v*, который мы использовали для создания *m*, мы изменим все *строки* матрицы (см. рисунки [2.1](#) и [2.2](#)).

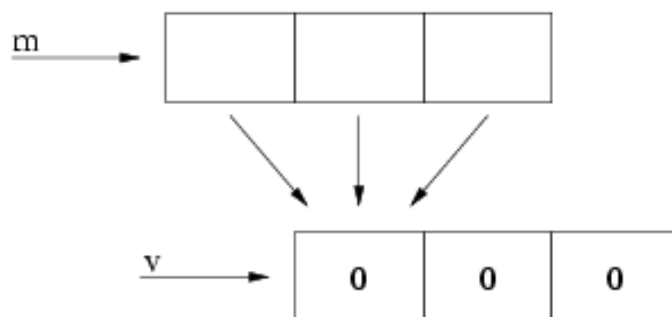


Рис. 2.1: Представление в памяти вектора с разделением элементов

```
# v.(0) <- 1 ;;
- : unit = ()
# m ;;
- : int array array = [|1; 0; 0|]; [|1; 0; 0|]; [|1; 0; 0|]|
```

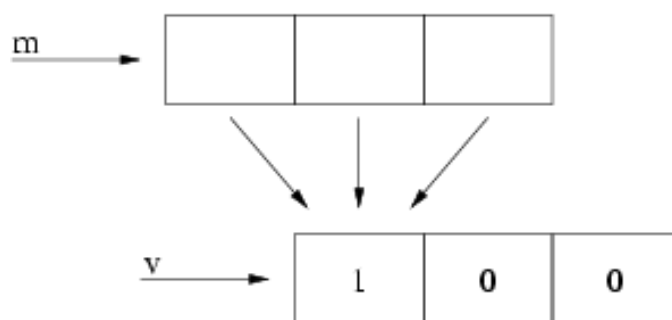


Рис. 2.2: Изменение элемента разделяемого вектора

Если значение инициализации вектора (второй аргумент функции *Array.create*) простое, атомарное, то оно копируется каждый раз и совместно используется если это значение структурное.

Мы называем **атомарным значением** то, размер которого не превышает стандартный размер значений Objective Caml, то есть **memory word** — целое, символ, булевый и конструкторы констант. Другие значения, называемые структурные, представлены указателем на зону памяти. Эта разница объяснена более детально в главе ?? стр. ??.

Векторы чисел с плавающей запятой есть особый случай. Несмотря на то что эти числа (**float**) — структурные значения, при создании вектора начальное значение копируется. Делается это для оптимизации. В главе ?? стр. ??, в которой обсуждается интерфейс с языком C мы обсудим эту проблему.

Не прямоугольная матрица

Матрица, вектор векторов, не обязательно должна быть прямоугольной. Действительно, мы можем заменить вектор на другой вектор с иным размером, что удобно для ограничения размера матрицы. Следующее значение `t` создает треугольную матрицу для коэффициентов треугольника Паскаля.

```
# let t = [|
    [|1|];
    [|1; 1|];
    [|1; 2; 1|];
    [|1; 3; 3; 1|];
    [|1; 4; 6; 4; 1|];
    [|1; 5; 10; 10; 5; 1|]
  |] ;;
val t : int array array =
  [| [|1|]; [|1; 1|]; [|1; 2; 1|]; [|1; 3; 3; 1|]; [|1; 4; 6; 4; 1|];
    [|1; 5; 10; 10; 5; 1|] |]
# t.(3) ;;
- : int array = [|1; 3; 3; 1|]
```

В этом примере элемент вектора `t` по индексу `i` есть вектор целых чисел размером `i+1`. Для манипуляции такой матрицей необходимо знать размер каждого элемента вектора.

Копирование векторов

При копировании вектора или конкатенации двух векторов мы получаем новый вектор. Модификация исходных векторов не повлияет на значение копий, кроме случая разделения значений рассмотренного ранее.

```
# let v2 = Array.copy v ;;
val v2 : int array = [|1; 0; 0|]
# let m2 = Array.copy m ;;
val m2 : int array array = [| [|1; 0; 0|]; [|1; 0; 0|]; [|1; 0; 0|];
  [|1; 0; 0|] |]
# v.(1) <- 352 ;;
- : unit = ()
# v2 ;;
```

```
- : int array = [|1; 0; 0|]
# m2 ;;
- : int array array = [| [|1; 352; 0|]; [|1; 352; 0|]; [|1; 352; 0|]|]
```

В приведенном примере видно что копия `m` содержит лишь указатели на `v`, если один из элементов `v` изменен, то `m2` тоже будет изменен. Конкатенация создает новый вектор длина которого равна сумме длин двух других.

```
# let mm = Array.append m m ;;
val mm : int array array =
  [| [|1; 352; 0|]; [|1; 352; 0|]; [|1; 352; 0|]; [|1; 352; 0|];
    [|1; 352; 0|]; [|1; 352; 0|]|]
# Array.length mm ;;
- : int = 6
# m.(0) <- Array.create 3 0 ;;
- : unit = ()
# m ;;
- : int array array = [| [|0; 0; 0|]; [|1; 352; 0|]; [|1; 352; 0|]|]
# mm ;;
- : int array array =
  [| [|1; 352; 0|]; [|1; 352; 0|]; [|1; 352; 0|]; [|1; 352; 0|];
    [|1; 352; 0|]; [|1; 352; 0|]|]
```

Однако, изменение `v`, разделяемого `m` и `mm`, приведет к изменению обеих матриц:

```
# v.(1) <- 18 ;;
- : unit = ()
- : int array array =
  [| [|1; 18; 0|]; [|1; 18; 0|]; [|1; 18; 0|]; [|1; 18; 0|]; [|1; 18; 0|]|]
```

2.1.2 Строки

Строки можно рассматривать как частный случай массива символов. Однако, по причинам использования памяти этот тип специализирован и синтаксис доступа к элементам изменен.

Синтаксис $expr_1 . [expr_2]$

Элементы строки могут быть физически изменены.

Синтаксис $expr_1 . [expr_2] <- expr_3$

```
# let s = "hello" ;;
val s : string = "hello"
# s.[2] ;;
- : char = 'l'
# s.[2] <- 'Z' ;;
- : unit = ()
# s ;;
- : string = "heZlo"
```

2.1.3 Изменяемые поля записей

Поля записи могут быть объявлены изменяемыми. Для этого достаточно указать при декларации типа поля ключевое слово **mutable**

Синтаксис **type** $name = \{ \dots; \text{mutable } name_i : t ; \dots \}$

Пример определения точки плоскости.

```
# type point = { mutable xc : float; mutable yc : float } ;;
type point = { mutable xc : float; mutable yc : float; }
# let p = { xc = 1.0; yc = 0.0 } ;;
val p : point = {xc = 1.; yc = 0.}
```

Для изменения значение поля объявленного **mutable** используйте следующий синтаксис.

Синтаксис $expr_1 . name <- expr_2$

Выражение $expr_1$ должно быть типа запись содержащее поле $name$. Операция модификации возвращает значение типа **unit**.

```
# p.xc <- 3.0 ;;
- : unit = ()
# p ;;
- : point = {xc = 3.; yc = 0.}
```

Напишем функцию перемещения точки, которая изменяет ее координаты. Здесь мы используем локальную декларацию с фильтром, для упорядочения побочных эффектов.

```
# let moveto p dx dy =
  let () = p.xc <- p.xc +. dx
  in p.yc <- p.yc +. dy ;;
val moveto : point -> float -> float -> unit = <fun>
# moveto p 1.1 2.2 ;;
- : unit = ()
# p ;;
- : point = {xc = 4.1; yc = 2.2}
```

Мы можем определить изменяемые или не изменяемые поля, только поля, помеченные ключевым словом **mutable**, будут изменяемые.

```
# type t = { c1 : int; mutable c2 : int } ;;
type t = { c1 : int; mutable c2 : int; }
# let r = { c1 = 0; c2 = 0 } ;;
val r : t = {c1 = 0; c2 = 0}
# r.c1 <- 1 ;;
The record field label c1 is not mutable
# r.c2 <- 1 ;;
- : unit = ()
# r ;;
- : t = {c1 = 0; c2 = 1}
```

Далее, мы приводим пример использования записей с изменяемыми полями и массивов для того чтобы реализовать структуру стека (см. [2.3.3 стр. 87](#)).

2.1.4 Указатели

Objective Caml предлагает полиморфный тип **ref**, рассматриваемый как указатель на любое значение. В Objective Caml указатель точнее будет называть **указатель на значение**. Указываемое значение может быть изменено. Тип **ref** определен как запись с изменяемым полем.

```
type 'a ref = mutable contents:'a
```

Создать ссылку на значение можно функцией *ref*. Указываемое значение может быть получено использованием префикса **!**. Для модификации значения используем инфиксную функцию **:=**.

```
# let x = ref 3 ;;
val x : int ref = {contents = 3}
# x ;;
- : int ref = {contents = 3}
# !x ;;
- : int = 3
# x := 4 ;;
- : unit = ()
# !x ;;
- : int = 4
# x := !x + 1 ;;
- : unit = ()
# !x ;;
- : int = 5
```

2.1.5 Полиморфизм и изменяемые значения

Тип `ref` параметризованный (см. 1.2.6), что позволяет создавать указатели на любой тип. Однако, существует несколько ограничений. Представим, что нет никаких ограничений и мы можем объявить

```
let x = ref [] ;;
```

В этом случае тип переменной `x` будет `'a list ref` и можем изменить значение способом не соответствующим статической типизации Objective Caml.

```
x := 1 :: !x ;;
x := true :: !x ;;
```

При этом получаем одну и ту же переменную типа `int list` в первой строке и `bool list` в другой.

Во избежании подобной ситуации, система типов Objective Caml использует новую категорию типа переменной — **слабо типизированные переменные**. Синтаксически, они отличаются предшествующим символом подчеркивания.

```
# let x = ref [] ;;
val x : '_a list ref = {contents = []}
```

Переменная типа `'_a` не является параметром типа, то есть ее тип не известен до момента ее реализации. Лишь в момент первого использования `x`, после объявления, тип будет окончательно зафиксирован.

```
# x := 0 :: !x ;;
- : unit = ()
# x ;;
- : int list ref = {contents = [0]}
```

Таким образом тип переменной `x` — `int list ref`.

Тип, содержащий неизвестную, является мономорфным, не смотря на то что его тип еще не был указан и невозможно реализовать то что неизвестно полиморфным типом.

```
# let x = ref [] ;;
val x : 'a list ref = {contents = []}
# x := (function y -> ()) :: !x ;;
- : unit = ()
# x ;;
- : ('a -> unit) list ref = {contents = [<fun>]}
```

В этом примере, не смотря на то что мы реализуем неизвестную типом *a priori* полиморфным (`'a -> unit`), тип остался мономорфным с новой неизвестной типа.

Это ограничение полиморфизма распространяется не только на указатели, но и на любое значение содержащее изменяемую часть: массивы, записи с полями объявленные **mutable**, и так далее. Все параметры типа, даже не имеющие никакого отношения к модифицируемым атрибутам, есть переменные слабого (weak) типа.

```
# type ('a, 'b) t = { ch1 : 'a list; mutable ch2 : 'b list } ◀
;;
type ('a, 'b) t = { ch1 : 'a list; mutable ch2 : 'b list; }
# let x = { ch1 = []; ch2 = [] } ;;
val x : ('a, 'b) t = {ch1 = []; ch2 = []}
```

Внимание

Эта модификация типа повлияет на чисто функциональные программы.

Если к полиморфной переменной применена полиморфная функция, то в результате мы получим переменную слабого типа, так как нельзя не учитывать то что функция может создать физически изменяемое значение. Иными словами, результат будет всегда мономорфный.

```
# (function x -> x) [] ;;  
- : 'a list = []
```

Тот же результат мы получим для частичных функций.

```
# let f a b = a ;;  
val f : 'a -> 'b -> 'a = <fun>  
# let g = f 1 ;;  
val g : '_a -> int = <fun>
```

Для получения полиморфного типа необходимо вычесть второй аргумент `f` и применить ее:

```
# let h x = f 1 x ;;  
val h : 'a -> int = <fun>
```

Действительно, выражения определяющее `h` есть функциональное выражение `function x -> f 1 x`. Его вычисление даст замыкание, которое не рискует произвести побочный эффект, так как тело функции не вычислено.

В общем случае, мы различаем выражение не расширяемое, расчет которого не приведет к побочному эффекту и другие расширяемые. Система типов Objective Caml классифицирует выражения в соответствии с синтаксисом:

- выражения не расширяемые состоят в основном из переменных, конструкторов не изменяемых значений и абстракций
- выражения расширяемые состоят из применений, конструкторов изменяемых значений. Так же к ним необходимо добавить структуры контроля: условный (*conditional*) или фильтр.

2.2 Ввод-вывод

Функции ввода-вывода вычисляют значение (чаще всего типа `unit`), однако в ходе вычисления, они изменяют состояние устройств ввода-вывода: изменение буфера клавиатуры, перерисовка экрана, запись в файл и так далее. Для каналов ввода и вывода определены для типа: `in_channel` и `out_channel`. При обнаружении конца файла возбуждается исключение `End_of_file`. Следующие константы соответствуют стандартным каналам ввода, вывода и ошибок UNIX: `stdin`, `stdout` и `stderr`.

2.2.1 Каналы

Функции ввода-вывода стандартной библиотеки Objective Caml, манипулируют каналами типа `in_channel` и `out_channel` соответственно. Для создания подобных каналов используется функция:

```
# open_in ;;
- : string -> in_channel = <fun>
# open_out ;;
- : string -> out_channel = <fun>
```

`open_in` открывает файл², если он не существует, возбуждается исключение `Sys_error`.

`open_out` создает указанный файл если такой не существует, если же он существует то его содержимое уничтожается.

```
# let ic = open_in "koala" ;;
val ic : in_channel = <abstr>
# let oc = open_out "koala" ;;
val oc : out_channel = <abstr>
```

Функции закрывающие каналы:

```
# close_in ;;
- : in_channel -> unit = <fun>
# close_out ;;
- : out_channel -> unit = <fun>
```

2.2.2 Чтение и запись

Стандартные функции чтения и записи:

```
# input_line ;;
- : in_channel -> string = <fun>
# input ;;
- : in_channel -> string -> int -> int -> int = <fun>
# output ;;
- : out_channel -> string -> int -> int -> unit = <fun>
```

- `input_line ic`: читает из входного канала `ic` символы до обнаружения возврата каретки или конца файла и возвращает в форме строки (не включая возврат каретки).

²с правом на чтение, при необходимости

- *input ic s p l*: считывает *l* символов из канала *ic* и помещает их в строку *s* со смещением *p*. Функция возвращает число прочитанных символов.
- *output oc s p l*: записывает в выходной канал *oc* часть строки *s* начиная с символа *p* длиной *l*.

Следующие функции чтения из стандартного ввода и записи в стандартный вывод:

```
# read_line ;;
- : unit -> string = <fun>
# print_string ;;
- : string -> unit = <fun>
# print_newline ;;
- : unit -> unit = <fun>
```

Другие значения простых типов данных могут быть прочтены или записаны, значения таких типов могут быть переведены в тип список символов.

Локальное объявление и порядок выполнения Попробуем вывести на экран выражение формы **let x = e1 in e2**. Зная, что в общем случае локальная переменная *x* может быть использована в *e2*, выражение *e1* вычислено первым и только затем *e2*. Если эти оба выражения есть императивные функции с побочным эффектом результат которых (), мы выполнили их в правильном порядке. В частности, так как нам известен тип возвращаемого результата *e1*: константа () типа *unit*, мы получим последовательный вывод на экран используя сопоставление с образцом '()'.

```
# let () = print_string "and one," in
  let () = print_string " and two," in
  let () = print_string " and three" in
    print_string " zero" ;;
and one, and two, and three zero- : unit = ()
```

2.2.3 Пример «Больше/Меньше»

В следующем примере мы приводим игру «Больше/Меньше», состоящую в поиске числа, после каждого ответа программа выводит на экран сообщение в соответствии с тем что предложенное число больше или меньше загаданного.

```
# let rec hilo n =
  let () = print_string "type a number: " in
  let i = read_int ()
  in
    if i = n then
      let () = print_string "BRAVO" in
      let () = print_newline ()
      in print_newline ()
    else
      let () =
        if i < n then
          let () = print_string "Higher"
          in print_newline ()
        else
          let () = print_string "Lower"
          in print_newline ()
      in hilo n ;;
val hilo : int -> unit = <fun>
```

Результат запуска

```
# hilo 64 ;;
type a number: 88
Lower
type a number: 44
Higher
type a number: 64
BRAVO

- : unit = ()
```

2.3 Структуры контроля

Операции ввода/вывода и изменяемые значения имеют побочный эффект. Их использование упрощено императивным стилем программирования с новыми структурами контроля. В этом параграфе мы поговорим о последовательных и итеративных структурах.

Нам уже приходилось встречать (см. [1.1.2](#) стр. 16) условную структуру контроля `if then` смысл которой такой же как и в императивных языках программирования. Например:

```
# let n = ref 1 ;;
val n : int ref = {contents = 1}
# if !n > 0 then n := !n - 1 ;;
- : unit = ()
```

2.3.1 Последовательность

Самая типичная императивная структура — последовательность, которая позволяет вычислять выражения разделенные точкой с запятой в порядке слева направо.

Синтаксис $expr_1 ; \dots ; expr_n$

Последовательность — это выражение, результат которого есть результат последнего выражения (здесь $expr_n$). Все выражения вычислены и побочный эффект каждой учтен.

```
# print_string "2 = "; 1+1 ;;
2 = - : int = 2
```

С побочным эффектом, мы получаем обычные конструкции императивного программирования.

```
# let x = ref 1 ;;
val x : int ref = {contents = 1}
# x := !x+1; x := !x*4; !x ;;
- : int = 8
```

В связи с тем что значения предшествующие точке запятой не сохранены, Objective Caml выводит предупреждение в случае если их тип отличен от типа `unit`.

```
# print_int 1; 2; 3 ;;
Warning: this expression should have type unit.
1- : int = 3
```

Чтобы избежать этого сообщения, можно воспользоваться функцией `ignore`.

```
# print_int 1; ignore 2; 3 ;;
1- : int = 3
```

Другое сообщение будет выведено в случае если забыт аргумент функции

```
# let g x y = x := y ;;
val g : 'a ref -> 'a -> unit = <fun>
# let a = ref 10 ;;
val a : int ref = {contents = 10}
# let u = 1 in g a ; g a u ;;
Warning: this function application is partial,
maybe some arguments are missing.
- : unit = ()
# let u = !a in ignore (g a) ; g a u ;;
Warning: this function application is partial,
maybe some arguments are missing.
- : unit = ()
```

Чаще всего мы используем скобки для определения зоны видимости.

Синтаксис `(expr)`

Синтаксис `begin expr end`

Теперь мы можем написать Больше/Меньше простым стилем (стр. 82).

```
# let rec hilo n =
  print_string "type a number: ";
  let i = read_int () in
  if i = n then print_string "BRAVO\n\n"
  else
    begin
      if i < n then print_string "Higher\n"
      else print_string "Lower\n" ;
      hilo n
    end ;;
val hilo : int -> unit = <fun>
```

2.3.2 Циклы

Структуры итеративного контроля тоже не принадлежат «миру» функционального программирования. Условие повторения цикла или выхода из него имеет смысл в случае когда есть физическое изменение памяти. Существует две структуры итеративного контроля: цикл **for** для ограниченного количества итераций и **while** для неограниченного. Структуры цикла возвращают константу `()` типа `unit`.

Цикл **for** может быть возрастающим (**to**) или убывающим (**downto**) с шагом в одну единицу.

Синтаксис	<pre>for name = <i>expr</i>₁ to <i>expr</i>₂ do <i>expr</i>₃ done for name = <i>expr</i>₁ downto <i>expr</i>₂ do <i>expr</i>₃ done</pre>
-----------	--

Тип выражений *expr*₁ и *expr*₂ — `int`. Если тип *expr*₃ не `unit`, компилятор выдаст предупреждение.

```
# for i=1 to 10 do print_int i; print_string " " done; ◀
    print_newline () ;;
1 2 3 4 5 6 7 8 9 10
- : unit = ()
# for i=10 downto 1 do print_int i; print_string " " done; ◀
    print_newline () ;;
10 9 8 7 6 5 4 3 2 1
- : unit = ()
```

Неограниченный цикл **while** имеет следующий синтаксис

Синтаксис	<code>while <i>expr</i>₁ do <i>expr</i>₂ done</code>
-----------	--

Тип выражения *expr*₁ должен быть `bool`. И, как в случае цикла **for**, если тип *expr*₂ не `unit`, компилятор выдаст предупреждение.

```
# let r = ref 1
    in while !r < 11 do
        print_int !r ;
        print_string " " ;
        r := !r+1
    done ;;
1 2 3 4 5 6 7 8 9 10 - : unit = ()
```

Необходимо помнить, что циклы — это такие же выражения как и любые другие, они вычисляют значение `()` типа `unit`.

```
# let f () = print_string "-- end\n" ;;
val f : unit -> unit = <fun>
# f (for i=1 to 10 do print_int i; print_string " " done) ;;
1 2 3 4 5 6 7 8 9 10 -- end
- : unit = ()
```


Обратите внимание на то, что строка ‘- end\n’ выводится после цифр 1...10: в подтверждение того что аргументы (в данном случае цикл) вычисляются перед передачей функции.

В императивном стиле тело цикла (выражение *expr*) не возвращает результата, а производит побочный эффект. В Objective Caml, если тип тела цикла не `unit`, то компилятор выдает сообщение :

```
# let s = [5; 4; 3; 2; 1; 0] ;;
val s : int list = [5; 4; 3; 2; 1; 0]
# for i=0 to 5 do List.tl s done ;;
Warning: this expression should have type unit.
- : unit = ()
```

2.3.3 Пример: реализация стека

В нашем примере структура данных `'a stack` будет запись с двумя полями: массив элементов и индекс первой свободной позиции в массиве.

```
# type 'a stack = { mutable ind:int; size:int; mutable elts: 
'a array } ;;
```

В поле `size` будет храниться максимальный размер стека.

Операции над стеком будут `init_stack` для инициализации, `push` для добавления и `pop` для удаления элемента.

```
# let init_stack n = { ind=0; size=n; elts=[] } ;;
val init_stack : int -> 'a stack = <fun>
```

Эта функция не может создавать не пустой массив, так для создания массива необходимо передать элемент. Поэтому поле `elts` получает пустой массив.

Два исключения добавлены на случай попытки удалить элемент из пустого стека и добавить элемент в полный. Они используются в функциях `pop` и `push`.

```

# exception Stack_empty ;;
exception Stack_empty
# exception Stack_full ;;
exception Stack_full
# let pop p =
    if p.ind = 0 then raise Stack_empty
    else (p.ind <- p.ind - 1; p.elts.(p.ind)) ;;
val pop : 'a stack -> 'a = <fun>
# let push e p =
    if p.elts = [| |] then
        (p.elts <- Array.create p.size e; p.ind <- 1)
    else if p.ind >= p.size then raise Stack_full
    else (p.elts.(p.ind) <- e; p.ind <- p.ind + 1) ;;
val push : 'a -> 'a stack -> unit = <fun>

```

Небольшой пример использования:

```

# let p = init_stack 4 ;;
val p : 'a stack = {ind = 0; size = 4; elts = [| |]}
# push 1 p ;;
- : unit = ()
# for i=2 to 5 do push i p done ;;
Exception: Stack_full.
# p ;;
- : int stack = {ind = 4; size = 4; elts = [|1; 2; 3; 4|]}
# pop p ;;
- : int = 4
# pop p ;;
- : int = 3

```

Чтобы избежать исключения `Stack_full` при переполнении стека, мы можем каждый раз увеличивать его размер. Для этого нужно чтобы поле `size` было модифицируемым.

```

# type 'a stack =
    {mutable ind:int; mutable size:int; mutable elts: 'a ←
      array} ;;
type 'a stack = {
    mutable ind : int;
    mutable size : int;
    mutable elts : 'a array;

```



```

}
# let init_stack n = {ind=0; size = max n 1; elts = [||]} ;;
val init_stack : int -> 'a stack = <fun>
# let n_push e p =
  if p.elts = [||] then
    begin
      p.elts <- Array.create p.size e;
      p.ind <- 1
    end
  else if p.ind >= p.size then
    begin
      let nt = 2 * p.size in
      let nv = Array.create nt e in
      for j=0 to p.size-1 do nv.(j) <- p.elts.(j) done;
      p.elts <- nv;
      p.size <- nt;
      p.ind <- p.ind + 1
    end
  else
    begin
      p.elts.(p.ind) <- e;
      p.ind <- p.ind + 1
    end ;;
val n_push : 'a -> 'a stack -> unit = <fun>

```

Однако необходимо быть осторожным со структурами которые могут беспрестанно увеличиваться. Вот небольшой пример стека, увеличивающегося по необходимости.

```

# let p = init_stack 4 ;;
val p : 'a stack = {ind = 0; size = 4; elts = [||]}
# for i=1 to 5 do n_push i p done ;;
- : unit = ()
# p ;;
- : int stack = {ind = 5; size = 8; elts = [|1; 2; 3; 4; 5; 5; 5; 5|]}
# p.stack ;;
Unbound record field label stack

```

В функции pop желательно добавить возможность уменьшения размера стека, это позволит сэкономить память.

2.3.4 Пример: расчет матриц

В этом примере мы определим тип «матрица» — двумерный массив чисел с плавающей запятой и несколько функций. Мономорфный тип `mat` это запись, состоящая из размеров и элементов матрицы. Функции `creat_mat`, `access_mat` и `mod_mat` служат для создания, доступа и изменения элементов матрицы.

```
# type mat = {n:int; m:int; t:float array array} ;;
type mat = { n : int; m : int; t : float array array; }
# let create_mat n m = {n=n; m=m; t = Array.create_matrix n m ◀
    0.0} ;;
val create_mat : int -> int -> mat = <fun>
# let access_mat m i j = m.t.(i).(j) ;;
val access_mat : mat -> int -> int -> float = <fun>
# let mod_mat m i j e = m.t.(i).(j) <- e ;;
val mod_mat : mat -> int -> int -> float -> unit = <fun>
# let a = create_mat 3 3 ;;
val a : mat =
    {n = 3; m = 3; t = [| [|0.; 0.; 0.]; [|0.; 0.; 0.]; [|0.; ◀
        0.; 0.]| |]}
# mod_mat a 1 1 2.0; mod_mat a 1 2 1.0; mod_mat a 2 1 1.0 ;;
- : unit = ()
# a ;;
- : mat =
    {n = 3; m = 3; t = [| [|0.; 0.; 0.]; [|0.; 2.; 1.]; [|0.; 1.; ◀
        0.]| |]}

```

Сумма двух матриц a и b есть матрица c , такая что

$$c_{ij} = a_{ij} + b_{ij}$$

```
# let add_mat p q =
    if p.n = q.n && p.m = q.m then
        let r = create_mat p.n p.m in
        for i = 0 to p.n-1 do
            for j = 0 to p.m-1 do
                mod_mat r i j (p.t.(i).(j) +. q.t.(i).(j))
            done
        done;
        r
    else failwith "add_mat : dimensions incompatible" ;;
val add_mat : mat -> mat -> mat = <fun>

```

```
# add_mat a a ;;
- : mat =
{n = 3; m = 3; t = [| [|0.; 0.; 0.]; [|0.; 4.; 2.]; [|0.; 2.; 0.]; |]}
```

Произведение двух матриц a и b есть матрица c , такая что $c_{ij} = \sum_{k=1}^{m_a} a_{ik} \cdot b_{ki}$

```
# let mul_mat p q =
  if p.m = q.n then
    let r = create_mat p.n q.m in
    for i = 0 to p.n-1 do
      for j = 0 to q.m-1 do
        let c = ref 0.0 in
        for k = 0 to p.m-1 do
          c := !c +. (p.t.(i).(k) *. q.t.(k).(j))
        done;
        mod_mat r i j !c
      done
    done;
    r
  else failwith "mul_mat : dimensions incompatible" ;;
val mul_mat : mat -> mat -> mat = <fun>
# mul_mat a a ;;
- : mat =
{n = 3; m = 3; t = [| [|0.; 0.; 0.]; [|0.; 5.; 2.]; [|0.; 2.; 1.]; |]}
```

2.4 Порядок вычисления аргументов

В функциональном языке программирования порядок вычисления аргументов не имеет значения. Из-за того что нет ни изменения памяти, ни приостановки вычисления, расчет одного аргумента не влияет на вычисление другого. Objective Caml поддерживает физически изменяемые значения и исключения, поэтому пренебречь порядком вычисления аргументов нельзя. Следующий пример специфичен для Objective Caml 2.04 ОС Linux на платформе Intel:

```
# let new_print_string s = print_string s; String.length s ;;
val new_print_string : string -> int = <fun>
```

```
# (+) (new_print_string "Hello ") (new_print_string "World!") ←
;;
World!Hello - : int = 12
```

По выводу на экран мы видим что вторая строка печатается после первой.

Таков же результат для исключений:

```
# try (failwith "function") (failwith "argument") with Failure ←
    s -> s ;;
- : string = "argument"
```

Если необходимо указать порядок вычисления аргументов, необходимо использовать локальные определения, форсируя таким образом порядок перед вызовом функции. Предыдущий пример может быть переписан следующим способом:

```
# let e1 = (new_print_string "Hello ")
  in let e2 = (new_print_string "World!")
    in (+) e1 e2 ;;
Hello World!- : int = 12
```

В Objective Caml порядок вычисления не указан, на сегодняшний день все реализации caml делают это слева направо. Однако рассчитывать на это может быть рискованно, в случае если в будущем язык будет реализован иначе.

Это вечный сюжет дебатов при концепции языка. Нужно ли специально не указывать некоторые особенности языка и предложить программистам не пользоваться ими, иначе они рискуют получить разные результаты для разных компиляторов. Или же необходимо их указать и, следовательно, разрешить программистам ими пользоваться, что усложнит компилятор и сделает невозможным некоторые оптимизации?

2.5 Калькулятор с памятью

Вернемся к нашему примеру с калькулятором описанным в предыдущей главе и добавим ему пользовательский интерфейс. Теперь мы будем вводить операции напрямую и получать результат сразу без вызова функции перехода (из одного состояния в другое) при каждом нажатии на кнопку.

Добавим 4 новые кнопки: 'C' которая очищает экран, 'M' для сохранения результата в памяти, 'm' для его вызова из памяти и 'OFF' для выключения калькулятора. Что соответствует следующему типу:

```
type key = Plus | Minus | Times | Div | Equals | Digit of int
          | Store | Recall | Clear | Off ;;
```

Теперь определим функцию, переводящую введенные символы в значение типа `key`. Исключение `Invalid_key` отлавливает все символы, которые не соответствуют кнопкам калькулятора. Функция `code` модуля `Char` переводит символ в код ASCII.

```
# exception Invalid_key ;;
exception Invalid_key
# let translation c = match c with
    '+' -> Plus
  | '-' -> Minus
  | '*' -> Times
  | '/' -> Div
  | '=' -> Equals
  | 'C' | 'c' -> Clear
  | 'M' -> Store
  | 'm' -> Recall
  | '0' | 'o' -> Off
  | '0'..'9' as c -> Digit ((Char.code c) - (Char.code '0'))
  | _ -> raise Invalid_key ;;
val translation : char -> key = <fun>
```

В императивном стиле, функция перехода (`transition`) не приведет к новому состоянию, а физически изменит текущее состояние калькулятора. Таким образом необходимо изменить тип `state` так, чтобы его поля были модифицируемые. Наконец, определим исключение `Key_off` для обработки нажатия на кнопку 'OFF'.

```
# type state = {
    mutable lcd : int; (* last computation done *)
    mutable lka : bool; (* last key activated *)
    mutable loa : key; (* last operator activated *)
    mutable vpr : int; (* value printed *)
    mutable mem : int (* memory of calculator *)
} ;;

# exception Key_off ;;
exception Key_off
# let transition s key = match key with
    Clear -> s.vpr <- 0
```

```

| Digit n -> s.vpr <- (if s.lka then s.vpr+10*n else n);
               s.lka <- true
| Store ->    s.lka <- false;
               s.mem <- s.vpr
| Recall ->   s.lka <- false;
               s.vpr <- s.mem
| Off -> raise Key_off
| _ -> let lcd = match s.loa with
               Plus ->   s.lcd + s.vpr
               | Minus -> s.lcd - s.vpr
               | Times -> s.lcd * s.vpr
               | Div ->   s.lcd / s.vpr
               | Equals -> s.vpr
               | _ -> failwith "transition: impossible"
match"
in
  s.lcd <- lcd;
  s.lka <- false;
  s.loa <- key;
  s.vpr <- s.lcd ;;
val transition : state -> key -> unit = <fun>

```

Определим функцию запуска калькулятора `go`, которая возвращает `'()`', так как нас интересует только ее эффект на окружение (ввод/вывод, изменение состояния). Ее аргумент есть константа `'()`', так как наш калькулятор автономен (он сам определяет свое начальное состояние) и интерактивен (данные необходимые для расчета вводятся с клавиатуры по мере необходимости). Переходы реализуются в бесконечном цикле (`while true do`) из которого мы можем выйти при помощи исключения `Key_off`.

```

# let go () =
  let state = { lcd=0; lka=false; loa=Equals; vpr=0; mem=0 }
  in try
    while true do
      try
        let input = translation (input_char stdin)
        in transition state input ;
        print_newline () ;
        print_string "result: " ;
        print_int state.vpr ;

```

```

        print_newline ()
    with
        Invalid_key -> () (* no effect *)
    done
with
    Key_off -> () ;;
val go : unit -> unit = <fun>

```

Заметим, что начальное состояние должно быть либо передано в параметре, либо объявлено локально внутри функции `go`, для того чтобы оно каждый раз инициализировалось при запуске этой функции. Если бы мы использовали значение `initial_state` в функциональном стиле, калькулятор начинал бы работать со старым состоянием, которое он имел перед выключением. Таким образом было бы не просто использовать два калькулятора в одной программе.

2.6 Упражнения

2.6.1 Двусвязный список

В функциональном программировании можно с легкостью манипулировать не циклическими структурами данных, такими как например списки. Однако, циклические структуры доставляют действительные трудности. Здесь мы определим двусвязный список, то есть каждому элементу списка известен предыдущий и последующий элемент:

1. Определить тип для такого списка с параметром, используя хотя бы одну запись с изменяемыми полями.
2. Написать функции `add` и `remove` добавляющие и удаляющие элемент из списка.

2.7 Резюме

В этой главе вы видели применение стилей императивного программирования (физически изменяемые значения, ввод-вывод, структуры итеративного контроля) в функциональном языке. Только **mutable** значения, такие как строки, векторы и записи с изменяемыми полями могут быть физически изменены. Другие значения не могут меняться после их создания. Таким образом мы имеем значения *read-only* для функциональной части и значения *read-write* для императивной.

Нужно отметить, что в случае отказа от использования императивных возможностей языка, это расширение функционального «ядра» языка не меняют его возможностей функционального программирования, за исключением некоторых, легко решаемых проблем с типами.

Глава 3

Функциональный и императивный стиль

Введение

Языки функционального и императивного программирования различаются контролем выполнения программы и управлением памятью.

- функциональная программа вычисляет выражение, в следствие чего мы получаем некоторый результат. Порядок в котором выполнены операции расчета и физическое представление данных не влияют на результат, он одинаков во всех случаях. При таком порядке вычислений, освобождение памяти в Objective Caml неявно осуществляется вызовом автоматического сборщика мусора или Garbage Collector (GC) (см. главу ??).
- императивная программа это список инструкций изменяющих состояние памяти. Каждый этап выполнения строго определен структурами контроля, указывающими на следующую инструкцию. Такие программы чаще всего манипулируют указателями на значение, чем самими значениями. Отсюда необходимость явного выделения и освобождения памяти, что может порой приводить к ошибкам доступа к памяти, однако ничто не запрещает использование GC.

Императивные языки предоставляют больший контроль над выполнением программы и памятью. Находясь ближе к реальной машине, такой код может быть эффективнее, но при этом код теряет в устойчивости выполнения. Функциональное программирование предоставляет более вы-

сокий уровень абстракции и таким образом лучший уровень устойчивости выполнения: типизация (динамическая или статическая) помогает избежать некорректных значений, автоматическая сборка мусора, хотя и замедляет скорость выполнения, обеспечивает правильное манипулирование областями памяти.

Исторически обе эти парадигмы программирования существовали в разных сферах: символьные программы для первого случая и числовые для второго. Однако, с тех времен кое-что изменилось, в частности появилась техника компиляции функциональных языков и выросла эффективность GC. С другой стороны, устойчивость выполнения стала важным критерием, иногда даже важнейшим критерием качества программного обеспечения. В подтверждение этому «коммерческий аргумент» языка Java: эффективность не должна преобладать над правильностью, оставаясь при этом благоразумно хорошей. Эта идея приобретает с каждым днем новых сторонников в мире производителей программного обеспечения.

Objective Caml придерживается этой позиции: он объединяет обе парадигмы, расширяя таким образом область своего применения и облегчая написание алгоритмов в том или ином стиле. Он сохраняет, однако, хорошие свойства правильности выполнения благодаря статической типизации, GC и механизму исключений. Исключения — это первая структура контроля выполнения позволяющая приостановить или продолжить расчет при возникновении определенных условий. Эта особенность находится на границе двух стилей программирования, хоть она и не изменяет результат, но может изменить порядок вычислений. Введение физически изменяемых значений может повлиять на чисто функциональную часть языка. Порядок вычисления аргументов функции становится определяемым если это вычисление производит побочный эффект. По этим причинам подобные языки называются «не чистые функциональные языки». Мы теряем часть абстракции, так как программист должен учитывать модель памяти и выполнение программы. Это не всегда плохо, в частности для читаемости кода. Однако, императивные особенности изменяют систему типов языка: некоторые функциональные программы, с теоретически правильными типами, не являются правильными на практике из-за введения ссылок (reference). Хотя такие программы могут быть легко переписаны.

План главы

В этой главе мы сравним функциональную и императивную модель Objective Caml по критерию контроля выполнения программы и представления значений в памяти. Смесь обоих стилей позволяет конструировать новые структуры данных. Это будет рассмотрено в первом разделе. Во втором разделе мы обсудим выбор между композицией функций (composition of functions) или последовательности (sequencing) с одной стороны и разделением (sharing) или копированием значений с другой. Третий раздел выявляет интерес смешивание двух стилей для создания функциональных изменяемых данных (mutable functional data), что позволит создавать не полностью вычисленные (evaluated) данные. В четвертом разделе рассмотрены *streams*, потенциально бесконечные потоки данных и их интеграция, посредством сопоставления с образцом.

3.1 Сравнение между функциональным и императивным стилями

Воспользуемся строками (`string`) и списками (`'a list`) для иллюстрации разницы между двумя стилями.

3.1.1 С функциональной стороны

`map` это одна из классических функций в среде функциональных языков (смотрим страницу 25). В чистом функциональном стиле она пишется так:

```
# let rec map f l = match l with
  []    -> []
  | h::q -> (f h) :: (map f q) ;;
val map : ('a -> 'b) -> 'a list -> 'b list = <fun>
```

Конечный список состоит из результатов применения функции `f` к элементам списка переданного в аргументе. Он рекурсивно создается указывая начальный элемент (голову) `(f t)` и последующую часть (хвост) `(map f q)`. В частности, программа не указывает какой из них будет вычислен первым.

При этом, для написания такой функции программисту не нужно знать физическое представление данных. Проблемы выделения памяти и разделения данных решаются Objective Caml без внешнего вмешательства программиста. Следующий пример иллюстрирует это:

```
# let example = [ "one" ; "two" ; "three" ] ;;
val example : string list = ["one"; "two"; "three"]
# let result = map (function x -> x) example ;;
val result : string list = ["one"; "two"; "three"]
```

Оба списка `example` и `result` содержат одинаковые значения:

```
# example = result ;;
- : bool = true
```

Оба значения имеют одинаковую структуру, хоть они и представлены в памяти по разному. Убедимся в этом при помощи проверки на физическое равенство:

```
# example == result ;;
- : bool = false
# (List.tl example) == (List.tl result) ;;
- : bool = false
```

3.1.2 Императивная сторона

Вернемся к предыдущему примеру и изменим строку списка `result`.

```
# (List.hd result).[1] <- 's' ;;
- : unit = ()
# result ;;
- : string list = ["ose"; "two"; "three"]
# example ;;
- : string list = ["ose"; "two"; "three"]
```

Определено, изменив список `result`, мы изменили список `example`. То есть знание физической структуры необходимо как только мы пользуемся императивными особенностями.

Рассмотрим теперь как порядок вычисления аргументов функции может стать ловушкой в императивном программировании. Определим структуру изменяемого списка, а так же функции создания, добавления и доступа:

```
# type 'a ilist = { mutable c : 'a list } ;;
type 'a ilist = { mutable c : 'a list; }
# let icreate () = { c = [] }
  let iempty l = (l.c = [])
```

```

let icons x y = y.c <- x::y.c ; y
let ihd x = List.hd x.c
let itl x = x.c <- List.tl x.c ; x ;;
val icreate : unit -> 'a ilist = <fun>
val iempty : 'a ilist -> bool = <fun>
val icons : 'a -> 'a ilist -> 'a ilist = <fun>
val ihd : 'a ilist -> 'a = <fun>
val itl : 'a ilist -> 'a ilist = <fun>
# let rec imap f l =
    if iempty l then icreate()
    else icons (f (ihd l)) (imap f (itl l)) ;;
val imap : ('a -> 'b) -> 'a ilist -> 'b ilist = <fun>

```

Несмотря на то что мы переняли общую структуру функции `map` предыдущего параграфа, с `imap` мы получим другой результат:

```

# let example = icons "one" (icons "two" (icons "three" ←
    (icreate())))) ;;
val example : string ilist = {c = ["one"; "two"; "three"]}
# imap (function x -> x) example ;;
Exception: Failure "hd".

```

В чем тут дело? Вычисление `(itl l)` произошло раньше чем `(ihd l)` и в последней итерации `imap`, список `l` пустой в момент обращения к его заголовку. Список `example` теперь пуст, хоть мы и не получили никакого результата:

```

# example ;;
- : string ilist = c = []

```

Проблема функции `imap` в недостаточном контроле смеси стилей программирования: мы предоставили системе выбор порядка вычисления. Переформулируем функцию `imap` явно указав порядок при помощи конструкции `let .. in ..`.

```

# let rec imap f l =
    if iempty l then icreate()
    else let h = ihd l in icons (f h) (imap f (itl l)) ;;
val imap : ('a -> 'b) -> 'a ilist -> 'b ilist = <fun>
# let example = icons "one" (icons "two" (icons "three" ←
    (icreate())))) ;;
val example : string ilist = {c = ["one"; "two"; "three"]}
# imap (function x -> x) example ;;
- : string ilist = {c = ["one"; "two"; "three"]}

```

Однако, начальный список опять утерян:

```
# example ;;
- : string ilist = c = []
```

Использование оператора последовательности (sequencing operator) и цикла есть другой способ явного указания порядка.

```
# let imap f l =
  let l_res = icreate()
  in while not (iempty l) do
    ignore (icons (f (ihd l)) l_res) ;
    ignore (itl l)
  done ;
  { l_res with c = List.rev l_res.c } ;;
val imap : ('a -> 'b) -> 'a ilist -> 'b ilist = <fun>
# let example = icons "one" (icons "two" (icons "three" ◀
  (icreate())))) ;;
val example : string ilist = {c = ["one"; "two"; "three"]}
# imap (function x -> x) example ;;
- : string ilist = {c = ["one"; "two"; "three"]}
```

Присутствие *ignore* — это факт того что нас интересует побочный эффект функций над ее аргументами, а не их результат. Так же было необходимо выстроить в правильном порядке элементы результата (функцией *List.rev*).

3.1.3 Рекурсия или итерация

Часто ошибочно ассоциируют рекурсию с функциональным стилем и императивный с итерацией. Чисто функциональная программа не может быть итеративной, так как значение условия цикла не меняется никогда. Тогда как императивная программа может быть рекурсивной: функция *imap* тому пример.

Значения аргументов функции хранятся во время ее выполнения. Если она вызовет другую функцию, то эта последняя сохранит и свои аргументы. Эти значения хранятся в стеке выполнения. При возврате из функции эти значения извлечены из стека. Зная что пространство памяти ограничено, мы можем дойти до ее предела, используя функцию с очень большой глубиной рекурсии. В подобных случаях Objective Caml возбуждает исключение *Stack_overflow*.

3.1. СРАВНЕНИЕ МЕЖДУ ФУНКЦИОНАЛЬНЫМ И ИМПЕРАТИВНЫМ СТИЛЯМИ 103

```
# let rec succ n = if n=0 then 1 else 1 + succ (n-1) ;;
val succ : int -> int = <fun>
# succ 100000 ;;
Stack overflow during evaluation (looping recursion?).
```

В итеративной версии место занимаемое функцией `succ_iter` в стеке не зависит от величины аргумента.

```
# let succ_iter n =
  let i = ref 0 in
    for j=0 to n do incr i done ;
    !i ;;
val succ_iter : int -> int = <fun>
# succ_iter 100000 ;;
- : int = 100001
```

У следующей версии функции *a priori* аналогичная глубина рекурсии, однако она успешно выполняется с тем же аргументом.

```
# let succ_tr n =
  let rec succ_aux n accu =
    if n = 0 then accu else succ_aux (n-1) (accu+1)
  in
    succ_aux 1 n ;;
val succ_tr : int -> int = <fun>
# succ_tr 100000 ;;
- : int = 100001
```

Данная функция имеет специальную форму рекурсивного вызова, так называемую **хвостовую рекурсию** при которой результат вызова функции будет результатом функции без дополнительных вычислений. Таким образом отпадает необходимость хранить аргументы функции во время вычисления рекурсивного вызова. Если Objective Caml распознал хвостовую рекурсивную функцию, то перед ее вызовом аргументы извлекаются из стека.

Распознавание хвостовой рекурсии существует во многих языках, однако это свойство просто необходимо функциональным языкам, где она очень часто используется.

3.2 Какой стиль выбрать?

Естественно, речь не идет о «священных» либо эстетических для каждого понятиях, *a priori* не существует стиля красивее или лучше чем другой. Однако, в зависимости от решаемой проблемы, один стиль может быть более подходящим или более адаптированным чем другой.

Первое правило — простота. Желаемый алгоритм (будь то в книге или лишь в голове программиста) уже определен в каком-то стиле, вполне естественно его и использовать при реализации.

Второй критерий — эффективность программы. Можно сказать что императивная программа (хорошо написанная) более эффективна чем ее функциональный аналог, но в очень многих случаях разница между ними не достаточно существенна для оправдания сложности кода императивного стиля, там где функциональный был бы естественен. Функция `map` есть хороший пример натурально выраженной в функциональном стиле проблемы и мы ни в чем не выиграем написав ее в императивном стиле.

3.2.1 Последовательность или композиция функций

Мы уже видели, что как только в программе появляются побочные эффекты, необходимо явно указывать порядок выполнения элементов программы. Это может быть сделано двумя способами:

функциональным опираясь на то что Objective Caml строгий язык, то есть аргумент вычислен до того как он будет передан функции. В выражение `(f (g x))` сначала вычисляется `(g x)` и затем передача этого результат функции `f`. В более сложных выражениях, промежуточный результат может быть наименован конструкцией `let in`, но принцип остается тем же: `let aux=(g x) in (f aux)`.

императивный используя последовательность или другую структуру контроля (цикл). В этом случае, результатом будет побочный эффект над памятью, а не значение возвращенное функцией: `aux:=(g x) ; (f!aux)`.

Давайте рассмотрим данную проблему выбора стиля на следующем примере. Рекурсивный алгоритм быстрой сортировки вектора описывается следующим образом:

1. Выбрать опорную точку: выбрать индекс элемента в векторе;

2. Переставить элементы вокруг этой точки: переставить элементы вектора так чтобы значения меньше значения в опорной точке были слева от нее, а большие значения справа;
3. Отсортировать (тем же алгоритмом) два полученных вектора: элементы после опорной точки и до нее.

Сортировка вектора — означает изменение его состояния, поэтому мы должны использовать императивный стиль, как минимум для манипуляции данными.

Начнем с определения функции переставляющей элементы вектора.

```
# let permute_element vec n p =
  let aux = vec.(n) in vec.(n) <- vec.(p) ; vec.(p) <- aux ~
  ;;
val permute_element : 'a array -> int -> int -> unit = <fun>
```

Выбор правильной точки опоры важен для эффективности алгоритма, но мы ограничимся самым простым способом: вернем индекс первого элемента вектора.

```
# let choose_pivot vec start finish = start ;;
val choose_pivot : 'a -> 'b -> 'c -> 'b = <fun>
```

Напишем теперь желаемый алгоритм переставляющий элементы вектора вокруг выбранной точки.

- Установить опорную точку в начало вектора;
- *i* индекс второго элемента вектора;
- *j* индекс последнего элемента вектора;
- Если элемент с индексом *j* больше чем значение в опорной точке, поменяем местами их значения и увеличим *i* на единицу, иначе уменьшим *j* на единицу;
- До тех пор пока *i* меньше чем *j* повторить предыдущую операцию;
- К этому этапу каждый элемент с индексом меньшим чем *i* (или *j*) меньше значения в опорной точке, а остальные элементы больше: если элемент с индексом *i* меньше чем опорное значение мы меняем их местами, иначе меняем с предыдущим элементом.

При реализации алгоритма мы естественно воспользуемся императивными управляющими структурами.

```
# let permute_pivot vec start finish ind_pivot =
  permute_element vec start ind_pivot ;
  let i = ref (start+1) and j = ref finish and pivot = ◀
  vec.(start) in
  while !i < !j do
    if vec.(!j) >= pivot then decr j
    else
      begin
        permute_element vec !i !j ;
        incr i
      end
    done ;
    if vec.(!i) > pivot then decr i ;
    permute_element vec start !i ;
    !i
  ;;
val permute_pivot : 'a array -> int -> int -> int -> int = ◀
  <fun>
```

Кроме эффекта над вектором, функция возвращает индекс опорной точки.

Нам остается лишь собрать воедино различные этапы и написать рекурсивный вызов для подвекторов.

```
# let rec quick vec start finish =
  if start < finish
  then
    let pivot = choose_pivot vec start finish in
    let place_pivot = permute_pivot vec start finish pivot ◀
    in
    quick (quick vec start (place_pivot-1)) (place_pivot+1) ◀
    finish
  else vec ;;
val quick : 'a array -> int -> int -> 'a array = <fun>
```

Здесь мы воспользовались двумя стилями. Выбранная опорная точка служит аргументом при перестановке вокруг нее и ее порядок в векторе после этой процедуры служит аргументом рекурсивного вызова. Зато,

полученный после перестановки вектор мы получаем в результате побочного эффекта, а не как возвращенное значение функции `permute_pivot`. Тогда как функция `quick` возвращает вектор и сортировка подвекторов реализуется композицией рекурсивных вызовов.

Теперь главная функция:

```
# let quicksort vec = quick vec 0 ((Array.length vec)-1) ;;
val quicksort : 'a array -> 'a array = <fun>
```

Это полиморфная функция так как порядковое отношение '`<`' элементов вектора само полиморфно.

```
# let t1 = [|4;8;1;12;7;3;1;9|] ;;
val t1 : int array = [|4; 8; 1; 12; 7; 3; 1; 9|]
# quicksort t1 ;;
- : int array = [|1; 1; 3; 4; 7; 8; 9; 12|]
# t1 ;;
- : int array = [|1; 1; 3; 4; 7; 8; 9; 12|]
# let t2 = [|"the"; "little"; "cat"; "is"; "dead"|] ;;
val t2 : string array = [|"the"; "little"; "cat"; "is"; "dead"|]
# quicksort t2 ;;
- : string array = [|"cat"; "dead"; "is"; "little"; "the"|]
# t2 ;;
- : string array = [|"cat"; "dead"; "is"; "little"; "the"|]
```

3.2.2 Общее использование или копии значений

До тех пор пока наши данные не изменяемые, нет необходимости знать используются они совместно или нет.

```
# let id x = x ;;
val id : 'a -> 'a = <fun>
# let a = [1; 2; 3] ;;
val a : int list = [1; 2; 3]
# let b = id a ;;
val b : int list = [1; 2; 3]
```

Является ли `b` копией `a` или же это один и тот же список не имеет значения, так как по любому это «неосязаемые» значения. Однако, если на место целых чисел, мы поместим изменяемые значения, необходимо будет знать скажется ли изменение одного значения на другое.

Реализация полиморфизма Objective Caml вызывает (causes) копию непосредственных значений (immediate values) и разделение (sharing) структурных значений. Хотя передача аргументов осуществляется копированием, в случае структурных значений передается лишь указатель. Как в случае с функцией `id`.

```
# let a = [|1; 2; 3|] ;;
val a : int array = [|1; 2; 3|]
# let b = id a ;;
val b : int array = [|1; 2; 3|]
# a.(1) <- 4 ;;
- : unit = ()
# a ;;
- : int array = [|1; 4; 3|]
# b ;;
- : int array = [|1; 4; 3|]
```

Здесь мы действительно имеем случай выбора стиля программирования, от которого зависит эффективность представления данных. С одной стороны, использование изменяемых значений позволяет немедленное манипулирование данными (без дополнительного выделения памяти). Однако это вынуждает, в некоторых случаях, делать копии там, где использование неизменяемого значения позволило бы разделение. Проиллюстрируем это двумя способами реализации списков.

```
# type 'a list_immutable = LInil | LIcons of 'a * 'a ◀
    list_immutable ;;
# type 'a list_mutable = LMnil | LMcons of 'a * 'a ◀
    list_mutable ref ;;
```

Фиксированные списки строго эквивалентны спискам Objective Caml, тогда как изменяемые больше в стиле C, где ячейка состоит из значения и ссылки на следующую ячейку.

С фиксированными списками существует единственный способ реализовать конкатенацию и он вынуждает копирование структуры первого списка, тогда как второй список может быть разделен с конечным списком.

```
# let rec concat l1 l2 = match l1 with
    LInil -> l2
  | LIcons (a, l11) -> LIcons (a, (concat l11 l2)) ;;
```

```

val concat : 'a list_immutable -> 'a list_immutable -> 'a ◀
  list_immutable = <fun>
# let li1 = LIcons (1, LIcons (2, LInil))
  and li2 = LIcons (3, LIcons (4, LInil)) ;;
val li1 : int list_immutable = LIcons (1, LIcons (2, LInil))
val li2 : int list_immutable = LIcons (3, LIcons (4, LInil))
# let li3 = concat li1 li2 ;;
val li3 : int list_immutable =
  LIcons (1, LIcons (2, LIcons (3, LIcons (4, LInil))))
# li1 == li3 ;;
- : bool = false
# let tllI l = match l with
  LInil          -> failwith "Liste vide"
  | LIcons (_, x) -> x ;;
val tllI : 'a list_immutable -> 'a list_immutable = <fun>
# tllI (tllI (li3)) == li2 ;;
- : bool = true

```

В этом примере мы видим что первые ячейки `li1` и `li3` различны, тогда как вторая часть `li3` есть именно `li2`.

С изменяемыми списками можно изменить аргументы (функция `concat_share`) или создать новое значение (функция `concat_copy`)

```

# let rec concat_copy l1 l2 = match l1 with
  LMnil -> l2
  | LMcons (x, l11) -> LMcons (x, ref (concat_copy !l11 l2)) ;;
val concat_copy : 'a list_mutable -> 'a list_mutable -> 'a ◀
  list_mutable = <fun>

```

Это решение, `concat_copy`, аналогично предыдущей функции `concat`. Вот второй вариант, с общим использованием.

```

# let concat_share l1 l2 =
  match l1 with
  LMnil -> l2
  | _ -> let rec set_last = function
    LMnil          -> failwith "concat_share : ◀
    impossible case!!"
    | LMcons (_, l) -> if !l=LMnil then l:=l2 ◀
  else set_last !l
  in

```

```

        set_last l1 ;
        l1 ;;
val concat_share : 'a list_mutable -> 'a list_mutable -> 'a ◀
    list_mutable = <fun>

```

Конкатенация с общим использованием не нуждается в выделении памяти (мы не используем конструктор `LMcons`), мы лишь ограничиваемся тем что последняя ячейка первого списка теперь указывает на второй список. При этом, конкатенация способна изменить аргументы переданные функции.

```

# let lm1 = LMcons (1, ref (LMcons (2, ref LMnil)))
  and lm2 = LMcons (3, ref (LMcons (4, ref LMnil))) ;;
val lm1 : int list_mutable =
  LMcons (1, {contents = LMcons (2, {contents = LMnil})})
val lm2 : int list_mutable =
  LMcons (3, {contents = LMcons (4, {contents = LMnil})})
# let lm3 = concat_share lm1 lm2 ;;
val lm3 : int list_mutable =
  LMcons (1,
    {contents =
      LMcons (2,
        {contents = LMcons (3, {contents = LMcons (4, {contents ◀
          = LMnil}})}})})

```

Мы получили ожидаемый результат для `lm3`, однако значение `lm1` изменено.

```

# lm1 ;;
- : int list_mutable =
LMcons (1,
  contents =
    LMcons (2,
      contents = LMcons (3, contents = LMcons (4, contents = LMnil))))

```

Таким образом это может повлиять на оставшуюся часть программы.

3.2.3 Критерии выбора

В чисто функциональной программе побочных эффектов не существует, это свойство лишает нас операций ввода/вывода, исключений и

изменяемых структур данных. Наше определение функционального стиля является менее ограниченным, то есть функция не изменяющая свое глобальное окружение может быть использована в функциональном стиле. Подобная функция может иметь локальные изменяемые значения (и значит следовать императивному стилю), но не должна изменять ни глобальные переменные ни свои аргументы. С внешней стороны, такие функции можно рассматривать как «черный ящик», чье поведение сравнимо с поведением чисто функциональной функции, с разницей что выполнение первой может быть приостановлено возбуждением исключения. В том же духе, изменяемое значение, которое больше не изменяется после инициализации, может быть использовано в функциональном стиле.

С другой стороны, программа написанная в императивном стиле, унаследует следующие достоинства Objective Caml: правильность статической типизации, автоматическое управление памятью, механизм исключений, параметризованный полиморфизм и вывод типов.

Выбор между стилем функциональным и императивным зависит от программного обеспечения которое вы хотите реализовать. Выбор может быть сделан в соответствии со следующими характеристиками:

выбор структур данных: от использования или нет изменяемых структур данных будет зависеть выбор стиля программирования. Действительно, функциональный стиль по своей природе не совместим с изменением значений. Однако, создание и просмотр значения не зависят от ее свойств. Таким образом мы возвращаемся к дискуссии «изменение на месте vs копия» в 3.2.2 на стр. 107, к которой мы вернемся для обсуждения критерия эффективности.

структура данных существует: если в программе необходимо менять изменяемые структуры данных (modify mutable data structures), то императивный стиль является единственно возможным. Однако, если необходимо лишь читать значения, то применение функционального стиля гарантирует целостность данных.

Использование рекурсивных структур данных подразумевает рекурсивные функции, которые могут быть определены используя особенности того или иного стиля программирования. Однако в общем случае бывает проще истолковывать создание значения следуя рекурсивному определению. Этот подход более близок к функциональному стилю, чем повторяющаяся обработка этого значения рекурсией. Функциональный стиль подразумевает определение интератора над структурой данных и этот фактор уменьшает время разработки и исполнения программы.

критерий эффективности: немедленное изменение без сомнения лучше чем создание значения. В случае когда эффективность кода является главенствующим критерием, чаша весов перевешивает в сторону императивного стиля. Однако отметим что совместное использование значений может оказаться нелегкой задачей и в конце концов более дорогостоящей, чем копирование значений с самого начала.

Чистая функциональность имеет определенную цену: частичное применение (*partial application*) и использование функций переданных в виде аргумента другой функции несет более серьезные накладные расходы в процессе выполнения программы, чем явное применение видимой функции. Стоит избегать использования этой функциональной особенности в тех местах, где критерий производительности является решающим.

критерий разработки: более высокий уровень абстракции функционального стиля программирования позволяет более быстрое написание компактного кода, содержащего меньше ошибок чем императивный вариант, который по своей природе является более «многословным». Таким образом, функциональный стиль является выгодным при разработке значительных программ. Независимость функции к своему контексту окружения позволяет разделить код на более маленькие части и тем самым облегчить его проверку и читаемость. Более высокая модульность функционального стиля плюс возможность передавать функции (а значит и обработку) в аргумент другим функциям повышает вероятность повторного использования программ.

Эти несколько замечаний подтверждают тот факт, что часто смешанное использование обоих стилей является рациональным выбором. Функциональный стиль быстрее в разработке и обеспечивает более простую организацию программы, однако части кода где необходимо повысить скорость выполнения лучше написать в императивном стиле.

3.3 Смесь стилей

Как мы уже заметили, язык программирования имеющий функциональные и императивные особенности дает свободу выбора наиболее подходящего стиля для реализации того или иного алгоритма. Конечно, мы можем использовать оба аспекта Objective Caml совместно в одной и той же функции. Именно этим мы сейчас и займемся.

3.3.1 Замыкание и побочные эффекты

Обычно функция с побочным эффектом рассматривается как процедура и она возвращает значение `()` типа `unit`. Однако, иногда бывает полезно произвести побочный эффект внутри функции и вернуть определенное значение. Мы уже использовали подобный коктейль стилей в функции `permute_pivot` в быстрой сортировке.

В следующем примере реализуем генератор символов, который создаст новый символ при каждом вызове функции. Мы используем счетчик, значение которого увеличивается с каждым вызовом.

```
# let c = ref 0 ;;
val c : int ref = {contents = 0}
# let reset_symb = function () -> c:=0 ;;
val reset_symb : unit -> unit = <fun>
# let new_symb = function s -> c:=!c+1 ; s^(string_of_int !c);;
val new_symb : string -> string = <fun>
# new_symb "VAR" ;;
- : string = "VAR1"
# new_symb "VAR" ;;
- : string = "VAR2"
# reset_symb () ;;
- : unit = ()
# new_symb "WAR" ;;
- : string = "WAR1"
# new_symb "WAR" ;;
- : string = "WAR2"
```

А теперь спрячем ссылку на `c` от всей программы следующим образом:

```
# let (reset_s, new_s) =
  let c = ref 0
  in let f1 () = c:=0
     and f2 s = c:=!c+1 ; s^(string_of_int !c)
     in (f1, f2) ;;
val reset_s : unit -> unit = <fun>
val new_s : string -> string = <fun>
```

Таким образом мы объявили пару функций, разделяющие локальную (для декларации) переменную `c`. Использование обеих функций производит тот же результат что и раньше.

```

# new_s "VAR" ;;
- : string = "VAR1"
# new_s "VAR" ;;
- : string = "VAR2"
# reset_s () ;;
- : unit = ()
# new_s "WAR" ;;
- : string = "WAR1"
# new_s "WAR" ;;
- : string = "WAR2"

```

Этот пример иллюстрирует способ представления замыкания. Замыкание можно рассматривать как пару, состоящую из кода (то есть часть **function**) и локальное окружение, содержащее значения свободных переменных замыкания с другой стороны. На диаграмме 3.1 мы можем увидеть представление в памяти замыканий `reset_s` и `new_s`.

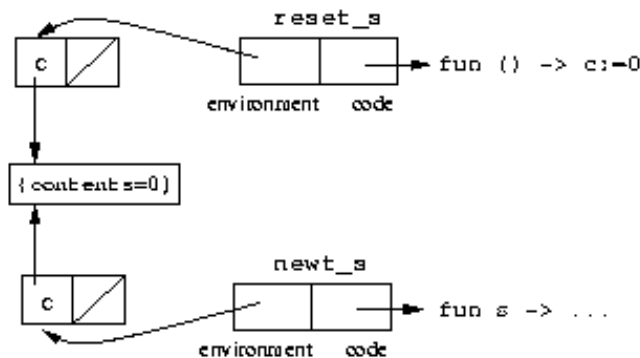


Рис. 3.1: Представление в памяти замыканий

Оба эти замыкания разделяют одно и то же окружение (значение `c`). Когда одно из них меняет ссылку `c`, то оно меняет содержимое памяти разделяемое с другим окружением.

3.3.2 Физическое изменение и исключения

Исключения позволяет отслеживать ситуацию, при которой дальнейшее продолжение программы невозможно. В подобном случае сборщик исключений позволит продолжить вычисление, зная что произошла ошибка. В момент возбуждения исключения может возникнуть проблема побочного эффекта с состоянием модифицируемых данных. Это состояние

не может быть гарантировано если произошли физические изменения в ответвлении неудавшегося вычисления.

Определим функцию увеличения (`++`), с аналогичным результатом что и в C:

```
# let (++) x = x:=!x+1; x ;;
val ( ++ ) : int ref -> int ref = <fun>
```

Следующий пример, иллюстрирует небольшой расчет в котором деление на ноль совпадает с побочным эффектом:

```
# let x = ref 2 ;;
val x : int ref = {contents = 2}
# (* 1 *)
  !((++) x) * (1/0) ;;
Exception: Division_by_zero.
# x ;;
- : int ref = {contents = 2}
# (* 2 *)
  (1/0) * !((++) x) ;;
Exception: Division_by_zero.
# x ;;
- : int ref = {contents = 3}
```

Переменная `x` не изменяется во время вычисления выражения `(*1*)`, тогда как она изменяется во время `(*2*)`. Если заранее не сохранить начальные значения, конструкция `try .. with ..` не должна (в части `with ..`) зависеть от изменяемых переменных, которые участвуют в вычислении возбуждившем исключение.

3.3.3 Изменяемые функциональные структуры данных

В функциональном программировании программа, как функциональное выражение, является в то же время данными — чтобы уяснить этот момент напишем список ассоциированных значений в виде функционального выражения. Этот список ассоциаций (`'a * 'b`) `list` можно рассматривать как частичную функцию из `'a` (множество ключей) в `'b` (множество соответствующих значений). Другими словами, функция `'a -> 'b`.

Пустой список это неопределенная функция, которую мы будем моделировать возбуждением исключения:

```
# let nil_assoc = function x -> raise Not_found ;;
val nil_assoc : 'a -> 'b = <fun>
```

Теперь напомним функцию `add_assoc` добавляющую элемент в список или другими словами расширим функцию новыми значениями.

```
# let add_assoc (k, v) l =
    function x -> if x = k then v else l x ;;
val add_assoc : 'a * 'b -> ('a -> 'b) -> 'a -> 'b = <fun>
# let l = add_assoc ('1', 1) (add_assoc ('2', 2) nil_assoc) ;;
val l : char -> int = <fun>
# l '2' ;;
- : int = 2
# l 'x' ;;
Exception: Not_found.
```

Перепишем функцию `mem_assoc`:

```
# let mem_assoc k l = try (l k) ; true
    with Not_found -> false ;;
val mem_assoc : 'a -> ('a -> 'b) -> bool = <fun>
# mem_assoc '2' l ;;
- : bool = true
# mem_assoc 'x' l ;;
- : bool = false
```

Однако, написание функции удаляющей элементы из списка, не совсем простое дело. У нас больше нет доступа к значениям входящим в замыкание. Для этого мы спрячем старое значение возбуждая исключения `Not_found`.

```
# let rem_assoc k l =
    function x -> if x = k then raise Not_found
        else l x ;;
val rem_assoc : 'a -> ('a -> 'b) -> 'a -> 'b = <fun>
# let l = rem_assoc '2' l ;;
val l : char -> int = <fun>
# l '2' ;;
Exception: Not_found.
```

Естественно, можно воспользоваться ссылками и побочным эффектом для использования таких значений, однако существует несколько предостережений.

```
# let add_assoc_again (k, v) l =
  l := (function x -> if x = k then v else !l x) ;;
val add_assoc_again : 'a * 'b -> ('a -> 'b) ref -> unit = <fun>
```

Мы получили функцию `l` которая указывает сама на саму себя и значит будет зацикливаться. Этот досадный побочный эффект есть результат того что разыменованное `!l` находится внутри замыкания **function** `x ->`. Значение `!l` вычислено при выполнении, а не компиляции. В этот момент, `l` указывает на значение измененное `add_assoc`. Необходимо исправить наше определение используя замыкание полученное определением `add_assoc`:

```
# let add_assoc_again (k, v) l =
  l := add_assoc (k, v) !l ;;
val add_assoc_again : 'a * 'b -> ('a -> 'b) ref -> unit = <fun>
# let l = ref nil_assoc ;;
val l : ('a -> 'b) ref = {contents = <fun>}
# add_assoc_again ('1', 1) l ;;
- : unit = ()
# add_assoc_again ('2', 2) l ;;
- : unit = ()
# !l '1' ;;
- : int = 1
# !l 'x' ;;
Exception: Not_found.
```

3.3.4 Ленивые изменяемые структуры

Смесь императивных особенностей с функциональным языком образует хорошие средства реализации языков программирования. В данном параграфе мы проиллюстрируем эту особенность в реализации структур данных с отсроченным вычислением. Такая структура данных не вычисляется полностью, вычисление продвигается в зависимости от использования структуры.

Отложенное вычисление, часто используемое в чистых функциональных языках, можно моделировать использованием функциональных значений (возможно, изменяемых). Выгода от использования данных с отложенным выполнением двойная; во первых вычисляется лишь то что необходимо для расчета, во вторых использование потенциально бесконечных данных.

Определим тип `vm` элементы которого либо уже вычисленное значение (конструктор `Imm`) либо значение которое будет вычислено (конструктор `Deferred`).

```
# type 'a v =
    Imm of 'a
  | Deferred of (unit -> 'a) ;;
# type 'a vm = { mutable c : 'a v } ;;
```

Откладывание вычислений может быть получено инкапсуляцией в замыкание. Функция вычисляющая такое значение должна или вернуть значение если оно уже вычислено или, в противном случае, вычислить и сохранить результат.

```
# let eval e = match e.c with
    Imm a -> a
  | Deferred f -> let u = f () in e.c <- Imm u ; u ;;
val eval : 'a vm -> 'a = <fun>
```

Операции задержки и активации вычисления также называют замораживанием и размораживанием значения.

Напишем условный контроль в виде функции:

```
# let if_deferred c e1 e2 =
    if eval c then eval e1 else eval e2 ;;
val if_deferred : bool vm -> 'a vm -> 'a vm -> 'a = <fun>
```

А теперь воспользуемся этим в рекурсивной функции подсчета факториала.

```
# let rec facr n =
    if_deferred
      {c=Deferred(fun () -> n = 0)}
      {c=Deferred(fun () -> 1)}
      {c=Deferred(fun () -> n*(facr(n-1)))} ;;
val facr : int -> int = <fun>
# facr 5 ;;
- : int = 120
```

Заметим, что классический `if` не может быть записан в виде функции. Действительно, определим функцию `if_function` следующим образом:

```
# let if_function c e1 e2 = if c then e1 else e2 ;;
val if_function : bool -> 'a -> 'a -> 'a = <fun>
```

Дело в том что все три аргумента вычисляются, это приводит к за-
цикливанию, так как рекурсивный вызов `fact(n-1)` всегда вычисляется,
даже в случае когда `n=0`.

```
# let rec fact n = if_function (n=0) 1 (n*fact(n-1)) ;;
val fact : int -> int = <fun>
# fact 5 ;;
Stack overflow during evaluation (looping recursion?).
```

Модуль Lazy

Трудности во внедрении замороженных значений происходят от кон-
струкции выражений с отложенным вычислением в контексте немедлен-
ного вычисления Objective Caml. Мы это увидели при попытке переоп-
ределить условное выражение. Нельзя написать функцию заморажива-
ющую значение при конструкции объекта типа `vm`.

```
# let freeze e = { c = Deferred (fun () -> e) } ;;
val freeze : 'a -> 'a vm = <fun>
```

Эта функция следует стратегии вычисления Objective Caml, то есть
выражение `e` вычисляется перед тем как создать замыкание `fun () -> e`.
Проиллюстрируем это в следующем примере:

```
# freeze (print_string "trace"; print_newline (); 4*5) ;;
trace
- : int vm = {c = Deferred <fun>}
```

По этой причине была введена следующая синтаксическая форма:

Синтаксис

<code>lazy expr</code>

Внимание

Эта синтаксическая конструкция может изменяться в следующих версиях.

Когда к выражению применяется ключевое слово **lazy** то создается
значение особого типа, который определен в модуле `Lazy`:

```
# let x = lazy (print_string "Helo"; 3*4) ;;
val x : int lazy_t = <lazy>
```

Выражение `(print_string "Hello")` не вычислено, так как не было никакого вывода на экран. При помощи функции *Lazy.force* мы можем вынудить вычисление выражения.

```
# Lazy.force x ;;
Helo- : int = 12
```

Тут мы замечаем, что значение `x` изменилось:

```
# x ;;
- : int lazy_t = lazy 12
```

Теперь это значение замороженного выражения, в данном случае `'12'`. Новый вызов функции *force* просто возвращает вычисленное значение:

```
# Lazy.force x ;;
- : int = 12
```

Строка `'Hello'` больше не выводится на экран.

«Бесконечные» структуры данных

Другой интерес использования отложенного вычисления состоит в возможности построения потенциально бесконечных структур данных, как на пример множество натуральных чисел. Вместо того чтобы конструировать каждое число, мы определим лишь первый элемент и способ получения следующего.

Определим настраиваемую (generic) структуру `'a enum` с помощью которой мы будем определять элементы множества.

```
# type 'a enum = { mutable i : 'a; f : 'a -> 'a } ;;
type 'a enum = { mutable i : 'a; f : 'a -> 'a; }
# let next e = let x = e.i in e.i <- (e.f e.i) ; x ;;
val next : 'a enum -> 'a = <fun>
```

Для того, чтобы получить множество натуральных чисел достаточно конкретизировать (instanciating) поля этой структуры.


```
# let nat = { i=0; f = fun x -> x + 1 } ;;
val nat : int enum = {i = 0; f = <fun>}
# next nat ;;
- : int = 0
# next nat ;;
- : int = 1
# next nat ;;
- : int = 2
```

Другой пример — ряд чисел Фибоначчи, который определен как:

$$\begin{cases} u_0 = 1 \\ u_1 = 1 \\ u_{n+2} = u_n + u_{n+1} \end{cases}$$

Функция, вычисляющая текущее значение, должна использовать значения u_{n-1} и u_{n-2} . Для этого воспользуемся состоянием s в следующем замыкании.

```
# let fib =
  let fx =
    let c = ref 0
    in fun v -> let r = !c + v
                 in c := v ; r
  in { i=1 ; f=fx } ;;
val fib : int enum = {i = 1; f = <fun>}
# for i=0 to 10 do
  print_int (next fib); print_string " " " done ;;
1 1 2 3 5 8 13 21 34 55 89 - : unit = ()
```

3.4 Поток данных

Потоки это потенциально бесконечная последовательность данных определенного рода. Вычисление части потока выполняется по необходимости для текущего вычисления — ленивые структуры данных.

stream абстрактный тип данных, реализация которого неизвестна. Мы манипулируем объектами этого типа при помощи функций конструкции и деструкции. Для удобства пользователя, Objective Caml предоставляет пользователю простые синтаксические конструкции для создания и доступа к элементам потока.

Внимание

Потоки это расширение языка не являющееся частью стабильного ядра Objective Caml.

3.4.1 Конструкция

Упрощенный синтаксис конструкции потоков похож на синтаксис конструкции списков или векторов. Пустой поток создается следующим способом:

```
# [< >] ;;
- : 'a Stream.t = <abstr>
```

Примечание

Для работы с потоками нужно сначала загрузить camlp4 препроцессор.
`# #load "camlp4o.cma";;`

Другой способ конструкции потока состоит в перечислении элементов этого потока, где перед каждым из них ставится апостроф ‘’.

```
# [< '0; '2; '4 >] ;;
- : int Stream.t = <abstr>
```

Выражения, перед которыми не стоит апостроф, рассматриваются как под-потоки.

```
# [< '0; [< '1; '2; '3 >]; '4 >] ;;
- : int Stream.t = <abstr>
# let s1 = [< '1; '2; '3 >] in [< s1; '4 >] ;;
- : int Stream.t = <abstr>
# let concat_stream a b = [< a; b >] ;;
val concat_stream : 'a Stream.t -> 'a Stream.t -> 'a Stream.t ←
    = <fun>
# concat_stream [< "if"; "c"; "then"; "1" >] [< "else"; ←
    "2" >] ;;
- : string Stream.t = <abstr>
```

Остальные функции сгруппированы в модуле `Stream`. На пример, функции `of_channel` и `of_string` возвращают поток состоящий из символов полученных из входного потока или строки символов.

```
# Stream.of_channel ;;
- : in_channel -> char Stream.t = <fun>
# Stream.of_string ;;
- : string -> char Stream.t = <fun>
```

Отложенное вычисление потоков позволяет использовать бесконечные структуры данных, как это было в случае с типом `'a enum` на стр. 120. Определим поток натуральных чисел при помощи первого элемента и функции вычисляющей поток из следующих элементов.

```
# let rec nat_stream n = [< 'n; nat_stream (n+1) >] ;;
val nat_stream : int -> int Stream.t = <fun>
# let nat = nat_stream 0 ;;
val nat : int Stream.t = <abstr>
```

3.4.2 Деструкция и сопоставление потоков

Элементарная операция `next` одновременно вычисляет, возвращает и извлекает первый элемент потока.

```
# for i=0 to 10 do
  print_int (Stream.next nat) ;
  print_string " "
done ;;
0 1 2 3 4 5 6 7 8 9 10 - : unit = ()
# Stream.next nat ;;
- : int = 11
```

Когда данные в потоке закончились возбуждается исключение.

```
# Stream.next [< >] ;;
Exception: Stream.Failure.
```

Для использования потоков Objective Caml предоставляет специальное сопоставление для потоков — **уничтожающее сопоставление** (*destructive matching*). Сопоставляемое значение вычисляется и удаляется из потока. Понятие исчерпываемости сопоставления (*exhaustive match*) не существует для потоков, так как мы используем ленивые и потенциально бесконечные структуры данных. Синтаксис сопоставления следующий:

Синтаксис	<pre>match <i>expr</i> with parser [<i>< 'p₁ ...></i>] -> <i>expr₁</i> ...</pre>
-----------	---

Функция `next` может быть переписана в следующей форме:

```
# let next s = match s with parser [< 'x >] -> x ;;
val next : 'a Stream.t -> 'a = <fun>
# next nat ;;
- : int = 12
```

Заметьте, что перечисление чисел началось с того места где мы остановились.

Существует другая форма синтаксиса для фильтрации функционального параметра типа `Stream.t`.

Синтаксис	<code>parser <i>p</i> -> ...</code>
-----------	--

Перепишем функцию `next` используя новый синтаксис.

```
# let next = parser [< 'x >] -> x ;;
val next : 'a Stream.t -> 'a = <fun>
# next nat ;;
- : int = 13
```

Мы можем сопоставлять пустой поток, но необходимо помнить о следующем: образец потока '`[<>]`' сопоставляется с каким угодно потоком. То есть, поток `s` всегда равен потоку `[< [<>]; s >]`. Поэтому нужно поменять обычный порядок сопоставления.

```
# let rec it_stream f s =
  match s with parser
    [< 'x; ss >] -> f x ; it_stream f ss
  | [< >] -> () ;;
val it_stream : ('a -> 'b) -> 'a Stream.t -> unit = <fun>
# let print_int1 n = print_int n ; print_string " " ;;
val print_int1 : int -> unit = <fun>
# it_stream print_int1 [< '1; '2; '3 >] ;;
1 2 3 - : unit = ()
```

Используя тот факт что сопоставление уничтожающее, перепишем предыдущую функцию.

```
# let rec it_stream f s =
  match s with parser
    [< 'x >] -> f x ; it_stream f s
  | [< >] -> () ;;
val it_stream : ('a -> 'b) -> 'a Stream.t -> unit = <fun>
# it_stream print_int1 [< '1; '2; '3 >] ;;
1 2 3 - : unit = ()
```

Несмотря на то что потоки ленивые, они добровольно и с «восторгом» отдадут свой первый элемент, который после этого будет утерян для потока. Эта особенность отображается на сопоставлении. Следующая функция есть попытка (обреченная на неудачу) вывода на экран двух чисел из потока, в конце потока может остаться один элемент.

```
# let print_int2 n1
  n2 =
    print_string "(" ; print_int n1 ; print_string "," ;
    print_int n2 ; print_string ")" ;;
val print_int2 : int -> int -> unit = <fun>
# let rec print_stream s =
  match s with parser
    [< 'x; 'y >] -> print_int2 x y ; print_stream s
  | [< 'z >] -> print_int1 z ; print_stream s
  | [< >] -> print_newline () ;;
Warning: this match case is unused.
val print_stream : int Stream.t -> unit = <fun>
# print_stream [< '1; '2; '3 >] ;;
(1,2)Exception: Stream.Error "".
```

Два первых элемента потока были выведены на экран без проблем, однако во время вычисления рекурсивного вызова (`print_stream [< 3 >]`) образец обнаружил `x`, который был «употреблен», но для `y` в потоке ничего нет. Это и привело к ошибке. Дело в том что второй образец абсолютно бесполезный, если поток не пустой то первый образец всегда совпадет.

Для того, чтобы получить ожидаемый результат необходимо упорядочить сопоставление.

```
# let rec print_stream s =
  match s with parser
    [< 'x >]
```

```

-> (match s with parser
    [< 'y >] -> print_int2 x y ; print_stream s
  | [< >] -> print_int1 x ; print_stream s)
| [< >] -> print_newline () ;;
val print_stream : int Stream.t -> unit = <fun>
# print_stream [< '1; '2; '3 >] ;;
(1,2)3
- : unit = ()

```

Если сопоставление не срабатывает на первом элементе образца, то фильтр работает как обычно.

```

# let rec print_stream s =
  match s with parser
    [< '1; 'y >] -> print_int2 1 y ; print_stream s
  | [< 'z >] -> print_int1 z ; print_stream s
  | [< >] -> print_newline () ;;
val print_stream : int Stream.t -> unit = <fun>
# print_stream [< '1; '2; '3 >] ;;
(1,2)3
- : unit = ()

```

Пределы сопоставления

Из-за своего свойства уничтожения сопоставление потоков отличается от сопоставления типов сумма. Давайте рассмотрим на сколько глубоко они отличаются.

Вот простейшая функция складывающая два элемента потока.

```

# let rec sum s =
  match s with parser
    [< 'n; ss >] -> n+(sum ss)
  | [< >] -> 0 ;;
val sum : int Stream.t -> int = <fun>
# sum [< '1; '2; '3; '4 >] ;;
- : int = 10

```

Однако, мы можем поглотить поток «изнутри» придав имя полученному результату.

```

# let rec sum s =
  match s with parser

```

```
    [< 'n; r = sum >] -> n+r  
  | [< >] -> 0 ;;  
val sum : int Stream.t -> int = <fun>  
# sum [< '1; '2; '3; '4 >] ;;  
- : int = 10
```

В главе ??, посвященной лексическому и синтаксическому анализу, мы рассмотрим другие примеры использования потоков. В частности, мы увидим как «поглощение» потока изнутри можно с выгодой использовать.

3.5 Резюме

В этой главе мы сравнили функциональный и императивный стили программирования. Основные различия состоят в контроле выполнения (неявный в функциональном и явный в императивном стиле) и представлении в памяти данных (явное разделение или копирование в императивном стиле, не имеющее такой важности в функциональном). Реализация алгоритмов в функциональном и императивном стилях подразумевает эти различия. Выбор между обоими стилями на самом деле приводит к их одновременному использованию. Это позволяет явно выразить представление замыкания, оптимизировать критические части программы и создать изменяемые функциональные данные. Физическое изменение значений в окружении замыкания помогает нам лучше понять что такое функциональное значение. Одновременное использование обоих стилей предоставляет мощные средства для реализации. Мы воспользовались этим при создании потенциально бесконечных данных.