作为Lisp 变体，Scheme 是一门非常简洁的计算语言。本文包括 Scheme 语言的特点、标准与实现、基本概念、数据类型和过程定义

一．Scheme 语言的特点

Scheme 语言是 LISP 语言的一个方言(或说成变种)，它诞生于 1975 年的 MIT，对于这个有近三十年历史的编程语言来说，它并没有象 C++，java，C#那样受到商业领域的青睐，在国内更是显为人知。但它在国外的计算机教育领域内却是有着广泛应用的，有很多人学的第一门计算机语言就是 Scheme 语言。

它是一个小巧而又强大的语言，作为一个多用途的编程语言，它可以作为脚本语言使用，也可以作为应用软件的扩展语言来使用，它具有元语言特性，还有很多独到的特色,以致于它被称为编程语言中的"皇后"。

下面是洪峰对 Scheme 语言的编程特色的归纳：

- 词法定界（Lexical Scoping）
- 动态类型（Dynamic Typing）
- 良好的可扩展性
- 尾递归（Tail Recursive）
- 函数可以作为值返回
- 支持一流的计算连续
- 传值调用（passing-by-value）
- 算术运算相对独立

本文的目的是让有编程基础（那怕是一点点）的朋友能尽快的掌握 Scheme 语言的语法规则，如果您在读完本文后，发现自己已经会用 Scheme 语言了，那么我的目的就达到了。

二．Scheme 语言的标准与实现

R5RS (Revised(5) Report on the Algorithmic Language Scheme)

Scheme 语言的语法规则的第 5 次修正稿，1998 年制定，即 Scheme 语言的现行标准，目前大多数 Scheme 语言的实现都将达到或遵循此标准，并且几乎都加入了一些属于自己的扩展特色。

Guile (GNU's extension language)

Guile 是 GNU 工程的一个项目，它是 GNU 扩展语言库，它也是 Scheme 语言的一个具体实现；如果你将它作为一个库打包，可以把它链接到你的应用程序中去，使你的应用程序具有自己的脚本语言，这个脚本语言目前就是 Scheme 语言。

Guile 可以在 LINUX 和一些 UNIX 系统上运行，下面是简单的安装过程：

下载 guile-1.6.4 版，文件名为 guile-1.6.4.tar.gz，执行下面的命令：

```
tar xvfz guile-1.6.4.tar.gz
cd guile-1.6.4
./configure
make
make install
```

如此，即可以执行命令 guile，进入 guile>提示符状态，输入调试 Scheme 程序代码了，本文的所有代码都是在 guile 下调试通过。

其它实现

除了 Guile 外，Scheme 语言的实现还有很多，如：GNU/MIT-Scheme， SCI，Scheme48，DrScheme 等，它们大多是开源的，可以自由下载安装使用，并且跨平台的实现也很多。你会发现既有象 basic 的 Scheme 语言解释器，也有将 Scheme 语言编译成 C 语言的编译器，也有象 JAVA 那样将 Scheme 语言代码编译成虚拟机代码的编译器。

三．基本概念

注释

Scheme 语言中的注释是单行注释，以分号[;]开始一直到行尾结束，其中间的内容为注释，在程序运行时不做处理，如：

```
   ; this is a scheme comment line.
```

标准的 Scheme 语言定义中没有多行注释，不过在它的实现中几乎都有。在 Guile 中就有多行注释，以符号组合"#!"开始，以相反的另一符号组合"!#"结束，其中内容为注释，如：

```
#!
there are scheme comment area.
you can write mulity lines here .
!#
```

注意的是，符号组合"#!"和"!#"一定分做两行来写。

Scheme 用做脚本语言

Scheme 语言可以象 sh，perl，python 等语言那样作为一种脚本语言来使用，用它来编写可执行脚本，在 Linux 中如果通过 Guile 用 Scheme 语言写可执行脚本，它的第一行和第二行一般是类似下面的内容：

```
#! /usr/local/bin/guile -s
!#
```

这样的话代码在运行时会自动调用 Guile 来解释执行，标准的文件尾缀是 ".scm"。

块(form)

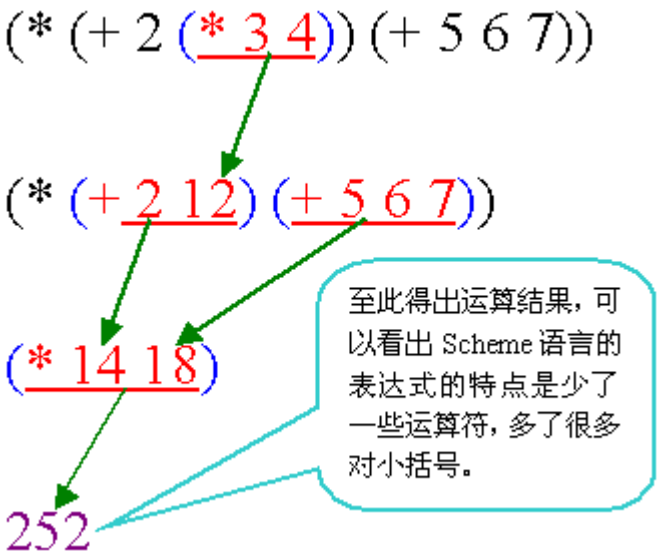
块(form)是 Scheme 语言中的最小程序单元，一个 Scheme 语言程序是由一个或多个 form 构成。没有特殊说明的情况下 form 都由小括号括起来，形如：

```
(define x 123)
(+ 1 2)
(* 4 5 6)
(display "hello world")
```

一个 form 也可以是一个表达式，一个变量定义，也可以是一个过程。

form 嵌套

Scheme 语言中允许 form 的嵌套，这使它可以轻松的实现复杂的表达式，同时也是一种非常有自己特色的表达式。下图示意了嵌套的稍复杂一点的表达式的运算过程：



变量定义

可以用 define 来定义一个变量，形式如下：

(define 变量名 值)

如：（define x 123），定义一个变量 x，其值为 123。

更改变量的值

可以用 set!来改变变量的值，格式如下：
(set! 变量名 值)

如：（set! x "hello"），将变量 x 的值改为"hello"。

Scheme 语言是一种高级语言，和很多高级语言(如 python，perl)一样，它的变量类型不是固定的，可以随时改变。

四．数据类型

1．简单数据类型

逻辑型(boolean)

最基本的数据类型，也是很多计算机语言中都支持的最简单的数据类型，只能取两个值：#t，相当于其它计算机语言中的 TRUE；#f，相当于其它计算机语言中的 FALSE。

Scheme 语言中的 boolean 类型只有一种操作：not。其意为取相反的值，即：

```
(not #f) => #t
(not #t) => #f
```

not 的引用，与逻辑非运算操作类似

```
guile> (not 1)
#f
guile> (not (list 1 2 3))
#f
guile> (not 'a)
#f
```

从上面的操作中可以看出来，只要 not 后面的参数不是逻辑型，其返回值均为#f。

数字型(number)

它又分为四种子类型：整型(integer)，有理数型(rational)，实型(real)，复

数型(complex)；它们又被统一称为数字类型(number)。

如：复数型(complex) 可以定义为 (define c 3+2i)
实数型（real）可以定义为 (define f 22/7)
有理数型（rational）可以定义为 (define p 3.1415)
整数型(integer) 可以定义为 (define i 123)

Scheme 语言中，数字类型的数据还可以按照进制分类，即二进制，八进制，十进制和十六进制，在外观形式上它们分别以符号组合 #b、 #o、 #d、 #x 来作为表示数字进制类型的前缀，其中表示十进制的#d 可以省略不写，如：二进制的 #b1010 ，八进制的 #o567，十进制的 123 或 #d123，十六进制的 #x1afc 。

Scheme 语言的这种严格按照数学定理来为数字类型进行分类的方法可以看出 Scheme 语言里面渗透着很深的数学思想，Scheme 语言是由数学家们创造出来的，在这方面表现得也比较鲜明。

字符型(char)

Scheme 语言中的字符型数据均以符号组合 "#\" 开始，表示单个字符，可以是字母、数字或"[! $ % & * + - . / : %lt; = > ? @ ^ _ ~]"等等其它字符，如：
#\A 表示大写字母 A，#\0 表示字符 0，
其中特殊字符有：#\space 表示空格符和 #\newline 表示换行符。

符号型(symbol)

符号类型是 Scheme 语言中有多种用途的符号名称，它可以是单词，用括号括起来的多个单词，也可以是无意义的字母组合或符号组合，它在某种意义上可以理解为 C 中的枚举类型。看下面的操作：

```
guile> (define a (quote xyz))  ;  定义变量 a 为符号类型，值为 xyz
guile> a
xyz
guile> (define xyz 'a)  ; 定义变量 xyz 为符号类型，值为 a
guile> xyz
a
```

此处也说明单引号' 与 quote 是等价的，并且更简单一些。符号类型与字符串不同的是符号类型不能象字符串那样可以取得长度或改变其中某一成员字符的值，但二者之间可以互相转换。

2. 复合数据类型

可以说复合数据类型是由基本的简单数据类型通过某种方式加以组合形成的数

据类型，特点是可以容纳多种或多个单一的简单数据类型的数据，多数是基于某一种数学模型创建的。

字符串(string) 由多个字符组成的数据类型，可以直接写成由双引号括起的内容，如："hello" 。下面是 Guile 中的字符串定义和相关操作：

```
guile> (define name "tomson")
guile> name
"tomson"
guile> (string-length name)  ; 取字符串的长度
6
guile> (string-set! name 0 #\g)  ; 更改字符串首字母(第 0 个字符)为小写字母 g (#\g)
guile> name
"gomson"
guile> (string-ref name 3)  ; 取得字符串左侧第 3 个字符（从 0 开始）
#\s
```

字符串还可以用下面的形式定义：

```
guile> (define other (string #\h #\e #\l #\l #\o ))
guile> other
"hello"
```

字符串中出现引号时用反斜线加引号代替，如："abc\"def" 。

点对(pair)

我把它译成"点对"，它是一种非常有趣的类型，也是一些其它类型的基础类型，它是由一个点和被它分隔开的两个所值组成的。形如： (1 . 2) 或 (a . b) ，注意的是点的两边有空格。

这是最简单的复合数据类型，同是它也是其它复合数据类型的基础类型，如列表类型（list）就是由它来实现的。

按照 Scheme 语言说明中的惯例，以下我们用符号组合 "=>" 来表示表达式的值。

它用 cons 来定义，如： (cons 8 9) => (8 . 9)

其中在点前面的值被称为 car ，在点后面的值被称为 cdr ，car 和 cdr 同时又成为取 pair 的这两个值的过程，如：

```
(define p (cons 4 5)) => (4 . 5)
(car p)   => 4
```

(cdr p) => 5

还可以用 set-car! 和 set-cdr! 来分别设定这两个值：

```
(set-car! p "hello")
(set-cdr! p "good")
```

如此，以前定义的 p 又变成了（"hello" . "good"）这个样子了。

列表(list)

列表是由多个相同或不同的数据连续组成的数据类型，它是编程中最常用的复合数据类型之一，很多过程操作都与它相关。下面是在 Guile 中列表的定义和相关操作：

```
 guile> (define la (list 1 2 3 4 ))
guile> la
(1 2 3 4)
guile> (length la)  ; 取得列表的长度
4
guile> (list-ref la 3)  ; 取得列表第 3 项的值（从 0 开始）
4
guile> (list-set! la 2 99)  ; 设定列表第 2 项的值为 99
99
guile> la
(1 2 99 4)
guile> (define y (make-list 5 6))  ;创建列表
guile> y
(6 6 6 6 6)
```

make-list 用来创建列表，第一个参数是列表的长度，第二个参数是列表中添充的内容；还可以实现多重列表，即列表的元素也是列表，如：(list (list 1 2 3) (list 4 5 6))。

列表与 pair 的关系

回过头来，我们再看看下面的定义：

```
guile> (define a (cons 1 (cons 2 (cons 3 '()))))
guile> a
(1 2 3)
```

由上可见，a 本来是我们上面定义的点对，最后形成的却是列表。事实上列表是在点对的基础上形成的一种特殊格式。

再看下面的代码：

```
guile> (define ls (list 1 2 3 4))
guile> ls
(1 2 3 4)
guile> (list? ls)
#t
guile> (pair? ls)
#t
```

由此可见，list 是 pair 的子类型，list 一定是一个 pair，而 pair 不是 list。

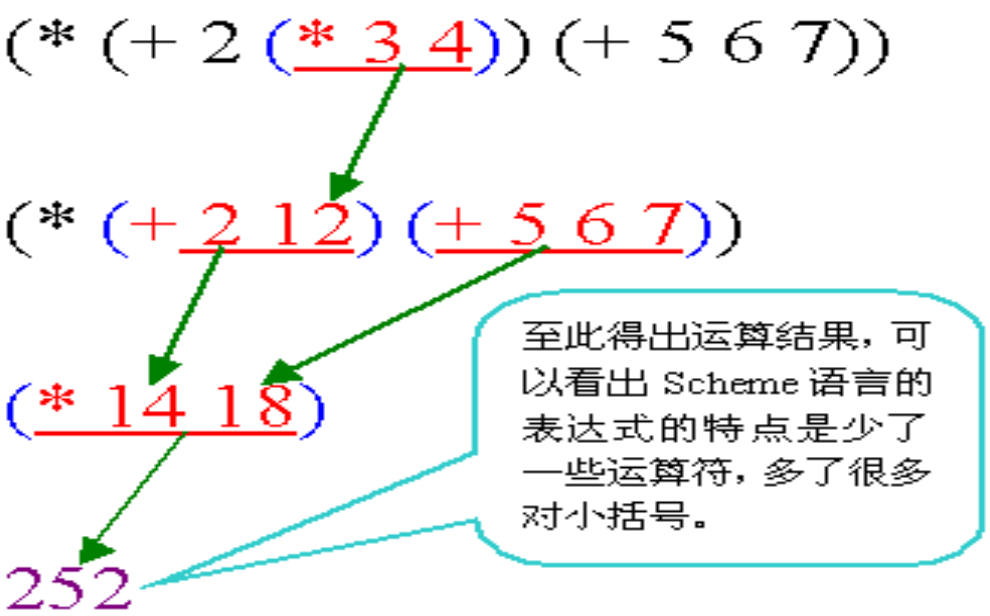
```
guile> (car ls)
1
guile> (cdr ls)
(2 3 4)
```

其 cdr 又是一个列表，可见用于 pair 的操作过程大多可以用于 list。

```
guile> (cadr ls)    ; 此"点对"对象的 cdr 的 car
2
guile> (cddr ls)    ; 此"点对"对象的 cdr 的 cdr
(3 4)
guile> (caddr ls)    ; 此"点对"对象的 cdr 的 cdr 的 car
3
guile> (cdddr ls)    ; 此"点对"对象的 cdr 的 cdr 的 cdr
(4)
```

上在的操作中用到的 cadr，cdddr 等过程是专门对 PAIR 型数据再复合形成的数据操作的过程，最多可以支持在中间加四位 a 或 d，如 cdddr，caaddr 等。

下图表示了由 pairs 定义形成的列表：



这个列表可以由 pair 定义为如下形式：

```
(define x (cons 'a (cons 'b (cons 'c (cons 'd '())))))
```

而列表的实际内容则为：(a b c d)

由 pair 类型还可以看出它可以轻松的表示树型结构，尤其是标准的二叉树。

向量（vector）

可以说是一个非常好用的类型 ，是一种元素按整数来索引的对象，异源的数据结构，在占用空间上比同样元素的列表要少，在外观上：

列表示为： (1 2 3 4)
VECTOR 表示为： #(1 2 3 4)
可以正常定义：(define v (vector 3 4 5))
也可以直接定义：(define v #(3 4 5))

vector 是一种比较常用的复合类型，它的元素索引从 0 开始，至第 n-1 结束，这一点有点类似 C 语言中的数组。

关于向量表（vector）的常用操作过程：

```
guile> (define v (vector 1 2 3 4 5))
guile> v
```

```
#(1 2 3 4 5)
guile> (vector-ref v 0) ; 求第 n 个变量的值
1
guile> (vector-length v)  ; 求 vector 的长度
5
guile> (vector-set! v 2 "abc")  ; 设定 vector 第 n 个元素的值
guile> v
#(1 2 "abc" 4 5)
guile> (define x (make-vector 5 6))  ; 创建向量表
guile> x
#(6 6 6 6 6)
```

make-vector 用来创建一个向量表，第一个参数是数量，后一个参数是添充的值，这和列表中的 make-list 非常相似。

我们可以看出，在 Scheme 语言中，每种数据类型都有一些基本的和它相关的操作过程，如字符串，列表等相关的操作，这些操作过程都很有规律，过程名的单词之间都用-号隔开，很容易理解。对于学过 C++的朋友来说，更类似于某个对象的方法，只不过表现的形式不同了。

3. 类型的判断、比较、运算、转换与方法

类型判断

Scheme 语言中所有判断都是用类型名加问号再加相应的常量或变量构成，形如：

(类型? 变量)

Scheme 语言在类型定义中有比较严格的界定，如在 C 语言等一些语言中数字 0 来代替逻辑类型数据 False，在 Scheme 语言中是不允许的。

以下为常见的类型判断和附加说明：

逻辑型：

```
(boolean? #t) => #t
(boolean? #f) => #t 因为#t 和#f 都是 boolean 类型，所以其值为#t
(boolean? 2)  => #f 因为 2 是数字类型，所以其值为 #f
```

字符型

```
(char? #\space) => #t
(char? #\newline) => #t 以上两个特殊字符：空格和换行
(char? #\f)  => #t 小写字母 f
```

(char? #\;) => #t 分号 ;
(char? #\5) => #t 字符 5 ，以上这些都是正确的，所以返回值都是 #t
(char? 5) => #f 这是数字 5 ，不是字符类型，所以返回 #f

数字型

(integer? 1) => #t
(integer? 2345) => #t
(integer? -90) => #t 以上三个数均为整数
(integer? 8.9) => #f 8.9 不整数
(rational? 22/7) => #t
(rational? 2.3) => #t
(real? 1.2) => #t
(real? 3.14159) => #t
(real? -198.34) => #t 以上三个数均为实数型
(real? 23) => #t 因为整型属于实型
(number? 5) => #t
(number? 2.345) => #t
(number? 22/7) => #t

其它型

(null? '()) => #t ; null 意为空类型，它表示为 '() ，即括号里什么都没有的符号
(null? 5) => #f
(define x 123) 定义变量 x 其值为 123
(symbol? x) => #f
(symbol? 'x) => #t ; 此时 'x 为符号 x，并不表示变量 x 的值

在 Scheme 语言中如此众多的类型判断功能，使得 Scheme 语言有着非常好的自省功能。即在判断过程的参数是否附合过程的要求。

比较运算

Scheme 语言中可以用<，>，<=，>=，= 来判断数字类型值或表达式的关系，如判断变量 x 是否等于零，它的形式是这样的：(= x 0) ，如 x 的值为 0 则表达式的值为#t，否则为#f。

还有下面的操作：

(eqv? 34 34) => #t
(= 34 34) => #t

以上两个 form 功能相同，说明 eqv? 也可以用于数字的判断。

在 Scheme 语言中有三种相等的定义，两个变量正好是同一个对象；两个对象具有相同的值；两个对象具有相同的结构并且结构中的内容相同。除了上面提到的符号判断过程和 eqv?外，还有 eq?和 equal?也是判断是否相等的过程。

eq?，eqv?，equal?

eq?，eqv?和 equal?是三个判断两个参数是否相等的过程，其中 eq?和 eqv?的功能基本是相同的，只在不同的 Scheme 语言中表现不一样。

eq?是判断两个参数是否指向同一个对象，如果是才返回#t；equal?则是判断两个对象是否具有相同的结构并且结构中的内容是否相同，它用 eq?来比较结构中成员的数量；equal?多用来判断点对，列表，向量表，字符串等复合结构数据类型。

```
guile> (define v (vector 3 4 5))
guile> (define w #(3 4 5))  ; w 和 v 都是 vector 类型，具有相同的值#(3 4 5)
guile> (eq? v w)
#f      ; 此时 w 和 v 是两个对象
guile> (equal? v w)
#t                  ; 符合 equal?的判断要求
```

以上操作说明了 eq? 和 equal? 的不同之处，下面的操作更是证明了这一点：

```
guile> (define x (make-vector 5 6))
guile> x
#(6 6 6 6 6)
guile> (eq? x x)     ; 是同一个对象，所以返回#t
#t
guile> (define z (make-vector 5 6))
guile> z
#(6 6 6 6 6)
guile> (eq? x z)     ; 不是同一个对象
#f
guile> (equal? x z)  ; 结构相同，内容相同，所以返回#t
#t
```

算术运算

Scheme 语言中的运算符有：
+ , - , * , / 和 expt (指数运算)
其中 - 和 / 还可以用于单目运算，如：

```
(- 4)  => -4
(/ 4)  => 1/4
```

此外还有许多扩展的库提供了很多有用的过程，

max 求最大 (max 8 89 90 213) => 213
min 求最小 (min 3 4 5 6 7) => 3
abs 求绝对值 (abs -7) ==> 7

除了 max，min，abs 外，还有很多数学运算过程，这要根据你用的 Scheme 语言的运行环境有关，不过它们大多是相同的。在 R5RS 中规定了很多运算过程，在 R5RS 的参考资料中可以很容易找到。

转换

Scheme 语言中用符号组合"->"来标明类型间的转换（很象 C 语言中的指针）的过程，就象用问号来标明类型判断过程一样。下面是一些常见的类型转换过程：

```
guile> (number->string 123)  ; 数字转换为字符串
"123"
guile> (string->number "456")  ; 字符串转换为数字
456
guile> (char->integer #\a)   ;字符转换为整型数，小写字母 a 的 ASCII 码值为 96
97
guile> (char->integer #\A)  ;大写字母 A 的值为 65
65
guile> (integer->char 97)  ;整型数转换为字符
#\a
guile> (string->list "hello")   ;字符串转换为列表
(#\h #\e #\l #\l #\o)
guile> (list->string (make-list 4 #\a)) ; 列表转换为字符串
"aaaa"
guile> (string->symbol "good")  ;字符串转换为符号类型
good
guile> (symbol->string 'better)  ;符号类型转换为字符串
"better"
```

五．过程定义

过程（Procedure）

在 Scheme 语言中，过程相当于 C 语言中的函数，不同的是 Scheme 语言过程是一种数据类型，这也是为什么 Scheme 语言将程序和数据作为同一对象处理的原因。如果我们在 Guile 提示符下输入加号然后回车，会出现下面的情况：

```
guile> +
#<primitive-procedure +>
```

这告诉我们"+"是一个过程，而且是一个原始的过程，即 Scheme 语言中最基础的过程，在 GUILE 中内部已经实现的过程，这和类型判断一样，如 boolean? 等，它们都是 Scheme 语言中最基本的定义。注意：不同的 Scheme 语言实现环境，出现的提示信息可能不尽相同，但意义是一样的。

define 不仅可以定义变量，还可以定义过程，因在 Scheme 语言中过程（或函数）都是一种数据类型，所以都可以通过 define 来定义。不同的是标准的过程定义要使用 lambda 这一关键字来标识。

Lambda 关键字

Scheme 语言中可以用 lambda 来定义过程，其格式如下：
(define 过程名 (lambda (参数 ...) (操作过程 ...)))

我们可以自定义一个简单的过程，如下：

```
(define add5 (lambda (x) (+ x 5)))
```

此过程需要一个参数，其功能为返回此参数加 5 的值，如：

```
(add5 11) => 16
```

下面是简单的求平方过程 square 的定义：

```
(define square (lambda (x)  (* x x)))
```

与 lambda 相同的另一种方式

在 Scheme 语言中，也可以不用 lambda，而直接用 define 来定义过程，它的格式为：
(define (过程名 参数) (过程内容 …))

如下面操作：

```
(define (add6 x) (+ x 6))
add6
#<procedure: add6 (x)> 说明 add6 是一个过程，它有一个参数 x
(add6 23)     => 29
```

再看下面的操作：

```
guile> (define fun
               (lambda(proc x y)
                      (proc x y)))
guile> fun
#<procedure fun (proc x y)>
guile> (fun * 5 6)
30
guile> (fun / 30 3)
10
```

更多的过程定义

上面定义的过程 fun 有三个参数，其中第一个参数 proc 也是一个操作过程（因为在 Scheme 语言中过程也是一种数据，可以作为过程的参数），另外两个参数是数值，所以会出现上面的调用结果。

```
guile> (define add
               (lambda (x y)
                      (+ x y)))
guile> add
#<procedure add (x y)>
guile> (fun add 100 200)
300
```

继续上面操作，我们定义一个过程 add，将 add 作为参数传递给 fun 过程，得出和(fun + 100 200)相同的结果。

```
guile> ((lambda (x) (+ x x)) 5)
10
```

上面的 (lambda(x) (+ x x)) 事实上是简单的过程定义，在后面直接加上操作参数 5，得出结果 10，这样实现了匿名过程，直接用过程定义来操作参数，得出运算结果。

通过上面的操作，相信你已初步了解了过程的用法。既然过程是一种数据类型，所以将过程作为过程的参数是完全可以的。以下过程为判断参数是否为过程，给出一个参数，用 procedure? 来判断参数是否为过程，采用 if 结构（关于 if 结构见下面的介绍）：

```
guile> (define isp
           (lambda (x)
                   (if (procedure? x) 'isaprocedure
```

```
'notaprocedure)))
guile> isp
#<procedure isp (x)>
guile> (isp 0)
notaprocedure
guile> (isp +)
isaprocedure
```

上面的过程就体现了 Scheme 语言的参数自省（辨别）能力，'0'是数字型，所以返回 notaprocedure；而'+'是一个最基础的操作过程，所以返回 isaprocedure。

过程的嵌套定义

在 Scheme 语言中，过程定义也可以嵌套，一般情况下，过程的内部过程定义只有在过程内部才有效，相当 C 语言中的局部变量。

如下面的代码的最终结果是 50：

```
(define fix
 (lambda (x y z)
  (define add
   (lambda (a b) (+ a b)))
  (- x (add y z))))
(display (fix 100 20 30))
```

此时过程 add 只在 fix 过程内部起做用，这事实上涉及了过程和变量的绑定，可以参考下面的关于过程绑定（let，let* 和 letrec）的介绍。

过程是初学者难理解的一个关键，随着过程参数的增加和功能的增强，过程的内容变得越来越复杂，小括号也会更多，如果不写出清晰的代码的话，读代码也会成为一个难题。

熟悉了 scheme 基本概念、数据类型和过程（函数）后， 下一部分我们来学习 scheme 的结构、递归调用和其他扩展功能。

作者简介

宋国伟，目前在吉林省德惠市信息中心从事网络维护工作，著有《GTK+2.0 编程范例》一书，热衷于 Linux 系统上的编程及相关的研究。 可以通过电子邮件地址 gwsong52@sohu.com 与他联系 。