

第六章 运行时存储空间的组织和管理

术语

- 过程的活动(activation)
过程的一次执行称为过程的一次活动
- 活动记录
过程的活动需要可执行代码和存放所需信息的存储空间, 后者称为活动记录

本章内容

- 讨论一个活动记录中的数据布局
- 程序执行过程中, 所有活动记录的组织方式

第六章 运行时存储空间的组织和管理

• 影响存储分配策略的语言特征

- 过程能否递归
- 当控制从过程的活动返回时, 局部变量的值是否要保留
- 过程能否访问非局部变量
- 过程调用的参数传递方式
- 过程能否作为参数被传递
- 过程能否作为结果值传递
- 存储块能否在程序控制下动态地分配
- 存储块是否必须显式地释放

6.1 局部存储分配

6.1.1 过程

语言概念:

过程定义、过程调用、形式参数、实在参数、活动的生存期

6.1 局部存储分配

6.1.2 名字的作用域和绑定

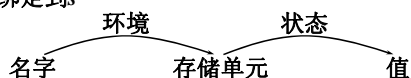
1、名字的作用域(scope)

- 一个声明起作用的程序部分称为该声明的作用域
- 即使一个名字在程序中只声明一次, 该名字在程序运行时也可能表示不同的数据对象

6.1 局部存储分配

2、环境和状态

- 环境把名字映射到左值, 而状态把左值映射到右值 (即名字到值有两步映射)
- 赋值改变状态, 但不改变环境
- 过程调用改变环境
- 如果环境将名字 x 映射到存储单元 s , 则说 x 被绑定到 s



6.1 局部存储分配

3、静态概念和动态概念的对应

静态概念	动态对应
过程的定义	过程的活动

6.1 局部存储分配

3、静态概念和动态概念的对应

静态概念	动态对应
过程的定义	过程的活动
名字的声明	名字的绑定

6.1 局部存储分配

3、静态概念和动态概念的对应

静态概念	动态对应
过程的定义	过程的活动
名字的声明	名字的绑定
声明的作用域	绑定的生存期

6.1 局部存储分配

6.1.3 活动记录(activation record)

活动记录的常见布局

返回
值
参 数
控 制 链
访 问 链
机 器 状 态
局 部 数 据
临 时 数 据

6.1 局部存储分配

6.1.4 局部数据的布局

- 字节是可编址内存的最小单位
- 变量所需的存储空间可以根据其类型而静态确定
- 一个过程所声明的局部变量，按这些变量声明时出现的次序，在局部数据域中依次分配空间
- 局部数据的地址可以用相对于活动记录中某个位置的地址来表示
- 数据对象的存储布局还有一个对齐问题

6.1 局部存储分配

- 例 在SPARC/Solaris工作站上下面两个结构体的size分别是24和16，为什么不一样？

```
typedef struct _a{          typedef struct _b{
    char c1;                char c1;
    long i;                 char c2;
    char c2;                long i;
    double f;               double f;
}a;                          }b;
对齐: char : 1, long : 4, double : 8
```

6.1 局部存储分配

- 例 在SPARC/Solaris工作站上下面两个结构体的size分别是24和16，为什么不一样？

```
typedef struct _a{          typedef struct _b{
    char c1; 0              char c1; 0
    long i;   4             char c2; 1
    char c2;  8             long i;  4
    double f; 16            double f; 8
}a;                          }b;
对齐: char : 1, long : 4, double : 8
```

6.1 局部存储分配

- 例 在X86/Linux机器的结果和SPARC/Solaris工作站不一样，是20和16。

```
typedef struct _a{          typedef struct _b{
    char c1; 0              char c1; 0
    long i;   4              char c2; 1
    char c2;  8              long i;   4
    double f; 12             double f; 8
}a;                          }b;
对齐: char : 1, long : 4, double : 4
```

6.1 局部存储分配

6.1.5 程序块

- 本身含有局部变量声明的语句
- 可以嵌套
- 最接近的嵌套作用域规则
- 并列程序块不会同时活跃
- 并列程序块的变量可以重叠分配

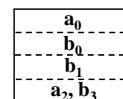
6.1 局部存储分配

```
main() /* 例 */
{
    /* begin of B0 */
    int a = 0;
    int b = 0;
    {
        /* begin of B1 */
        int b = 1;
        {
            /* begin of B2 */
            int a = 2;
        } /* end of B2 */
        {
            /* begin of B3 */
            int b = 3;
        } /* end of B3 */
    } /* end of B1 */
} /* end of B0 */
```

6.1 局部存储分配

```
main() /* 例 */
{ /* begin of B0 */
    int a = 0;
    int b = 0;
    { /* begin of B1 */
        int b = 1;
        { /* begin of B2 */
            int a = 2;
        } /* end of B2 */
        { /* begin of B3 */
            int b = 3;
        } /* end of B3 */
    } /* end of B1 */
} /* end of B0 */
```

声 明	作 用 域
int a = 0;	$B_0 - B_2$
int b = 0;	$B_0 - B_1$
int b = 1;	$B_1 - B_3$
int a = 2;	B_2
int b = 3;	B_3



重叠分配存储单元

6.2 全局栈式存储分配

本节介绍

- 介绍程序运行时所需的各个活动记录在存储空间的分配策略
- 描述过程的目标代码怎样访问绑定到局部名字的存储单元
- 介绍三种分配策略
 - 静态分配策略
 - 栈式分配策略
 - 堆式分配策略

6.2 全局栈式存储分配

6.2.1 运行时内存的划分



6.2 全局栈式存储分配

1、静态分配

- 名字在程序被编译时绑定到存储单元，不需要运行时的任何支持
- 绑定的生存期是程序的整个运行期间

6.2 全局栈式存储分配

2、静态分配给语言带来限制

- 递归过程不被允许
- 数据对象的长度和它在内存中位置的限制，必须是在编译时可以知道的
- 数据结构不能动态建立

6.2 全局栈式存储分配

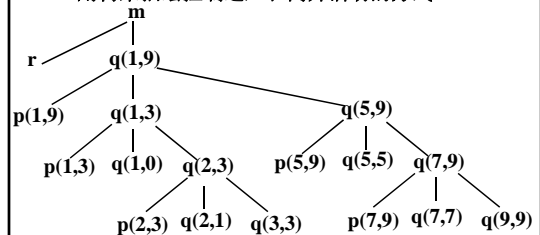
- 例 C程序的外部变量、静态局部变量以及程序中出现的常量都可以静态分配
- 声明在函数外面
 - 外部变量 — 静态分配
 - 静态外部变量 — 静态分配
- 声明在函数里面
 - 静态局部变量 — 也是静态分配
 - 自动变量 — 不能静态分配

6.2 全局栈式存储分配

6.2.2 活动树和运行栈

1、活动树

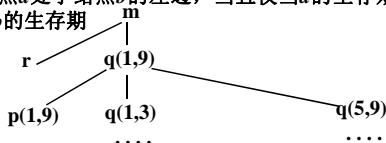
- 用树来描绘控制进入和离开活动的方式



6.2 全局栈式存储分配

• 活动树的特点

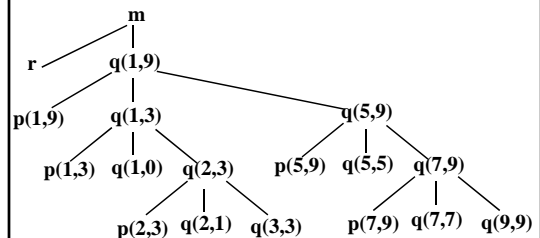
- 每个结点代表某过程的一个活动
- 根结点代表主程序的活动
- 结点 a 是结点 b 的父结点，当且仅当控制流从 a 的活动进入 b 的活动
- 结点 a 处于结点 b 的左边，当且仅当 a 的生存期先于 b 的生存期



6.2 全局栈式存储分配

- 当前活跃着的过程活动可以保存在一个栈中

- 例 控制栈的内容: $m, q(1,9), q(1,3), q(2,3)$

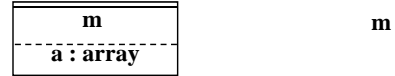


6.2 全局栈式存储分配

- 2、运行栈：把控制栈中的信息拓广到包括过程活动所需的所有局部信息（即活动记录）

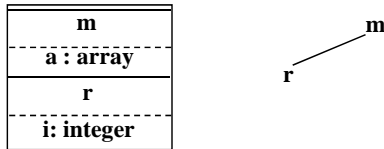
6.2 全局栈式存储分配

- 2、运行栈：把控制栈中的信息拓广到包括过程活动所需的所有局部信息（即活动记录）



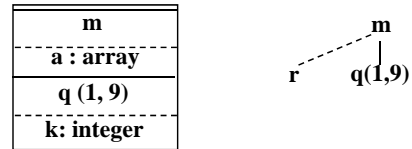
6.2 全局栈式存储分配

- 2、运行栈：把控制栈中的信息拓广到包括过程活动所需的所有局部信息（即活动记录）



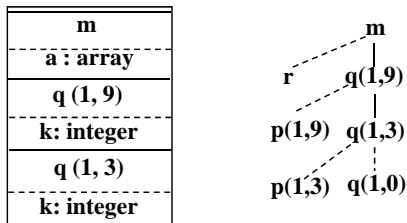
6.2 全局栈式存储分配

- 2、运行栈：把控制栈中的信息拓广到包括过程活动所需的所有局部信息（即活动记录）



6.2 全局栈式存储分配

- 2、运行栈：把控制栈中的信息拓广到包括过程活动所需的所有局部信息（即活动记录）



6.2 全局栈式存储分配

6.2.3 调用序列(calling sequence)

- 过程调用和过程返回都需要执行一些代码来管理活动记录栈，保存或恢复机器状态等
- 过程调用序列
过程调用时执行的分配活动记录，把信息填入它的域中，使被调用过程可以开始执行的代码
- 过程返回序列
被调用过程返回时执行的恢复机器状态，释放被调用过程活动记录，使调用过程能够继续执行的代码
- 调用序列和返回序列常常都分成两部分，分处于调用过程和被调用过程中

6.2 全局栈式存储分配

- 即使是同一种语言，过程调用序列、返回序列和活动记录中各域的排放次序，也会因实现而异

- 设计这些序列和活动记录的一些原则

- 以活动记录中间的某个位置作为基地址
- 长度能较早确定的域放在活动记录的中间

返 回 值
参 数
控 制 链
访 问 链
机 器 状 态
局 部 数 据
临 时 数 据

6.2 全局栈式存储分配

- 即使是同一种语言，过程调用序列、返回序列和活动记录中各域的排放次序，也会因实现而异

- 设计这些序列和活动记录的一些原则

- 一般把临时数据域放在局部数据域的后面
- 把参数域和可能的返回值域放在紧靠调用者活动记录的地方

返 回 值
参 数
控 制 链
访 问 链
机 器 状 态
局 部 数 据
临 时 数 据

6.2 全局栈式存储分配

- 即使是同一种语言，过程调用序列、返回序列和活动记录中各域的排放次序，也会因实现而异

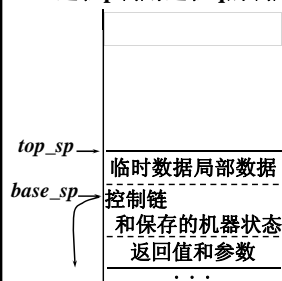
- 设计这些序列和活动记录的一些原则

- 用同样的代码来执行各个活动的保存和恢复

返 回 值
参 数
控 制 链
访 问 链
机 器 状 态
局 部 数 据
临 时 数 据

6.2 全局栈式存储分配

1、过程p调用过程q的调用序列



6.2 全局栈式存储分配

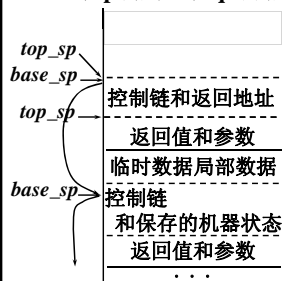
1、过程p调用过程q的调用序列



(1) p计算实参，依次放入栈顶，并在栈顶留出放返回值的空间。top_sp的值在此过程中被改变

6.2 全局栈式存储分配

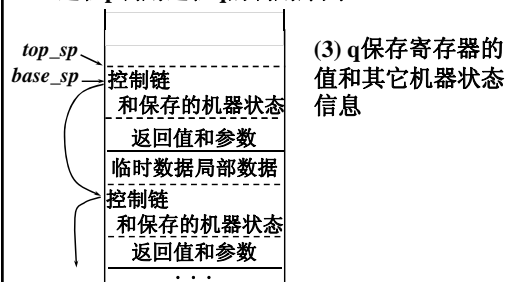
1、过程p调用过程q的调用序列



(2) p把返回地址和当前base_sp的值存入q的活动记录中，建立q的访问链，增加base_sp的值

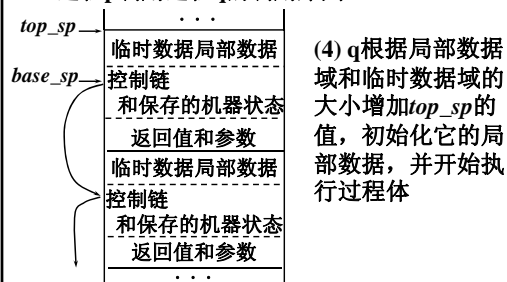
6.2 全局栈式存储分配

1、过程p调用过程q的调用序列



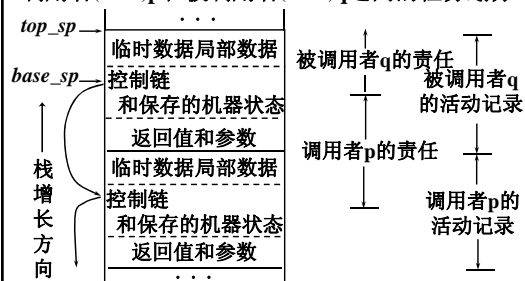
6.2 全局栈式存储分配

1、过程p调用过程q的调用序列



6.2 全局栈式存储分配

调用者(caller)p和被调用者(callee)q之间的任务划分



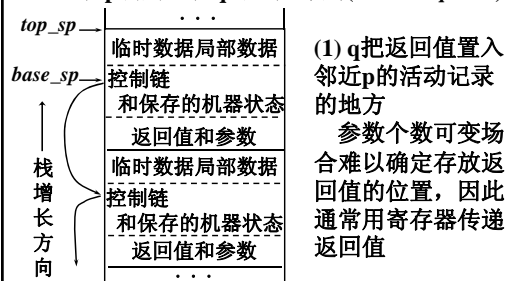
6.2 全局栈式存储分配

2、过程p调用过程q的返回序列



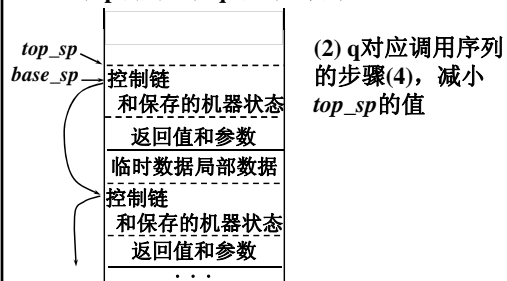
6.2 全局栈式存储分配

2、过程p调用过程q的返回序列(return sequence)



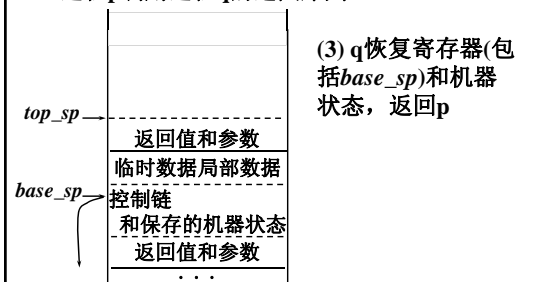
6.2 全局栈式存储分配

2、过程p调用过程q的返回序列



6.2 全局栈式存储分配

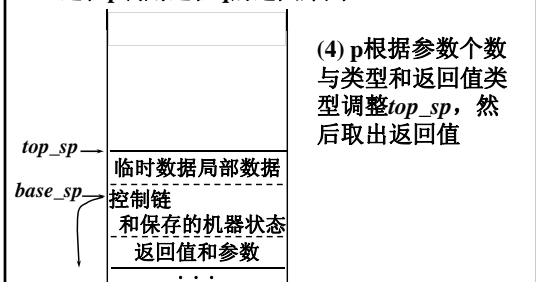
2、过程p调用过程q的返回序列



(3) q恢复寄存器(包括`base_sp`)和机器状态, 返回p

6.2 全局栈式存储分配

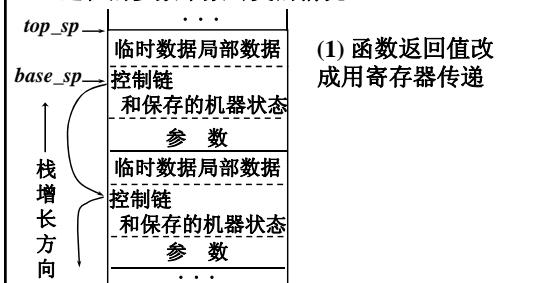
2、过程p调用过程q的返回序列



(4) p根据参数个数与类型和返回值类型调整`top_sp`, 然后取出返回值

6.2 全局栈式存储分配

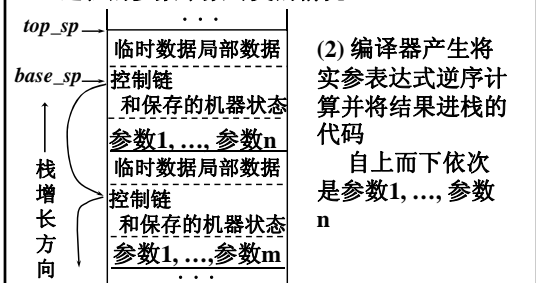
3、过程的参数个数可变的情况



(1) 函数返回值改成用寄存器传递

6.2 全局栈式存储分配

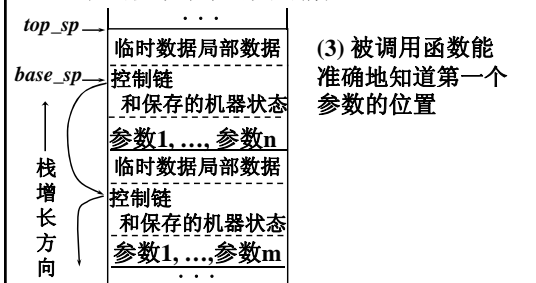
3、过程的参数个数可变的情况



(2) 编译器产生将实参表达式逆序计算并将结果进栈的代码
自上而下依次是参数1, ..., 参数n

6.2 全局栈式存储分配

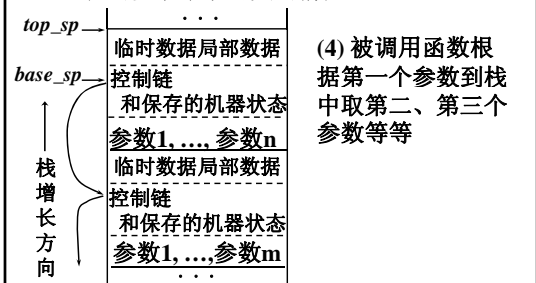
3、过程的参数个数可变的情况



(3) 被调用函数能准确地知道第一个参数的位置

6.2 全局栈式存储分配

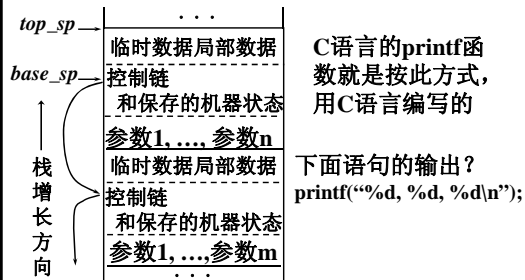
3、过程的参数个数可变的情况



(4) 被调用函数根据第一个参数到栈中取第二、第三个参数等等

6.2 全局栈式存储分配

3、过程的参数个数可变的情况



6.2 全局栈式存储分配

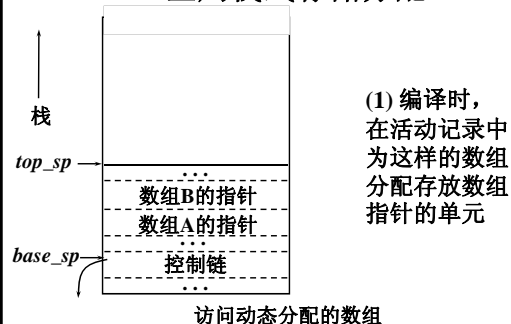
6.2.4 栈上可变长数据(variable-length data)

活动记录的长度在编译时不能确定的情况

- 例: 局部数组的大小要等到过程激活时才能确定

备注: Java语言的实现是将它们分配在堆上

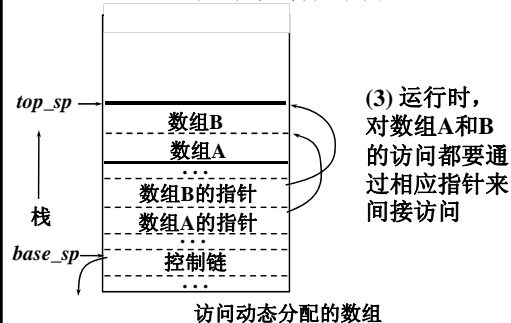
6.2 全局栈式存储分配



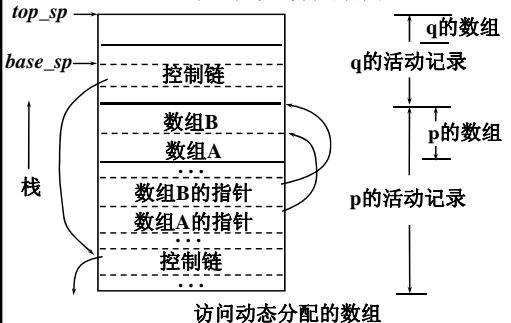
6.2 全局栈式存储分配



6.2 全局栈式存储分配



6.2 全局栈式存储分配



6.2 全局栈式存储分配

6.2.5 悬空引用

悬空引用：引用某个已被释放的存储单元

6.2 全局栈式存储分配

6.2.5 悬空引用

悬空引用：引用某个已被释放的存储单元

例：main中引用p指向的对象

```
main() {           |   int * dangle () {  
    int *q;         |       int j = 20;  
    q = dangle ();  |       return &j;  
}                  |   }
```

6.3 非局部名字的访问

本节介绍

- 无过程嵌套的静态作用域（C语言）
- 有过程嵌套的静态作用域（Pascal语言）
- 动态作用域（Lisp语言）

6.3 非局部名字的访问

6.3.1 无过程嵌套的静态作用域

- 过程体中的非局部引用可以直接使用静态确定的地址（非局部数据此时就是全局数据）
- 局部变量在栈顶的活动记录中，可以通过 *base_sp* 指针来访问
- 无须深入栈中取数据，无须访问链
- 过程可以作为参数来传递，也可以作为结果来返回

6.3 非局部名字的访问

6.3.2 有过程嵌套的静态作用域

图6.14 P177

```
sort  
  readarray  
  exchange  
  quicksort  
    partition
```

6.3 非局部名字的访问

6.3.2 有过程嵌套的静态作用域

- 过程嵌套深度

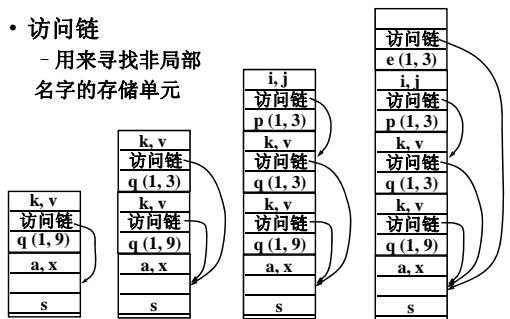
```
sort           1  
  readarray    2  
  exchange     2  
  quicksort    2  
    partition   3
```

- 变量的嵌套深度：它的声明所在过程的嵌套深度作为该名字的嵌套深度

6.3 非局部名字的访问

- 访问链

- 用来寻找非局部名字的存储单元



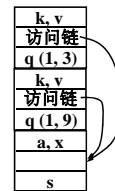
6.3 非局部名字的访问

- 访问非局部名字的存储单元

- 假定过程p的嵌套深度为 n_p ，它引用嵌套深度为 n_a 的变量a， $n_a \leq n_p$ ，如何访问a的存储单元

```

sort          1
readarray     2
exchange      2
quicksort     2
partition     3
    
```



6.3 非局部名字的访问

- 访问非局部名字的存储单元

- 假定过程p的嵌套深度为 n_p ，它引用嵌套深度为 n_a 的变量a， $n_a \leq n_p$ ，如何访问a的存储单元

- 从栈顶的活动记录开始，追踪访问链 $n_p - n_a$ 次
- 到达a的声明所在过程的活动记录
- 访问链的追踪用间接操作就可完成

```

sort          1
readarray     2
exchange      2
quicksort     2
partition     3
    
```

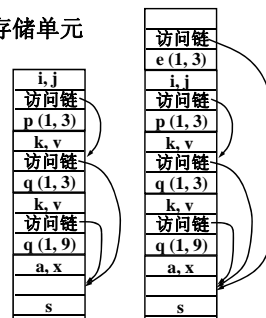
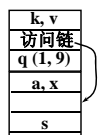


6.3 非局部名字的访问

- 访问非局部名字的存储单元

```

sort
readarray
exchange
quicksort
partition
    
```



6.3 非局部名字的访问

- 过程p对变量a访问时，a的地址由下面的二元组表示：

- $(n_p - n_a, a \text{在活动记录中的偏移})$

6.3 非局部名字的访问

- 建立访问链

- 假定嵌套深度为 n_p 的过程p调用嵌套深度为 n_x 的过程x

(1) $n_p < n_x$ 的情况

```

sort          1
readarray     2
exchange      2
quicksort     2
partition     3
    
```

这时x肯定
就声明在p中

6.3 非局部名字的访问

• 建立访问链

- 假定嵌套深度为 n_p 的过程 p 调用嵌套深度为 n_x 的过程 x

(1) $n_p < n_x$ 的情况

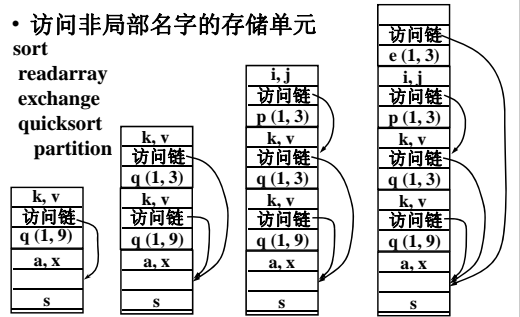
- 被调用过程的访问链必须指向调用过程的活动记录的访问链

解决：诸如在`quicksort(1, 3)`调用`partition(1, 3)`时如何确定后者活动记录中的访问链的问题

6.3 非局部名字的访问

• 访问非局部名字的存储单元

sort
readarray
exchange
quicksort
partition



6.3 非局部名字的访问

• 建立访问链

- 假定嵌套深度为 n_p 的过程 p 调用嵌套深度为 n_x 的过程 x

(2) $n_p \geq n_x$ 的情况

sort	1	这时 p 和 x 的
readarray	2	嵌套深度分别
exchange	2	为1, 2, ...,
quicksort	2	$n_x - 1$ 的外围过
partition	3	程肯定相同

6.3 非局部名字的访问

• 建立访问链

- 假定嵌套深度为 n_p 的过程 p 调用嵌套深度为 n_x 的过程 x

(2) $n_p \geq n_x$ 的情况

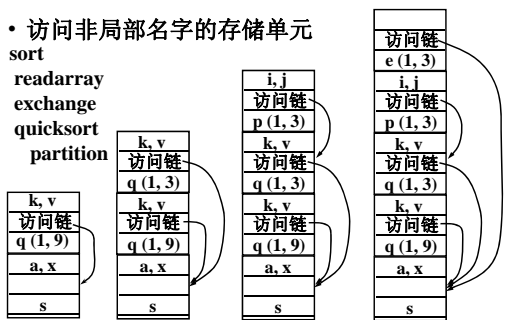
- 追踪访问链 $n_p - n_x + 1$ 次，到达了静态包围 x 和 p 的且离它们最近的那个过程的最新活动记录
- 所到达的活动记录就是 x 的活动记录中的访问链应该指向的那个活动记录

解决：诸如在`partition(1, 3)`调用`exchange(1, 3)`时如何确定后者活动记录中的访问链的问题

6.3 非局部名字的访问

• 访问非局部名字的存储单元

sort
readarray
exchange
quicksort
partition



6.3 非局部名字的访问

program param(input, output); (过程作为参数)

procedure b(function h(n: integer): integer);

begin writeln(h(2)) end {b};

procedure c;

var m: integer;

function f(n: integer): integer;

begin f := m+n end {f};

begin m := 0; b(f) end {c};

begin

c 函数 f 作为参数传递时，怎样

end. 在 f 被激活时建立它的访问链

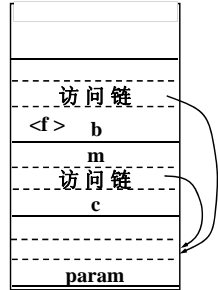
6.3 非局部名字的访问

program param(input, output); (过程作为参数)

```

procedure b(function h(...)
begin writeln(h(2)) end;
procedure c;
var m: integer;
function f(n: integer)...
begin f := m+n end {f};
begin m := 0; b(f) end {c};
begin
c    从b的访问链难以建立
    立f的访问链
end.

```



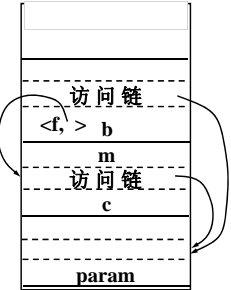
6.3 非局部名字的访问

program param(input, output); (过程作为参数)

```

procedure b(function h(...)
begin writeln(h(2)) end;
procedure c;
var m: integer;
function f(n: integer)...
begin f := m+n end {f};
begin m := 0; b(f) end {c};
begin
c    f作为参数传递时,
    它的起始地址连同它
    的访问链一起传递
end.

```



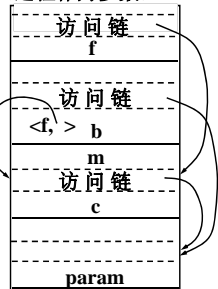
6.3 非局部名字的访问

program param(input, output); (过程作为参数)

```

procedure b(function h(...)
begin writeln(h(2)) end;
procedure c;
var m: integer;
function f(n: integer)...
begin f := m+n end {f};
begin m := 0; b(f) end {c};
begin
c    b调用f时, 用传递
    过来的访问链来建立f
    的访问链
end.

```



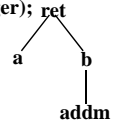
6.3 非局部名字的访问

program ret(input, output); (过程作为返回值)

```

var f: function(integer): integer;
function a: function(integer): integer;
var m: integer;
function addm(n: integer): integer;
begin return m+n end;
begin m := 0; return addm end;
procedure b(g: function(integer): integer); ret
begin writeln(g(2)) end;
begin
f := a; b(f)
end.

```



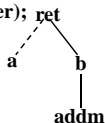
6.3 非局部名字的访问

program ret(input, output); (过程作为返回值)

```

var f: function(integer): integer;
function a: function(integer): integer;
var m: integer;
function addm(n: integer): integer;
begin return m+n end;
begin m := 0; return addm end;
procedure b(g: function(integer): integer); ret
begin writeln(g(2)) end;
begin
f := a; b(f)
end.

```



执行addm时, a
的活动记录已不存在,
取不到m的值

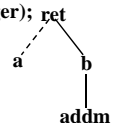
6.3 非局部名字的访问

program ret(input, output); (~~过程作为返回值~~)

```

var f: function(integer): integer;
function a: function(integer): integer;
var m: integer;
function addm(n: integer): integer;
begin return m+n end;
begin m := 0; return addm end;
procedure b(g: function(integer): integer); ret
begin writeln(g(2)) end;
begin
f := a; b(f)
end.

```



执行addm时, a
的活动记录已不存在,
取不到m的值

6.3 非局部名字的访问

- C语言的函数声明不能嵌套，函数不论在什么情况下激活，要访问的数据分成两种情况
 - 非静态局部变量（包括形式参数），它们分配在活动记录栈顶的那个活动记录中
 - 外部变量（包括定义在其它源文件之中的外部变量）和静态的局部变量，它们都分配在静态数据区
 - 因此C语言允许函数（的指针）作为返回值

6.3 非局部名字的访问

6.3.3 动态作用域

- 被调用过程的非局部名字 a 和它在调用过程中引用的是同样的存储单元
 - 基于运行时的调用关系
 - 而不是基于静态作用域来确定
- 新的绑定仅为被调用过程的局部名字建立，这些名字在被调用过程的活动记录中占用存储单元
 - 这一点与静态作用域没有区别

6.3 非局部名字的访问

```

program dynamic(input, output);
  dynamic
  var r: real;
  procedure show;
    begin write(r: 5: 3) end;
  procedure small;
    var r: real;
    begin r := 0.125; show end;
begin
  r := 0.25;
  show; small; writeln;
  show; small; writeln
end.
  
```

6.3 非局部名字的访问

```

program dynamic(input, output);
  dynamic
  var r: real;
  procedure show;
    begin write(r: 5: 3) end;
  procedure small;
    var r: real;
    begin r := 0.125; show end;
begin
  r := 0.25;
  show; small; writeln;
  show; small; writeln
end.
  
```

6.3 非局部名字的访问

```

program dynamic(input, output);
  dynamic
  var r: real;
  procedure show;
    begin write(r: 5: 3) end;
  procedure small;
    var r: real;
    begin r := 0.125; show end;
begin
  r := 0.25;
  show; small; writeln;
  show; small; writeln
end.
  
```

6.3 非局部名字的访问

实现动态作用域的方法

- 深访问
 - 用控制链搜索运行栈，寻找包含该非局部名字的第一个活动记录
- 浅访问
 - 为每个名字在静态分配的存储空间中保存它的当前值
 - 当过程 p 的新活动出现时， p 的局部名字 n 使用在静态数据区分配给 n 的存储单元。 n 的先前值保存在 p 的活动记录中，当 p 的活动结束时再恢复

6.3 非局部名字的访问

```

program dynamic(input, output);  dynamic
var r: real;
procedure show;
begin write(r: 5: 3) end;
procedure small;
var r: real;
begin r := 0.125; show end;
begin (绿色表示已执行部分)
  r := 0.25;
  show; small; writeln;
  show; small; writeln
end.

```

静态区 使用值的
栈区 暂存值的

6.3 非局部名字访问

```

program dynamic(input, output);  dynamic
var r: real;
procedure show;
begin write(r: 5: 3) end;
procedure small;
var r: real;
begin r := 0.125; show end;
begin (绿色表示已执行部分)
  r := 0.25;
  show; small; writeln;
  show; small; writeln
end.

```

静态区 使用值的
栈区 暂存值的

6.3 非局部名字访问

```

program dynamic(input, output);  dynamic
var r: real;
procedure show;
begin write(r: 5: 3) end;
procedure small;
var r: real;
begin r := 0.125; show end;
begin (绿色表示已执行部分)
  r := 0.25;
  show; small; writeln;
  show; small; writeln
end.

```

静态区 使用值的
栈区 暂存值的

6.3 非局部名字访问

```

program dynamic(input, output);  dynamic
var r: real;
procedure show;
begin write(r: 5: 3) end;
procedure small;
var r: real;
begin r := 0.125; show end;
begin (绿色表示已执行部分)
  r := 0.25;
  show; small; writeln;
  show; small; writeln
end.

```

静态区 使用值的
栈区 暂存值的

6.3 非局部名字访问

```

program dynamic(input, output);  dynamic
var r: real;
procedure show;
begin write(r: 5: 3) end;
procedure small;
var r: real;
begin r := 0.125; show end;
begin (绿色表示已执行部分)
  r := 0.25;
  show; small; writeln;
  show; small; writeln
end.

```

静态区 使用值的
栈区 暂存值的

6.3 非局部名字访问

```

program dynamic(input, output);  dynamic
var r: real;
procedure show;
begin write(r: 5: 3) end;
procedure small;
var r: real;
begin r := 0.125; show end;
begin (绿色表示已执行部分)
  r := 0.25;
  show; small; writeln;
  show; small; writeln
end.

```

静态区 使用值的
栈区 暂存值的

6.4 参数传递

6.4.1 值调用(call by value)

- 实参的右值传给被调用过程
- 值调用可以如下实现
 - 把形参当作所在过程的局部名看待，形参的存储单元在该过程的活动记录中
 - 调用过程计算实参，并将其右值放入被调用过程形参的存储单元中

6.4 参数传递

6.4.2 引用调用(call by reference)

- 实参的左值传给被调用过程
- 引用调用可以如下实现：
 - 把形参当作所在过程的局部名看待，形参的存储单元在该过程的活动记录中
 - 调用过程计算实参，把实参的左值放入被调用过程形参的存储单元
 - 在被调用过程的目标代码中，任何对形参的引用都是通过传给该过程的指针来间接引用实参

6.4 参数传递

6.4.3 换名调用(call by name)

从概念上说，每次调用时，用实参表达式对形参进行正文替换，然后再执行

```
procedure swap(var x, y: integer);  
  var temp: integer;  
  begin  
    temp := x;  
    x := y;  
    y := temp  
  end
```

6.4 参数传递

6.4.3 换名调用

从概念上说，每次调用时，用实参表达式对形参进行正文替换，然后再执行

```
procedure swap(var x, y: integer);  
  var temp: integer; 例如： 调用swap(i, a[i])  
  begin  
    temp := x;  
    x := y;  
    y := temp  
  end
```

6.4 参数传递

6.4.3 换名调用

从概念上说，每次调用时，用实参表达式对形参进行正文替换，然后再执行

```
procedure swap(var x, y: integer);  
  var temp: integer; 例如： 调用swap(i, a[i])  
  begin 替换结果: temp := i;  
    temp := x; i := a[i];  
    x := y; a[i] := temp  
    y := temp  
  end
```

6.4 参数传递

6.4.3 换名调用

从概念上说，每次调用时，用实参表达式对形参进行正文替换，然后再执行

```
procedure swap(var x, y: integer);  
  var temp: integer; 例如： 调用swap(i, a[i])  
  begin 替换结果: temp := i;  
    temp := x; i := a[i];  
    x := y; a[i] := temp  
    y := temp 交换两个数据的程序  
  end 并非总是正确
```


6.5 堆 管 理

堆式分配(heap allocation)

- 堆用来存放生存期不确定的数据
 - C++和Java允许程序员用new创建对象，它们的生存期没有被约束在创建它们的过程活动的生成期之内
 - 实现内存回收是内存管理器的责任
- 堆空间的回收有两种不同方式
 - 程序显式释放空间：free (C) 或删除 (C++)
 - 垃圾收集器自动收集 (Java)。11.3节介绍垃圾收集算法，本课程不做介绍

6.5 堆 管 理

6.5.1 内存管理器(memory manager)

- 内存管理器把握的基本信息是堆中空闲空间
 - 分配函数
 - 回收函数
- 内存管理器应具有下列性质
 - 空间有效性：极小化程序需要的堆空间总量
 - 程序有效性：较好地利用内存子系统，使得程序能运行得快一些
 - 低开销：分配和回收操作所花时间在程序执行时间中的比例尽量小

6.5 堆 管 理

6.5.2 计算机内存分层(memory hierarchy)

典型大小		典型访问时间
> 2千兆字节	虚拟内存(磁盘)	3–15微秒
256兆–2千兆字节	物理内存	100–150纳秒
128千–4兆字节	2级缓存	40–60纳秒
16–64千字节	1级缓存	5–10纳秒
32字	寄存器(处理器)	1纳秒

6.5 堆 管 理

6.5.2 计算机内存分层

- 现代计算机都设计成程序员不用关心内存子系统的细节就可以写出正确的程序
- 程序的效率不仅取决于被执行的指令数，还取决于执行每条指令需要多长时间
- 执行一条指令的时间区别非常可观
- 差异源于硬件技术的基本局限：构造不了大容量的高速存储器
- 数据以块（缓存行、页）为单位在相邻层次之间进行传送
- 数据密集型程序可从恰当利用内存子系统中获益

6.5 堆 管 理

6.5.3 程序局部性(locality in programs)

- 大多数程序的大部分时间在执行一小部分代码，并且仅涉及一小部分数据
- 时间局部性(temporal locality)
 - 程序访问的内存单元在很短的时间内可能再次被程序访问
- 空间局部性(spatial locality)
 - 毗邻被访问单元的内存单元在很短的时间内可能被访问

6.5 堆 管 理

6.5.3 程序局部性

- 即使知道哪些指令会被频繁执行，最快的缓存也可能没有大到足以把它们同时放在其中，因此必须动态调整最快缓存的内容
- 把最近使用的指令保存在缓存是一种较好的最优利用内存分层的策略
- 改变数据布局或计算次序也可以改进程序数据访问的时间和空间局部性

6.5 堆 管 理

例：一个结构体大数组 分拆成若干个数组

```

struct student {
    int num;
    char name[20];
    ...
}
struct student st[10000];
    
```

```

int num[10000];
char name[10000][20];
...
    
```

- 若是顺序处理每个结构体的多个域，左边方式的数据局部性较好
- 若是先顺序处理每个结构的num域，再处理每个结构的name域，...，则右边方式的数据局部性较好
- 最好是按左边方式编程，由编译器决定是否需要将数据按右边方式布局

6.5 堆 管 理

6.5.4 手工回收(manual deallocation)请求

- 程序员在程序中显式释放堆块来达到回收堆块的目的
 - 内存泄漏：没有释放程序已经引用不到的堆块
只要内存没有用尽，它就不影响程序的正确性
自动无用单元收集通过回收所有无用单元来摆脱内存泄漏
 - 悬空引用(dangling pointer)：引用已经被释放的堆块
过分热心地释放数据对象而引起
悬空引用容易导致不会被捕获的错误

本 章 要 点

- 影响存储分配策略的语言特征
- 各种存储分配策略，主要了解静态分配和动态栈式分配
- 活动记录中各种数据域的作用和布局
- 非局部数据访问的实现方法
- 各种参数传递方式及其实现
- 堆管理

例 题 1

一个C语言程序及其在X86/Linux操作系统上的编译结果如下。根据所生成的汇编程序来解释程序中四个变量的存储分配、生存期、作用域和置初值方式等方面的区别

```

static long aa = 10;
short bb = 20;
func() {
    static long cc = 30;
    short dd = 40;
}
    
```

static long aa = 10;	例 题 1	func() {
short bb = 20;		static long cc = 30;
		short dd = 40; }
.data		.align 4
.align 4		.type cc.2,@object
.type aa,@object		.size cc.2,4
.size aa,4	cc.2:	
aa:		.long 30
.long 10		.text
.globl bb		.align 4
.align 2		.globl func
.type bb,@object	func:	
.size bb,2		...
bb:		movw \$40,-2(%ebp)
.value 20		...

static long aa = 10;	例 题 1	func() {
short bb = 20;		static long cc = 30;
		short dd = 40; }
.data		.align 4
.align 4		.type cc.2,@object
.type aa,@object		.size cc.2,4
.size aa,4	cc.2:	
aa:		.long 30
.long 10		.text
.globl bb		.align 4
.align 2		.globl func
.type bb,@object	func:	
.size bb,2		...
bb:		movw \$40,-2(%ebp)
.value 20		...

static long aa = 10; short bb = 20;	例 题	func() { 1 static long cc = 30; short dd = 40; }
.data		.align 4
.align 4		.type cc.2,@object
.type aa,@object		.size cc.2,4
.size aa,4		cc.2:
aa:		.long 30
.long 10		.text
.globl bb		.align 4
.align 2		.globl func
.type bb,@object		func:
.size bb,2		...
bb:		movw \$40,-2(%ebp)
.value 20		...

static long aa = 10; short bb = 20;	例 题	func() { 1 static long cc = 30; short dd = 40; }
.data		.align 4
.align 4		.type cc.2,@object
.type aa,@object		.size cc.2,4
.size aa,4		cc.2:
aa:		.long 30
.long 10		.text
.globl bb		.align 4
.align 2		.globl func
.type bb,@object		func:
.size bb,2		...
bb:		movw \$40,-2(%ebp)
.value 20		...

static long aa = 10; short bb = 20;	例 题	func() { 1 static long cc = 30; short dd = 40; }
.data		.align 4
.align 4		.type cc.2,@object
.type aa,@object		.size cc.2,4
.size aa,4		cc.2:
aa:		.long 30
.long 10		.text
.globl bb		.align 4
.align 2		.globl func
.type bb,@object		func:
.size bb,2		...
bb:		movw \$40,-2(%ebp)
.value 20		...

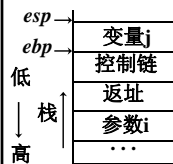
static long aa = 10; short bb = 20;	例 题	func() { 1 static long cc = 30; short dd = 40; }
.data		.align 4
.align 4		.type cc.2,@object
.type aa,@object		.size cc.2,4
.size aa,4		cc.2:
aa:		.long 30
.long 10		.text
.globl bb		.align 4
.align 2		.globl func
.type bb,@object		func:
.size bb,2		...
bb:		movw \$40,-2(%ebp)
.value 20		...

例 题 2

```
func(i)
long i;
{
    long j;
    j = i - 1;
    func(j);
}
```

例 题 2

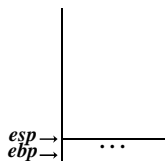
```
func:
pushl %ebp 老的基地址指针压栈
movl %esp,%ebp 修改基地址指针
subl $4,%esp 为j分配空间
movl 8(%ebp),%edx 取i到寄存器
decl %edx i - 1
movl %edx,-4(%ebp) i - 1 → j
movl -4(%ebp),%eax
pushl %eax 把实参j的值压栈
call func 函数调用
addl $4,%esp 恢复栈顶指针
L1:
leave 即 mov ebp, esp; pop ebp
ret 即 pop eip(下条指令地址)
```



例 题 2

```
func(i)      func:
long i;      pushl %ebp  老的基地址指针压栈
{            movl %esp,%ebp 修改基地址指针
  long j;    subl $4,%esp   为j分配空间
  j = i - 1;  movl 8(%ebp),%edx 取i到寄存器
  func(j);   decl %edx      i - 1
}            movl %edx,-4(%ebp)  i - 1 ⇒ j
             movl -4(%ebp),%eax
             pushl %eax 把实参j的值压栈
             call func  函数调用
             addl $4,%esp 恢复栈顶指针

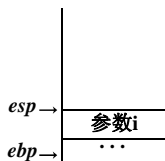
             L1:
             leave 即 mov ebp, esp; pop ebp
             ret   即 pop eip(下条指令地址)
```



例 题 2

```
func(i)      func:
long i;      pushl %ebp  老的基地址指针压栈
{            movl %esp,%ebp 修改基地址指针
  long j;    subl $4,%esp   为j分配空间
  j = i - 1;  movl 8(%ebp),%edx 取i到寄存器
  func(j);   decl %edx      i - 1
}            movl %edx,-4(%ebp)  i - 1 ⇒ j
             movl -4(%ebp),%eax
             pushl %eax 把实参j的值压栈
             call func  函数调用
             addl $4,%esp 恢复栈顶指针

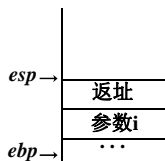
             L1:
             leave 即 mov ebp, esp; pop ebp
             ret   即 pop eip(下条指令地址)
```



例 题 2

```
func(i)      func:
long i;      pushl %ebp  老的基地址指针压栈
{            movl %esp,%ebp 修改基地址指针
  long j;    subl $4,%esp   为j分配空间
  j = i - 1;  movl 8(%ebp),%edx 取i到寄存器
  func(j);   decl %edx      i - 1
}            movl %edx,-4(%ebp)  i - 1 ⇒ j
             movl -4(%ebp),%eax
             pushl %eax 把实参j的值压栈
             call func  函数调用
             addl $4,%esp 恢复栈顶指针

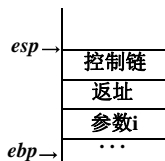
             L1:
             leave 即 mov ebp, esp; pop ebp
             ret   即 pop eip(下条指令地址)
```



例 题 2

```
func(i)      func:
long i;      pushl %ebp  老的基地址指针压栈
{            movl %esp,%ebp 修改基地址指针
  long j;    subl $4,%esp   为j分配空间
  j = i - 1;  movl 8(%ebp),%edx 取i到寄存器
  func(j);   decl %edx      i - 1
}            movl %edx,-4(%ebp)  i - 1 ⇒ j
             movl -4(%ebp),%eax
             pushl %eax 把实参j的值压栈
             call func  函数调用
             addl $4,%esp 恢复栈顶指针

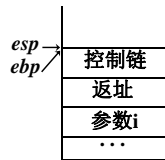
             L1:
             leave 即 mov ebp, esp; pop ebp
             ret   即 pop eip(下条指令地址)
```



例 题 2

```
func(i)      func:
long i;      pushl %ebp  老的基地址指针压栈
{            movl %esp,%ebp 修改基地址指针
  long j;    subl $4,%esp   为j分配空间
  j = i - 1;  movl 8(%ebp),%edx 取i到寄存器
  func(j);   decl %edx      i - 1
}            movl %edx,-4(%ebp)  i - 1 ⇒ j
             movl -4(%ebp),%eax
             pushl %eax 把实参j的值压栈
             call func  函数调用
             addl $4,%esp 恢复栈顶指针

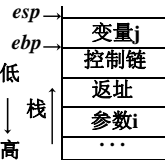
             L1:
             leave 即 mov ebp, esp; pop ebp
             ret   即 pop eip(下条指令地址)
```



例 题 2

```
func(i)      func:
long i;      pushl %ebp  老的基地址指针压栈
{            movl %esp,%ebp 修改基地址指针
  long j;    subl $4,%esp   为j分配空间
  j = i - 1;  movl 8(%ebp),%edx 取i到寄存器
  func(j);   decl %edx      i - 1
}            movl %edx,-4(%ebp)  i - 1 ⇒ j
             movl -4(%ebp),%eax
             pushl %eax 把实参j的值压栈
             call func  函数调用
             addl $4,%esp 恢复栈顶指针

             L1:
             leave 即 mov ebp, esp; pop ebp
             ret   即 pop eip(下条指令地址)
```



例题 2

返回序列之一
返回序列之二

```

func(i)
long i;
{
    long j;
    j = i - 1;
    func(j);
}

```

esp →

ebp →

...

```

func:
    pushl %ebp  老的基地址指针压栈
    movl %esp,%ebp 修改基地址指针
    subl $4,%esp  为j分配空间
    movl 8(%ebp),%edx 取i到寄存器
    decl %edx      i - 1
    movl %edx,-4(%ebp)  i - 1 ⇒ j
    movl -4(%ebp),%eax
    pushl %eax  把实参j的值压栈
    call func   函数调用
    addl $4,%esp  恢复栈顶指针
L1:
    leave  即 mov ebp, esp; pop ebp
    ret    即 pop eip(下条指令地址)

```

例题 3

下面的程序运行时输出3个整数。试从运行环境和printf的实现来分析，为什么此程序会有3个整数输出？

```

main()
{
    printf(“%d, %d, %d\n”);
}

```

例题 4

```

main()
{
    char *cp1, *cp2;

    cp1 = "12345";
    cp2 = "abcdefghij";
    strcpy(cp1, cp2);
    printf("cp1 = %s\ncp2 = %s\n", cp1, cp2);
}

```

在某些系统上的运行结果是：
cp1 = abcdefghij
cp2 = ghij
为什么cp2所指的串被修改了？

例题 4

因为常量串“12345”和“abcdefghij”连续分配在常数区

执行前：

```

1 2 3 4 5 \0 a b c d e f g h i j \0
↑      ↑
cp1     cp2

```

例题 4

因为常量串“12345”和“abcdefghij”连续分配在常数区

执行前：

```

1 2 3 4 5 \0 a b c d e f g h i j \0
↑      ↑
cp1     cp2

```

执行后：

```

a b c d e f g h i j \0 f g h i j \0
↑      ↑
cp1     cp2

```

例题 4

因为常量串“12345”和“abcdefghij”连续分配在常数区

执行前：

```

1 2 3 4 5 \0 a b c d e f g h i j \0
↑      ↑
cp1     cp2

```

执行后：

```

a b c d e f g h i j \0 f g h i j \0
↑      ↑
cp1     cp2

```

现在的编译器大都把程序中的串常量单独存放在只读数据段中，因此运行时不会报错

```
func(i,j,f,e)
short i,j; float f,e;
{
    short il,jl; float fl,e1;
    printf(&i,&j,&f,&e);
    printf(&il,&jl,&fl,&e1);
}
main()
{
    short i,j; float f,e;
    func(i,j,f,e);
}
Address of i,j,f,e = ...36, ...42, ...44, ...54 (八进制数)
Address of il,jl,fl,e1 = ...26, ...24, ...20, ...14
```

```
func(i,j,f,e)           Sizes of short, int, long, float,
short i,j; float f,e;   double = 2, 4, 4, 4, 8
{                         (在SPARC/SUN工作站上)
    short i1,j1; float f1,e1;
    printf(&i,&j,&f,&e);
    printf(&i1,&j1,&f1,&e1);
}
main()
{
    short i,j; float f,e;
    func(i,j,f,e);
}
Address of i,j,f,e = ...36, ...42, ...44, ...54 (八进制数)
Address of i1,j1,f1,e1 = ...26, ...24, ...20, ...14
```

func(i,j,f,e)	Sizes of short, int, long, float,
short i,j; float f,e;	double = 2, 4, 4, 4, 8
{	(在SPARC/SUN工作站上)
short il,jl; float fl,e1;	
printf(&i,&j,&f,&e);	
printf(&i1,&j1,&f1,&e1);	
}	
main()	为什么4个形式参数i,j,f,e的地址
{	间隔和它们类型的大小不一致
short i,j; float f,e;	
func(i,j,f,e);	
}	
Address of i,j,f,e = ...36, ...42, ...44, ...54 (八进制数)	
Address of il,jl,fl,e1 = ...26, ...24, ...20, ...14	

- 当用传统的参数声明方式时，编译器不检查实参和形参的个数和类型是否一致，由程序员自己负责
- 但对形参和实参是不同的整型，或不同的实型
 - 编译器试图保证运行时能得到正确结果
 - 条件是：若需数据类型转换时，不出现溢出
- 编译器的做法
 - 把整型或实型数据分别提升到long和double类型的数据，再传递到被调用函数
 - 被调用函数根据形参所声明的类型，决定是否要将传来的实参向低级别类型转换

Diagram illustrating memory layout for long, short, float, and double types. The top part shows 'long' (4 cells) and 'short' (2 cells) with '低地址放高位' (low address for high bit) on the left and '高地址放低位' (high address for low bit) on the right. The bottom part shows 'float' (4 cells) and 'double' (8 cells) with '低地址放高位' on the left and '高地址放低位' on the right.

在main函数中参数压栈时的观点

在func函数中存取形式参数时的观点

栈的增长方向

参数在栈中的情况

i, 4个字节: 2个字节, 起始地址36

j, 4个字节: 2个字节, 起始地址42

f, 8个字节: 4个字节, 起始地址44

e, 8个字节: 4个字节, 起始地址54

例 题 6

下面程序为什么死循环（在SPARC/SUN工作站上）？

```
main() { addr(); loop(); }
long *p;
loop()
{
    long i,j;
    j=0;
    for(i=0;i<10;i++){ (*p)--; j++; }
}
addr() { long k; k=0; p=&k;}
```

例 题 6

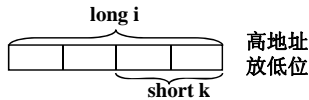
将long *p改成short *p，long k 改成short k 后，循环体执行一次便停止，为什么？

```
main() { addr(); loop(); }
short *p;
loop()
{
    long i,j;
    j=0;
    for(i=0;i<10;i++){ (*p)--; j++; }
}
addr() { short k; k=0; p=&k;}
```

例 题 6

将long *p改成short *p，long k 改成short k 后，循环体执行一次便停止，为什么？

```
main() { addr(); loop(); }
short *p;
loop()
{
    long i,j;
    j=0;
    for(i=0;i<10;i++){ (*p)--; j++; }
}
addr() { short k; k=0; p=&k;}
```



例 题 7

```
main()
{ func(); printf("Return from func\n"); }
```

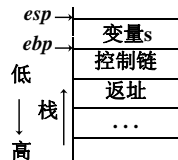
```
func()
{ char s[4];
  strcpy(s,"12345678"); printf("%s\n",s); }
```

在X86/Linux操作系统上的运行结果如下：
12345678
Return from func
Segmentation fault (core dumped)

例 题 7

```
main()
{ func(); printf("Return from func\n"); }
```

```
func()
{ char s[4];
  strcpy(s,"12345678");
  printf("%s\n",s);
}
```

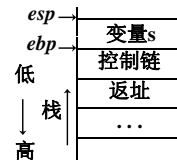


例 题 7

```
main()
{ func(); printf("Return from func\n"); }
```

```
func()
{ char s[4];
  strcpy(s,"123456789");
  printf("%s\n",s);
}
```

123456789
Segmentation fault (core dumped)



例 题 8

```
int fact(i)      | main()
int i;           | {
{               |   printf("%d\n", fact(5));
  if(i==0)       |   printf("%d\n", fact(5,10,15));
    return 1;    |   printf("%d\n", fact(5.0));
  else          |   printf("%d\n", fact());
    return i*fact(i-1); | }
}
```

该程序在X86/Linux机器上的运行结果如下：

```
120
120
1
Segmentation fault (core dumped)
```

例 题 8

请解释下面问题：

- 第二个fact调用：结果为什么没有受参数过多的影响？
- 第三个fact调用：为什么用浮点数5.0作为参数时结果变成1？
- 第四个fact调用：为什么没有提供参数时会出现 Segmentation fault？

例 题 8

请解释下面问题：

- 第二个fact调用：结果为什么没有受参数过多的影响？

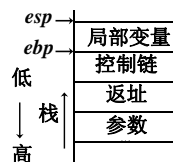
解答：参数表达式逆序计算并进栈，fact能够取到第一个参数

例 题 8

请解释下面问题：

- 第三个fact调用：为什么用浮点数5.0作为参数时结果变成1？

解答：参数5.0转换成双精度数进栈，占8个字节
它低地址的4个字节看成整数时正好是0

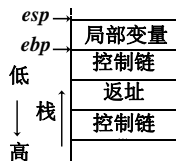


例 题 8

请解释下面问题：

- 第四个fact调用：为什么没有提供参数时会出现 Segmentation fault？

解答：由于没有提供参数，而main函数又无局部变量，fact把老ebp（控制链）（main的活动记录中保存的ebp）当成参数，它一定是一个很大的整数，使得活动记录栈溢出



习 题

- 第一次：6.3, 6.4, 6.5
- 第二次：6.6, 6.9, 6.12
- 第三次：6.16, 6.18, 6.23