



中国科学技术大学
University of Science and Technology of China



《编译原理与技术》

运行时存储空间的组织和管理

计算机科学与技术学院

李诚

29/11/2018



□术语

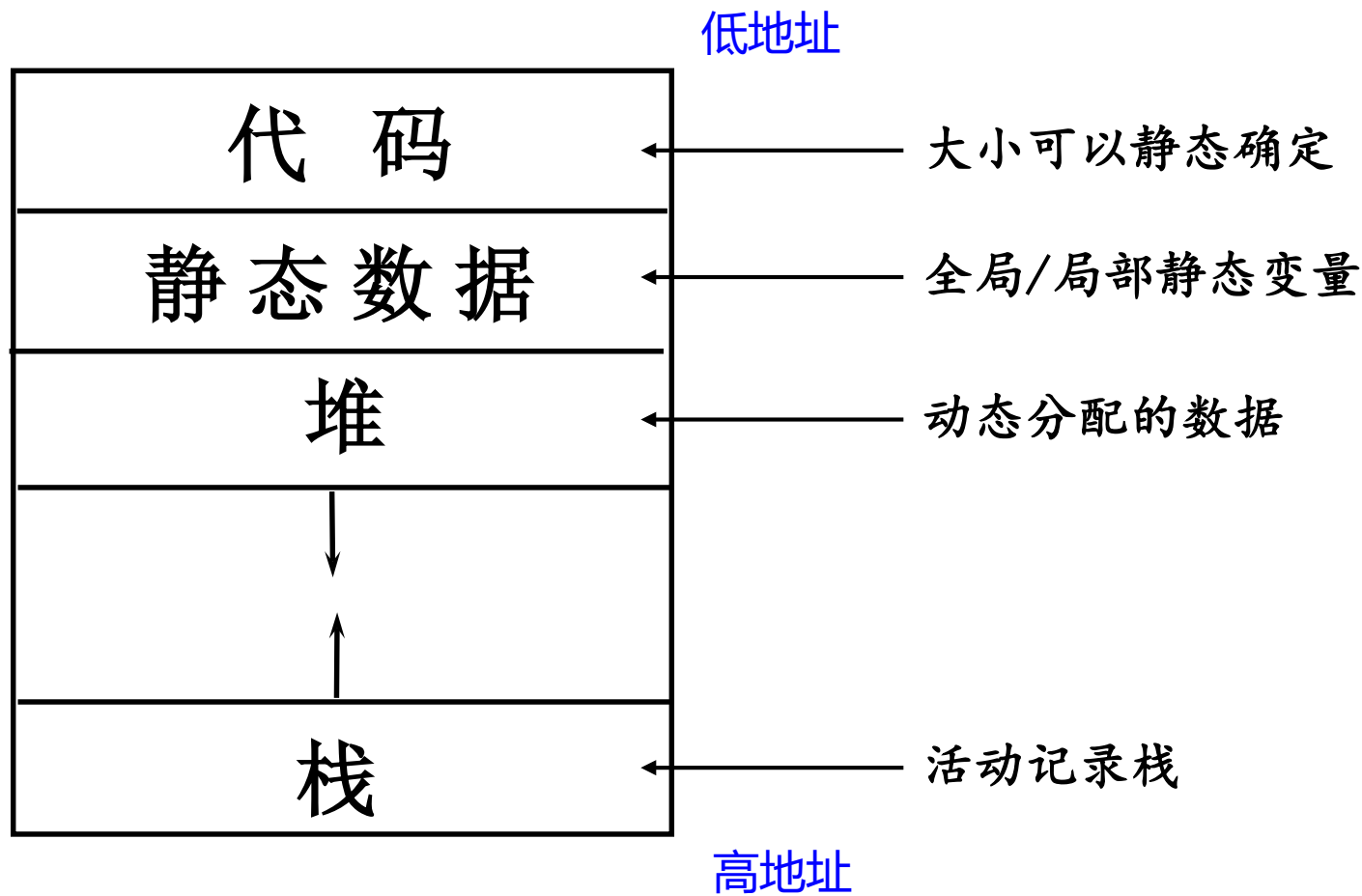
- ❖ 过程的活动(activation): 过程的一次执行
- ❖ 活动记录: 过程的活动需要可执行代码和存放所需信息的存储空间,后者称为活动记录

□本章内容

- ❖ 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
- ❖ 非局部名字的管理、参数传递方式、堆管理



程序的存储分配





- 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被动态地分配
- 存储块是否必须被显式地释放



□过程

- FORTRAN的子例程(subroutine)
- PASCAL的过程/函数(procedure/function)
- C的函数

□过程的激活（调用）与终止（返回）

□过程的执行需要：

代码段 + 活动记录（过程运行所需的额外信息，如参数，局部数据，返回地址等）

□ **基本概念：作用域与生存期**

□ **活动记录的常见布局**

❖ 字节寻址、类型、次序、对齐

□ **程序块：同名情况的处理**



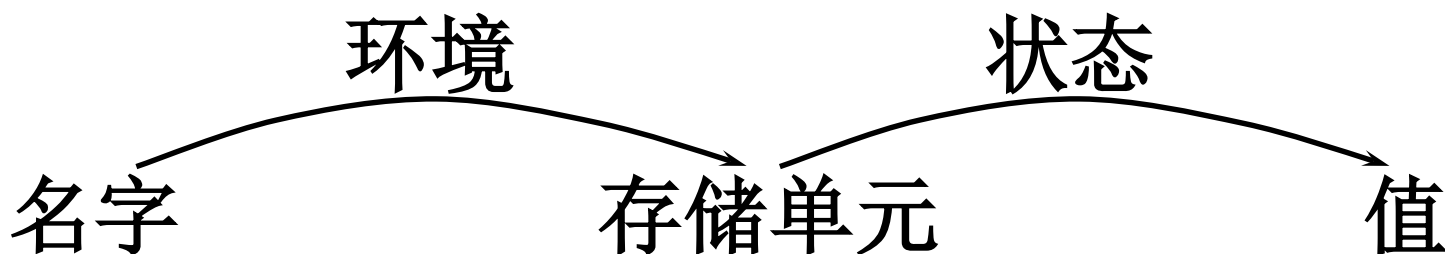
□名字的作用域

- ❖ 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
- 如下图代码中的n

```
int f(int n){  
    if (n<0) error("arg<0");  
    else if (n==0) return 1;  
    else return n*f(n-1);  
}
```

□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





□ 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



□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

□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

□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{
```

```
    char c1;    0
```

```
    long i;     4
```

```
    char c2;    8
```

```
    double f;  16
```

```
}a;
```

```
typedef struct _b{
```

```
    char c1;    0
```

```
    char c2;    1
```

```
    long i;     4
```

```
    double f;   8
```

```
}b;
```

对齐: char : 1, long : 4, double : 8

□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{
```

```
    char c1;    0
```

```
    long i;     4
```

```
    char c2;    8
```

```
    double f;   12
```

```
}a;
```

```
typedef struct _b{
```

```
    char c1;    0
```

```
    char c2;    1
```

```
    long i;     4
```

```
    double f;   8
```

```
}b;
```

对齐: char : 1, long : 4, double : 4



□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

```
main() /* 例 */
{
    int a = 0;
    int b = 0;
    {
        int b = 1;
        {
            int a = 2;
        }
        {
            int b = 3;
        }
    }
}
```

/ begin of B_0 */*

/ begin of B_1 */*

/ begin of B_2 */*

/ end of B_2 */*

/ begin of B_3 */*

/ end of B_3 */*

/ end of B_1 */*

/ end of B_0 */*

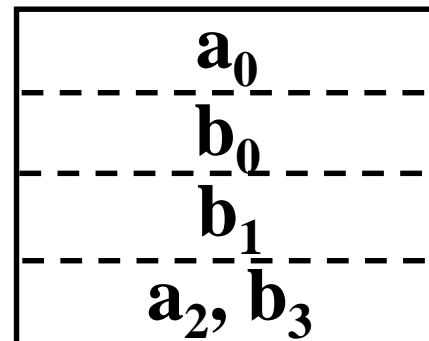


```

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





- **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**



□ 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



□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



□主要有两种策略

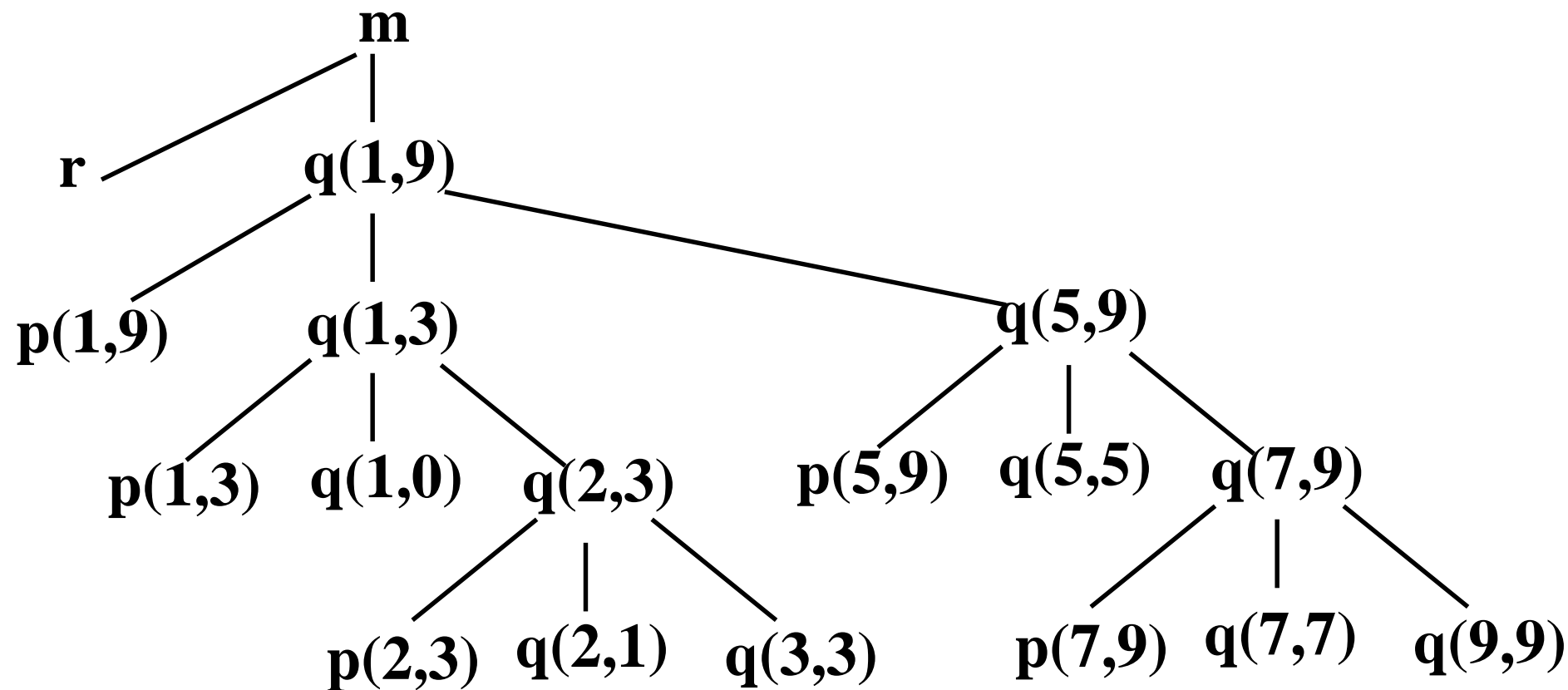
- ❖ 栈式存储：与过程调用返回有关，涉及过程的局部变量以及过程活动记录
- ❖ 堆存储：关系到部分生存周期较长的数据



- 介绍程序运行时所需的各个活动记录在存储空间中的分配策略
- 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

□ 活动树

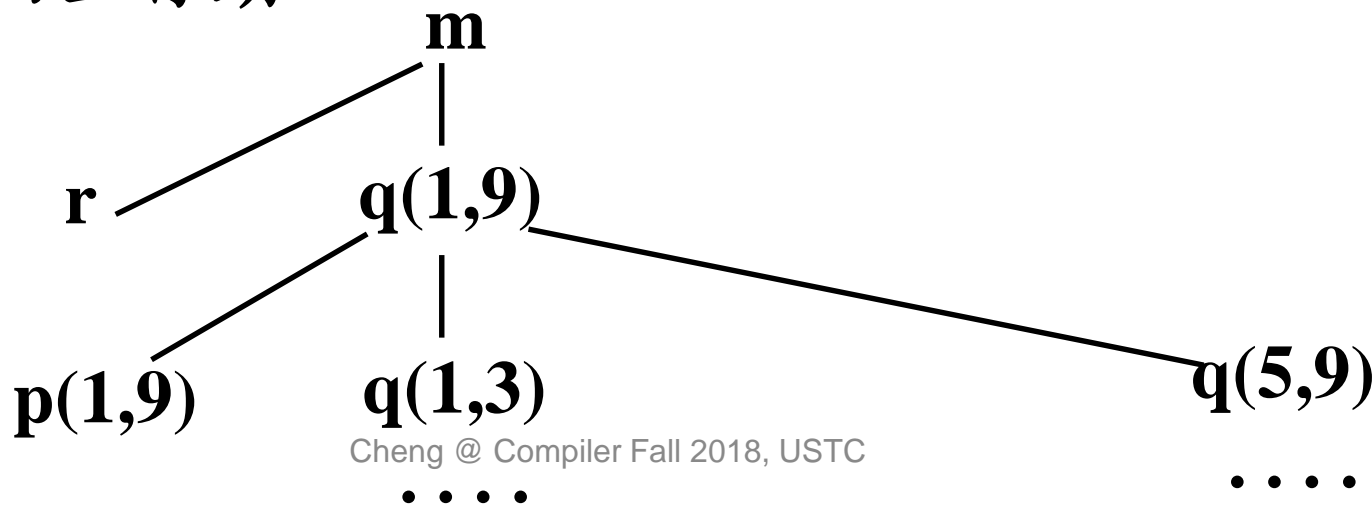
❖ 用树来描绘控制进入和离开活动的方式





□活动树的特点

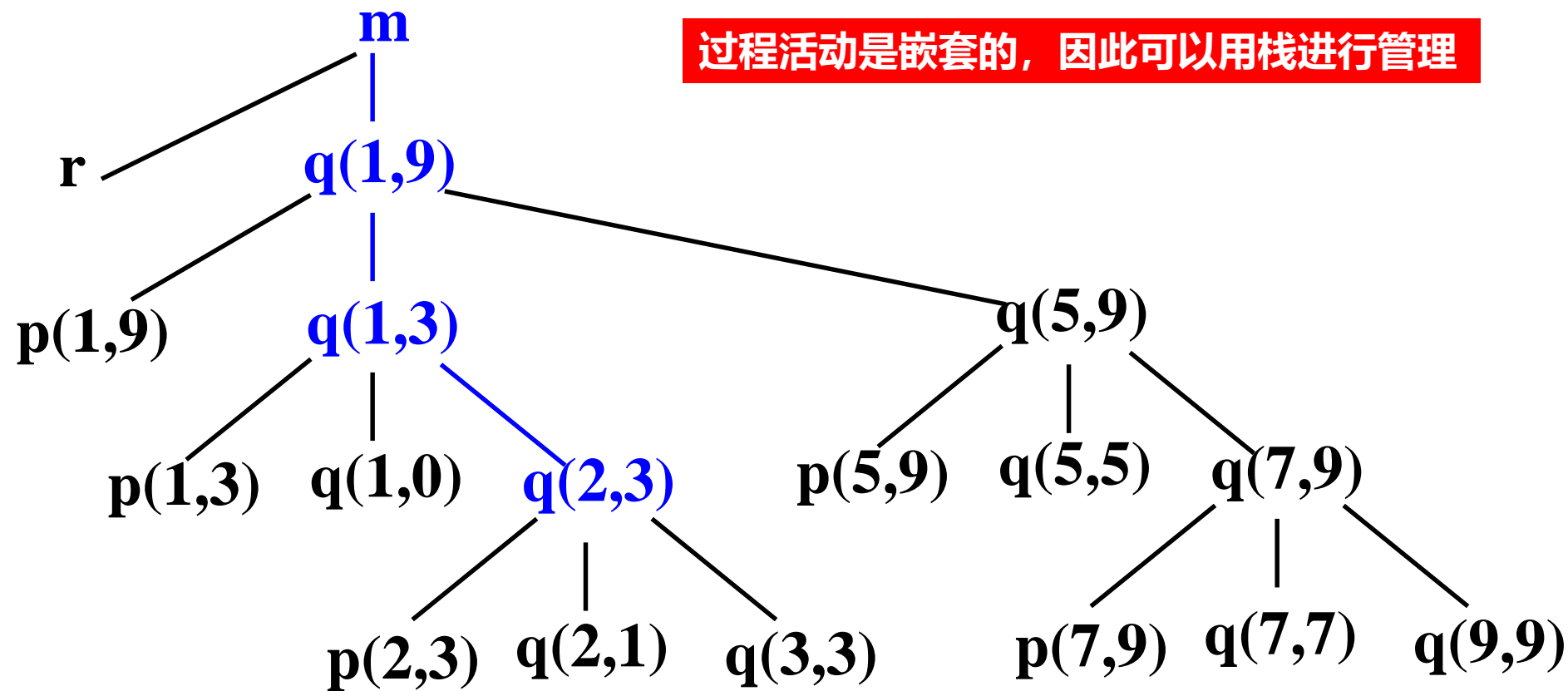
- ❖ 每个结点代表某过程的一个活动
- ❖ 根结点代表主程序的活动
- ❖ 结点 a 是结点 b 的父结点，当且仅当控制流从 a 的活动进入 b 的活动
- ❖ 结点 a 处于结点 b 的左边，当且仅当 a 的生存期先于 b 的生存期



□当前活跃着的过程活动可以保存在一个栈中

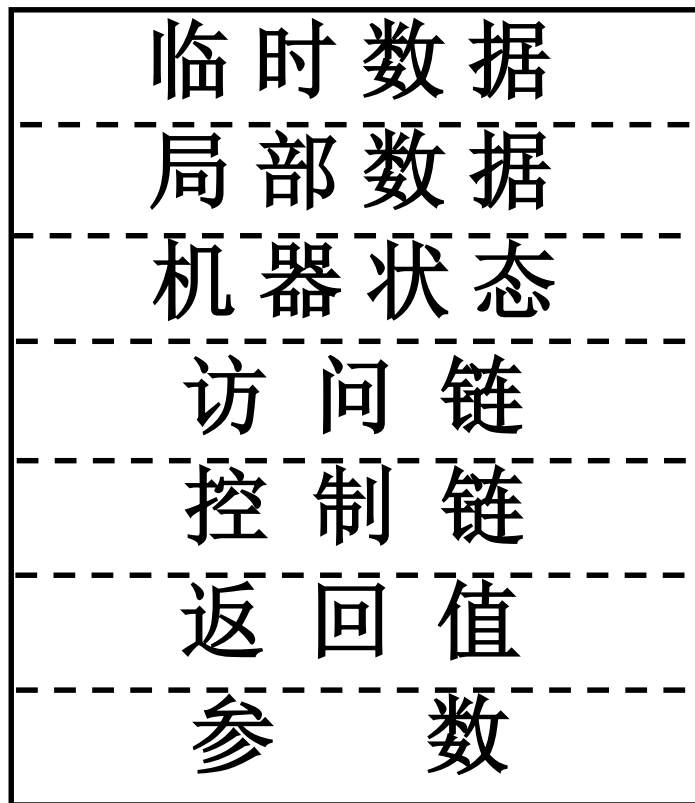
❖例 控制栈的内容： $m, q(1, 9), q(1, 3), q(2, 3)$

过程活动是嵌套的，因此可以用栈进行管理





□ 活动记录的常见布局





□活动记录 - AR (Activation Record)

❖ 是一连续存储区域，用于管理与存放和程序单元执行相关的重要信息。

□AR中的内容

- ❖ **临时区域**：用以保存无法存放在寄存器中的临时计算结果
- ❖ **局部数据区**：源程序中程序单元声明的局部变量对应在此区域
- ❖ **机器状态保存区**：存有此次调用之前的机器状态信息，如寄存器内容，程序指令计数器值ip（返回地址）等

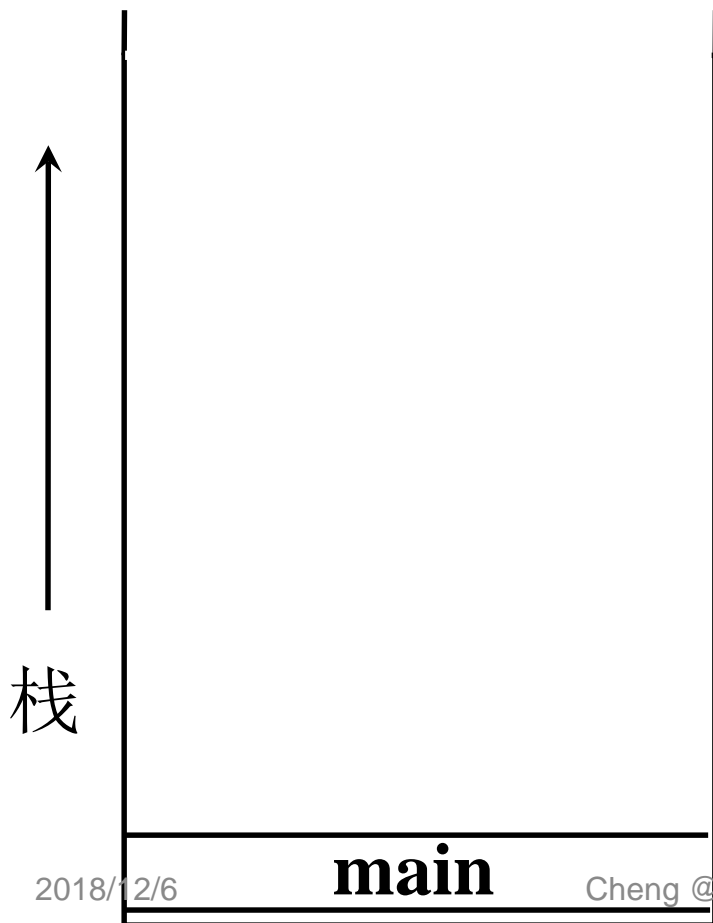


□AR中的内容

- ❖ **可选访问链（静态链）**。当前程序单元可以访问的（静态程序中）外围程序单元的活动记录链。
- ❖ **控制链（动态链）**。程序单元的活动记录按它们的生成（或调用）次序串成链。
- ❖ **实在参数**
- ❖ **返回值**

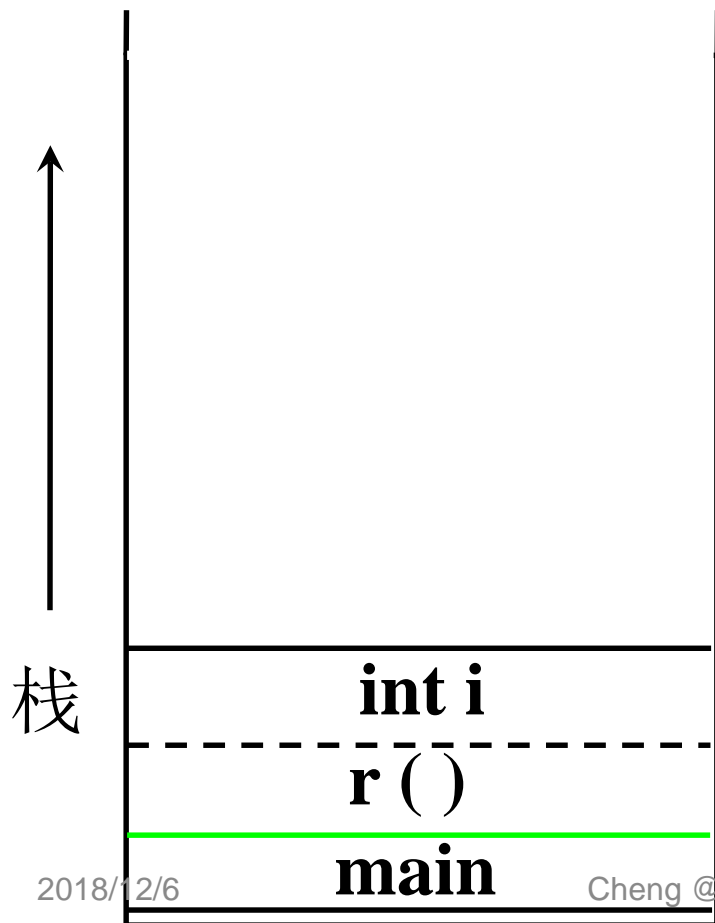
□运行栈：把控制栈中的信息拓广到包括过程活动所需的所有局部信息（即活动记录）

□ **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）**

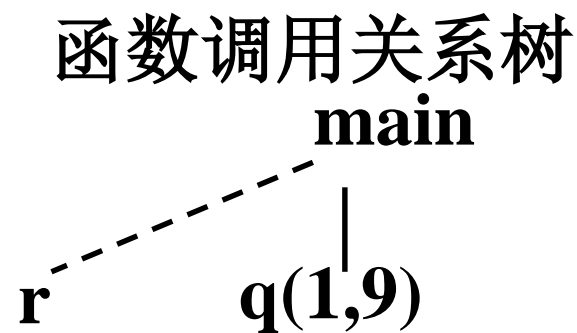
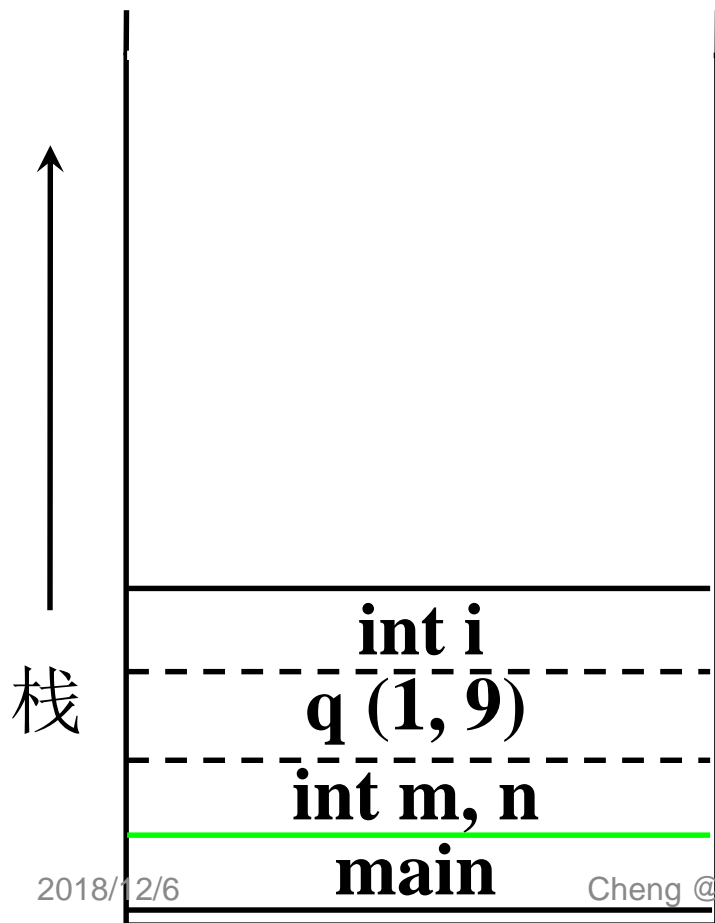


函数调用关系树
main

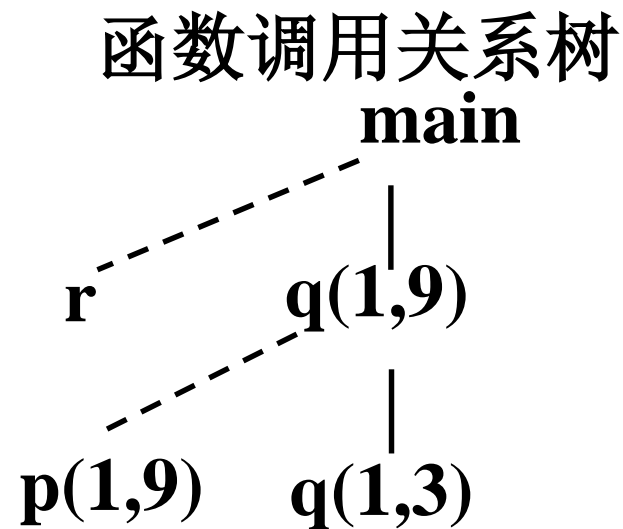
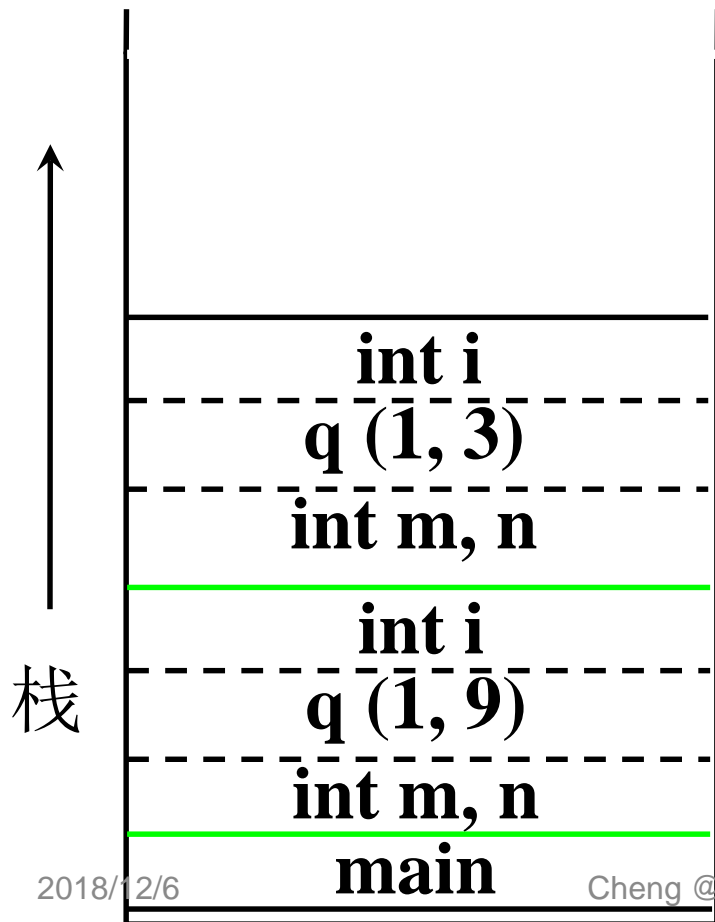
□运行栈：把控制栈中的信息拓广到包括过程活动所需的所有局部信息（即活动记录）



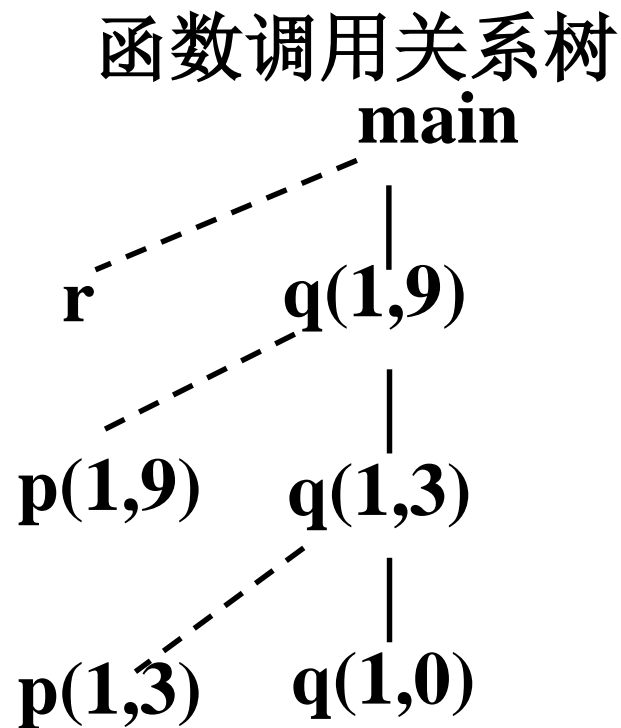
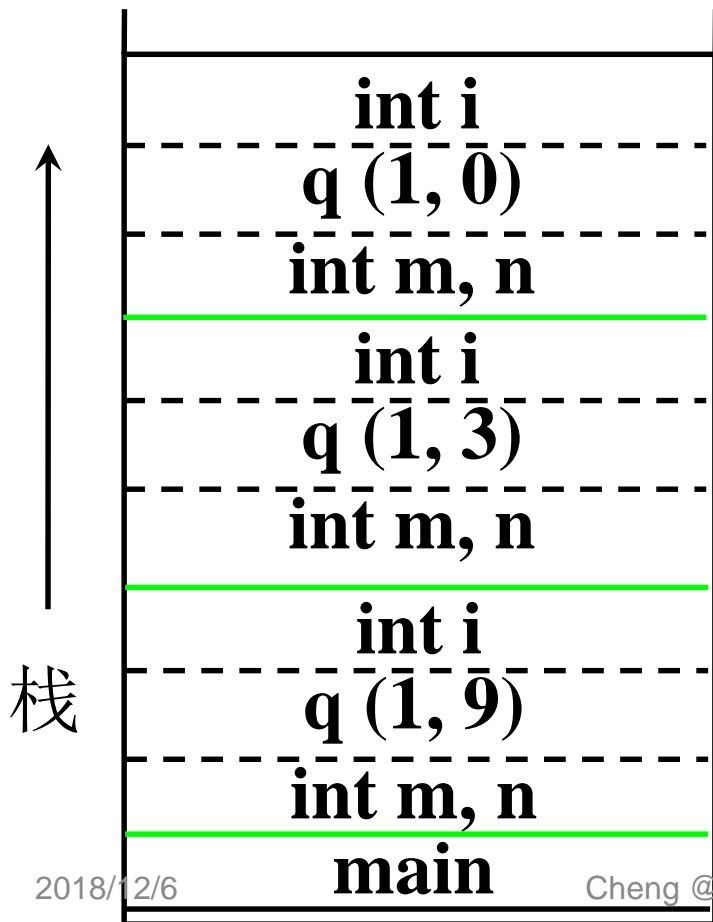
□运行栈：把控制栈中的信息拓广到包括过程活动所需的所有局部信息（即活动记录）



□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）



□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）



□ 代码序列

❖ 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(calling sequence)

❖ 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(return sequence)

❖ 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的活动记录中**

□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

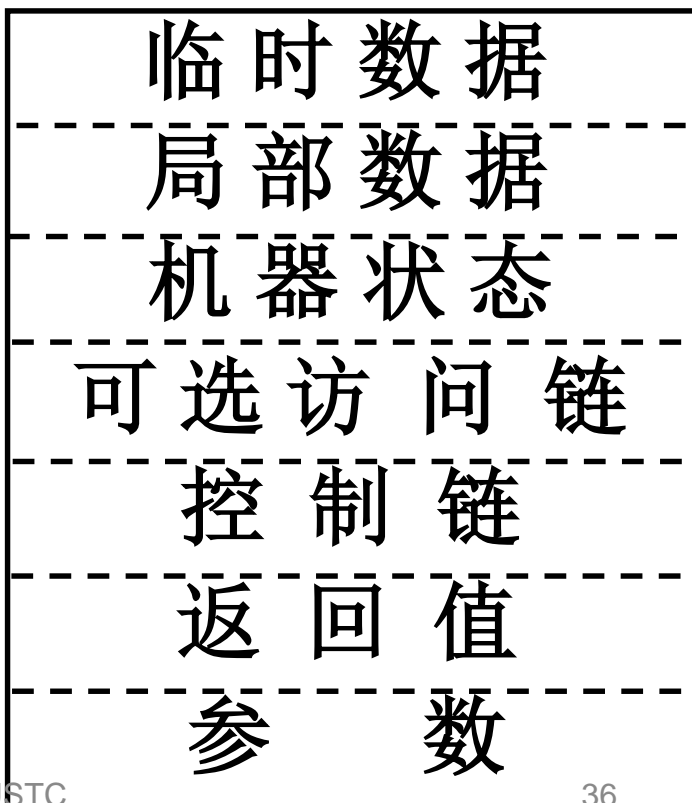
- ❖以活动记录中间的某个位置作为基地址(控制链)
- ❖长度能较早确定的域放在活动记录的中间



□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
- ❖把参数域和可能有的返回
值域放在紧靠调用者活动
记录的地方





□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

临时数据

局部数据

机器状态

访问链

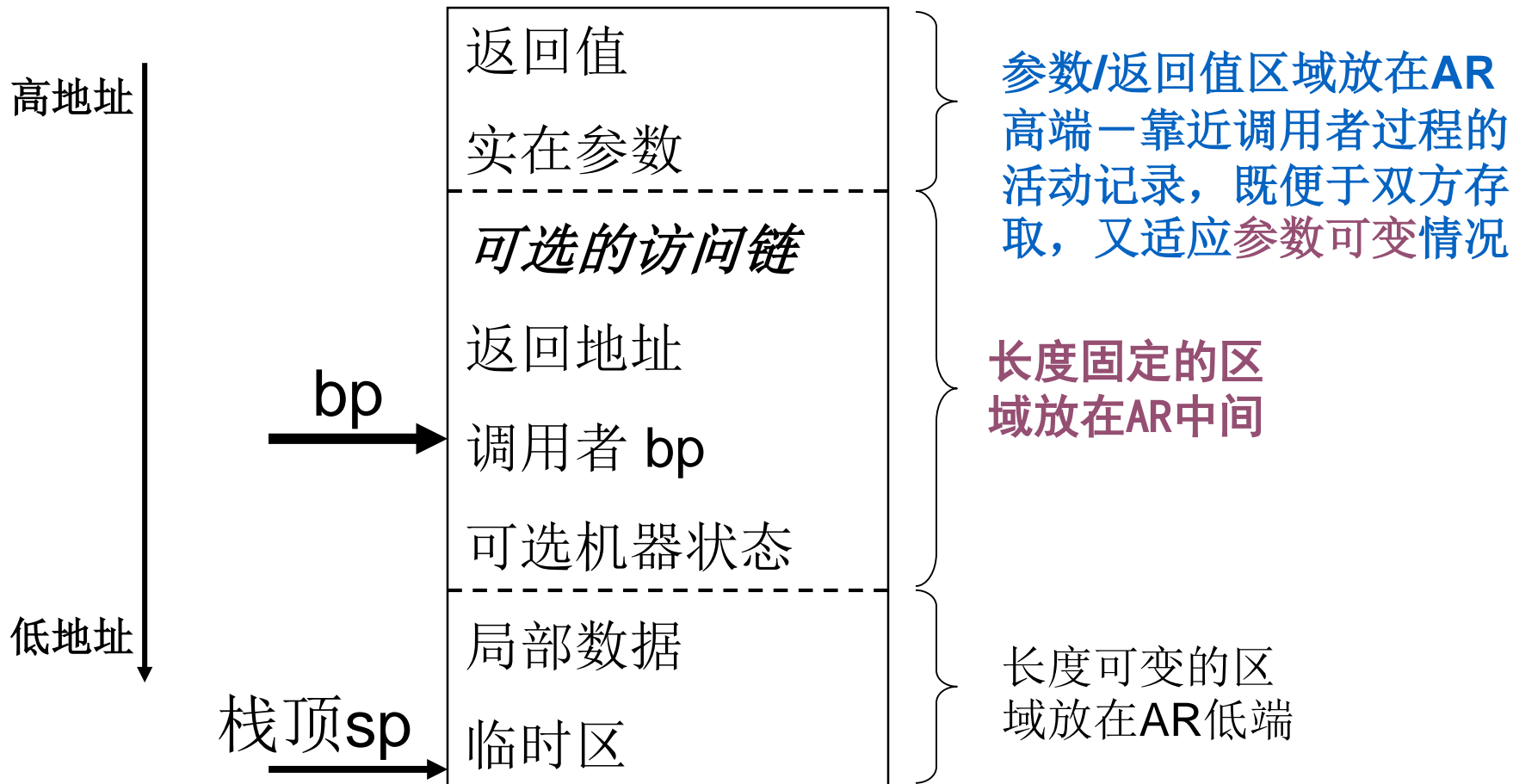
控制链

返回值

参数

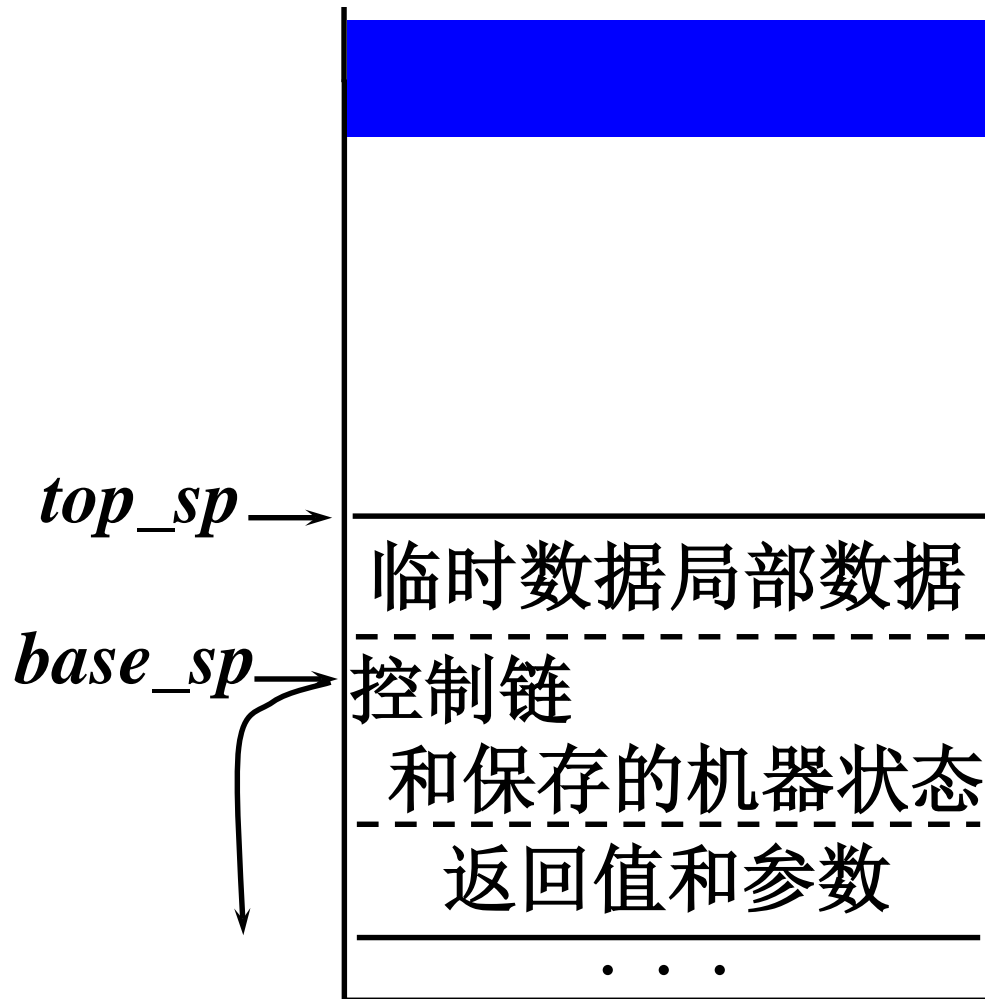


□ 栈式分配下的AR内容布局





□ 过程p调用过程q的调用序列



- ❖ **top_sp**: 栈顶寄存器，如esp、rsp
- ❖ **base_sp**: 基址寄存器，如ebp、rbp
- ❖ **PC**: 程序计数器，如eip、rip

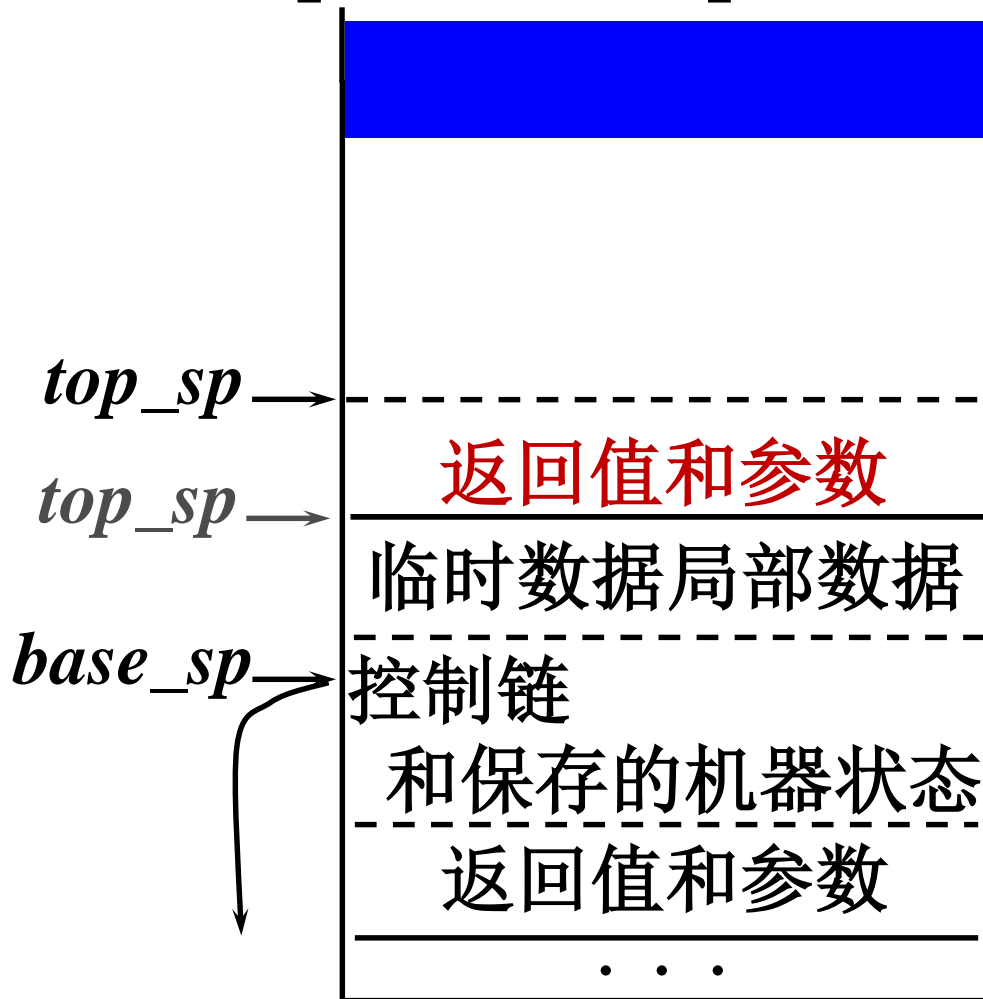
□ 常见寄存器

寄存器	16位	32位	64位
累加寄存器	AX	EAX	RAX
基址寄存器	BX	EBX	RBX
计数寄存器	CX	ECX	RCX
数据寄存器	DX	EDX	RDX
堆栈基指针	BP	EBP	RBP
变址寄存器	SI	ESI	RSI
堆栈顶指针	SP	ESP	RSP
指令寄存器	IP	EIP	RIP

source: <https://tinyurl.com/yavdmta8>



□ 过程p调用过程q的调用序列(栈往上增长)



(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



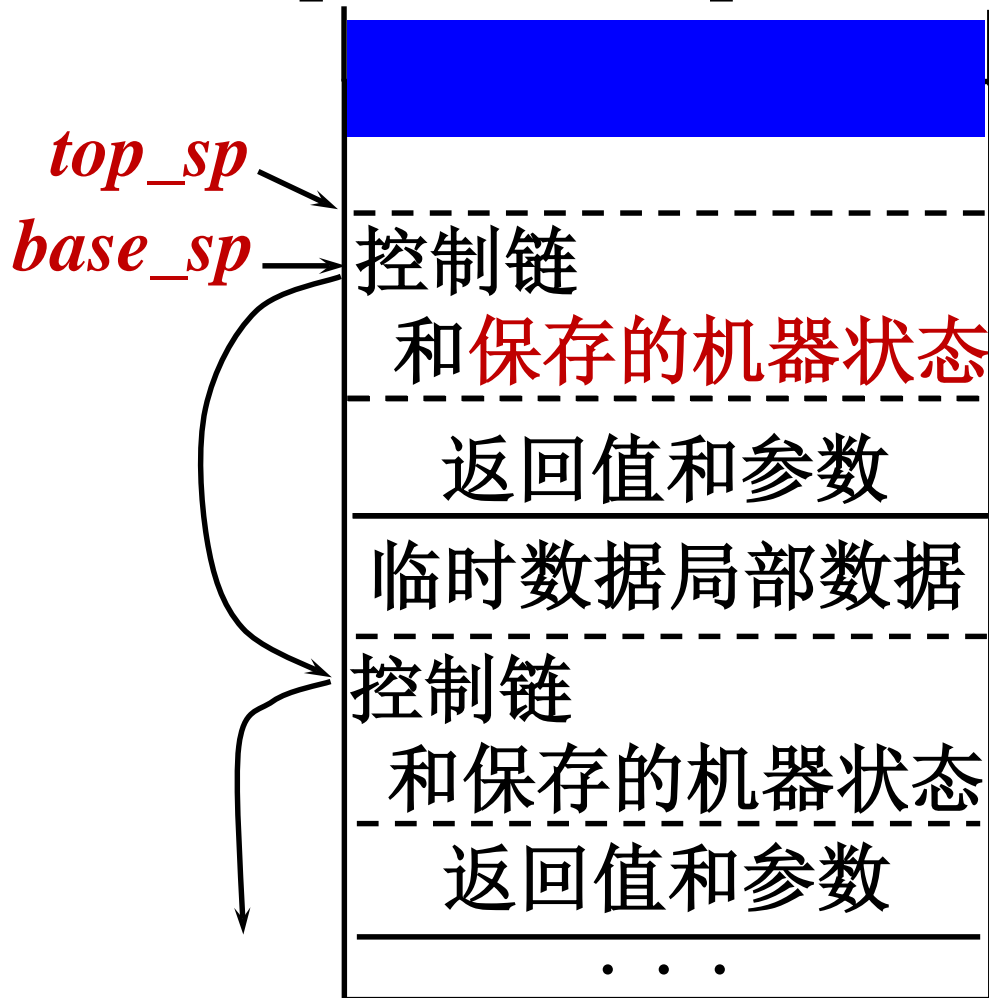
□ 过程p调用过程q的调用序列



(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**，改变 $base_sp$ 的值



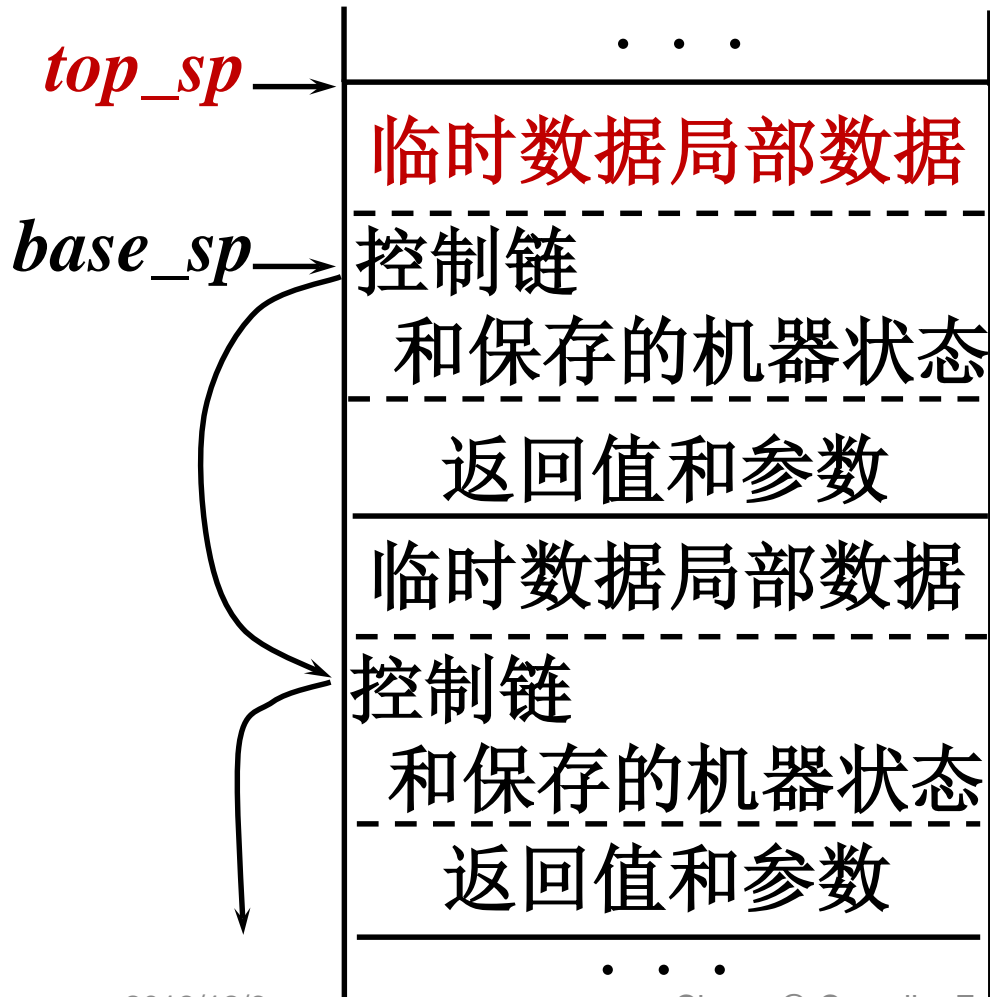
□ 过程p调用过程q的调用序列



(3) q保存寄存器的值和其它机器状态信息



□过程p调用过程q的调用序列

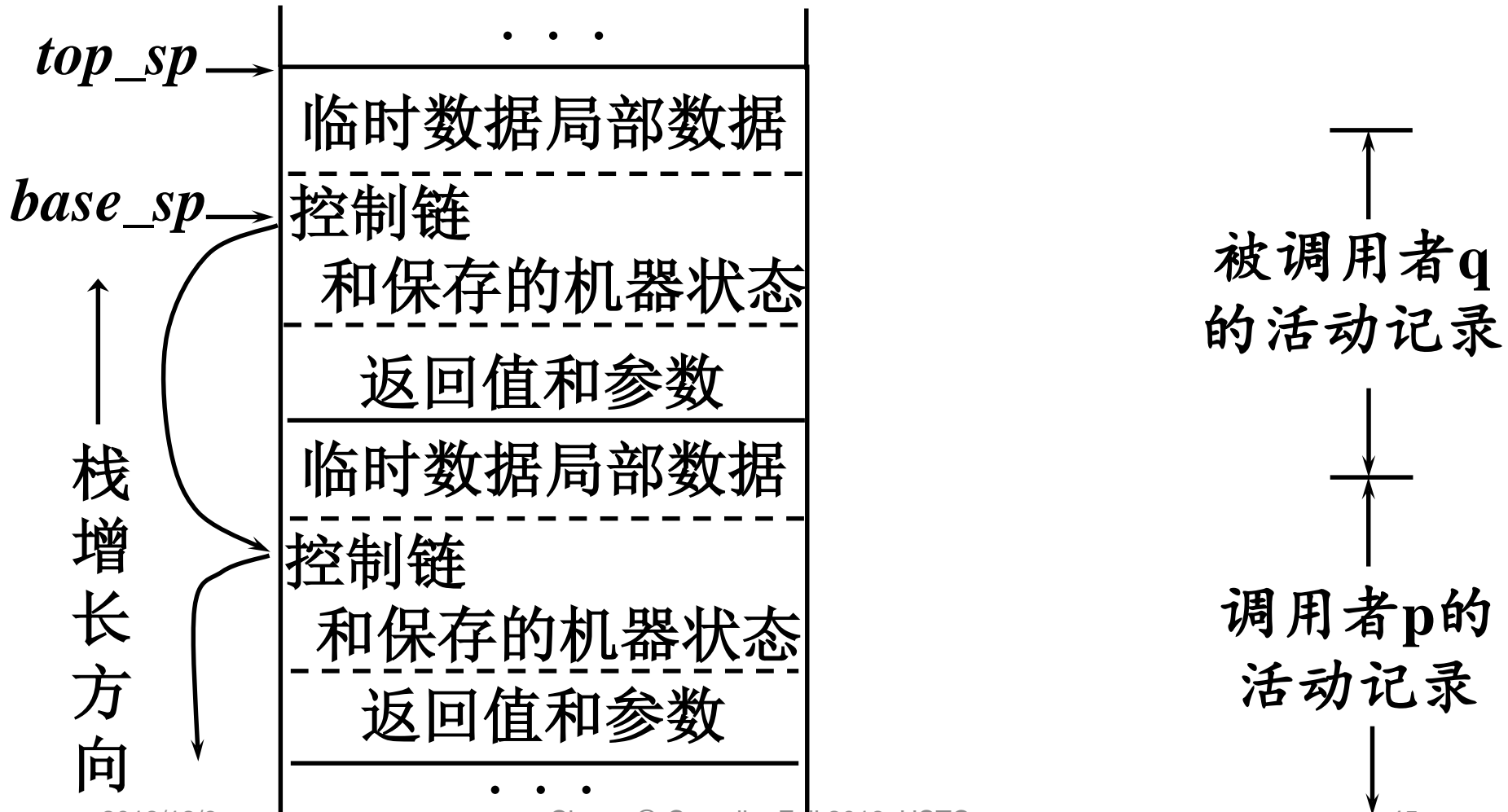


base_sp不变, 指向活动记录中间

(4) q根据局部数据域和临时数据域的大小减小 top_sp 的值, 初始化它的局部数据, 并开始执行过程体

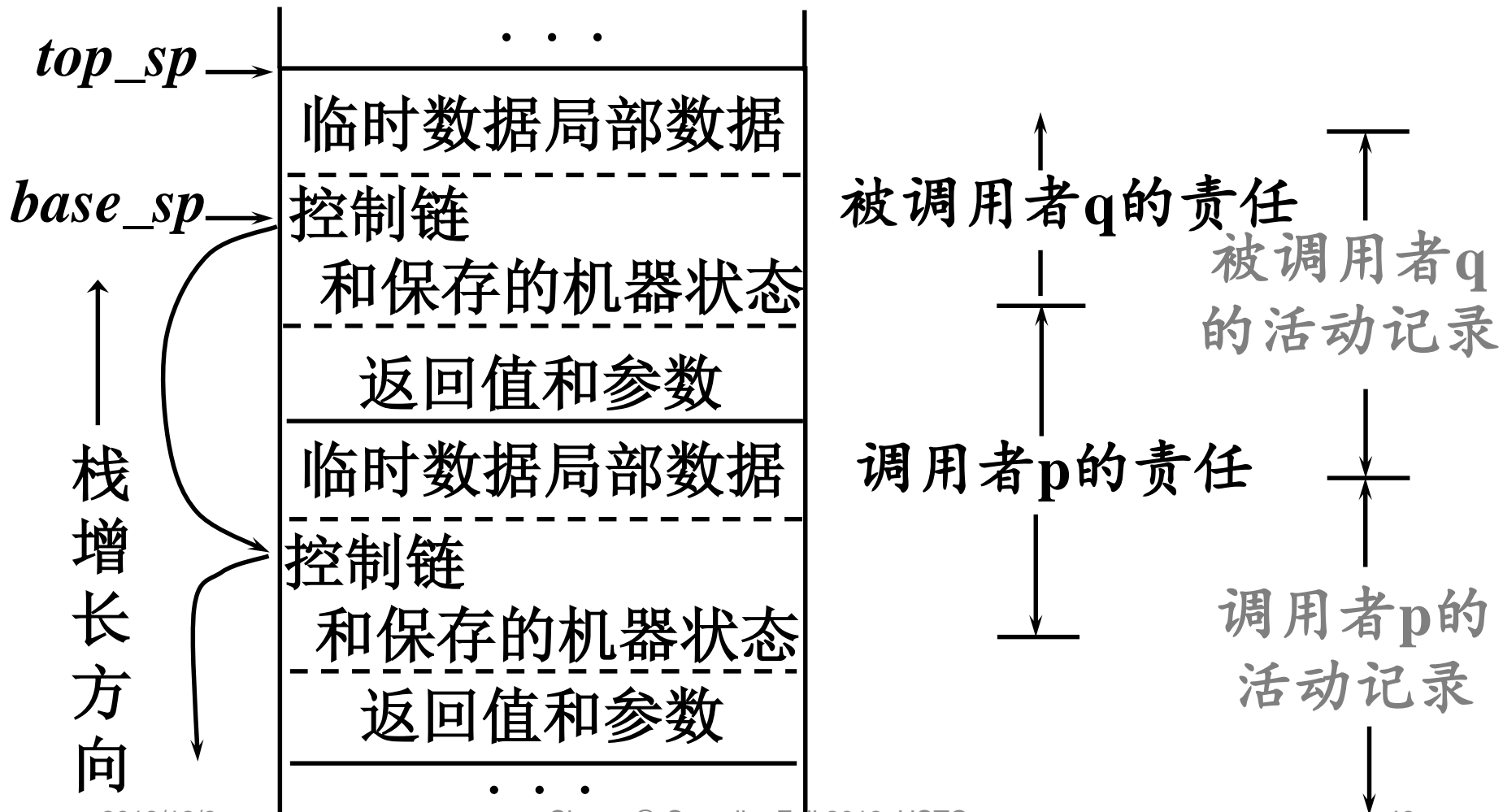


调用者p和被调用者q之间的任务划分



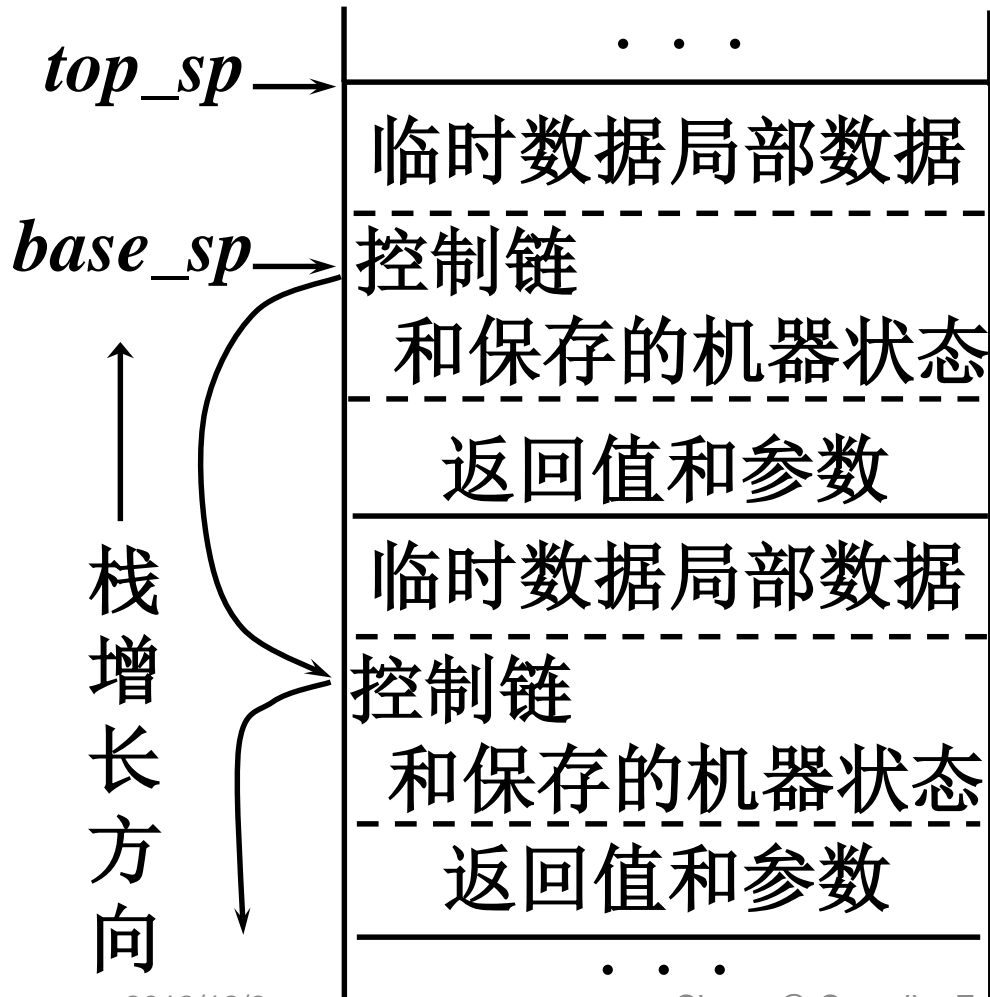


调用者p和被调用者q之间的任务划分



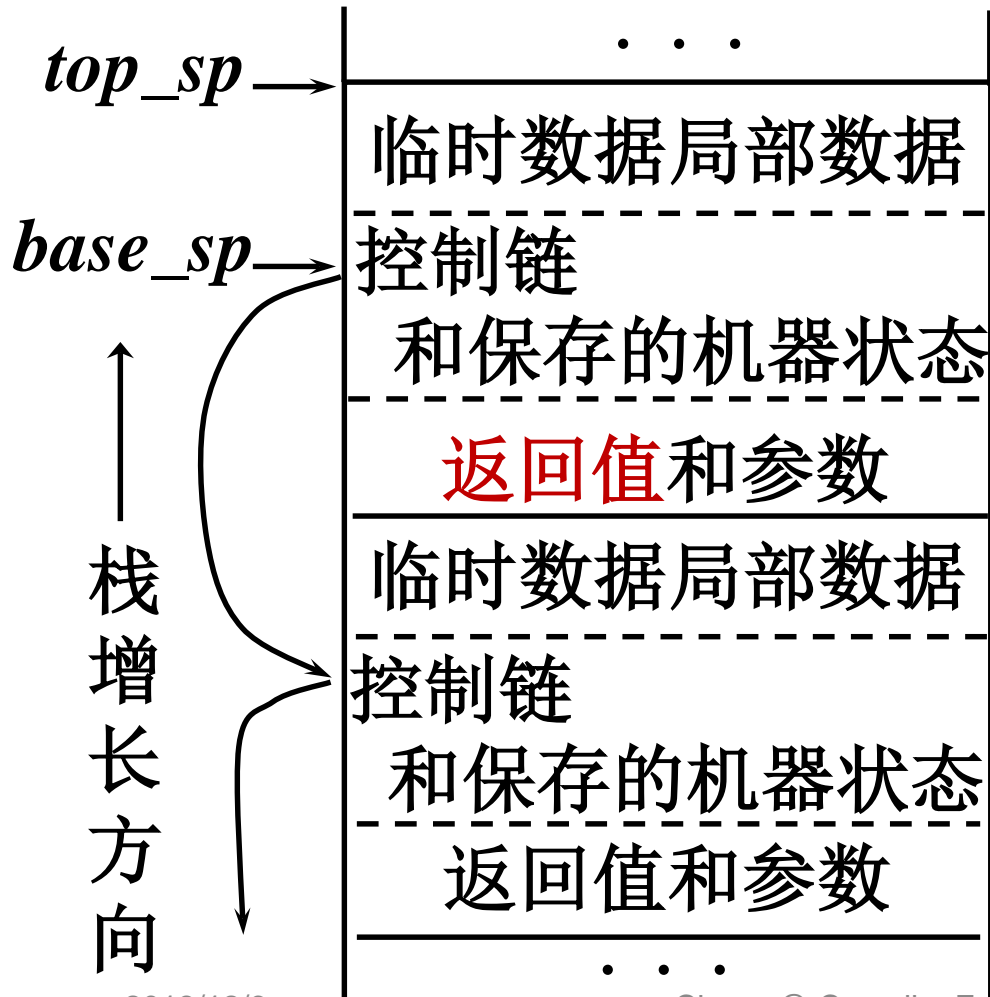


□ 过程p调用过程q的返回序列





□过程p调用过程q的返回序列

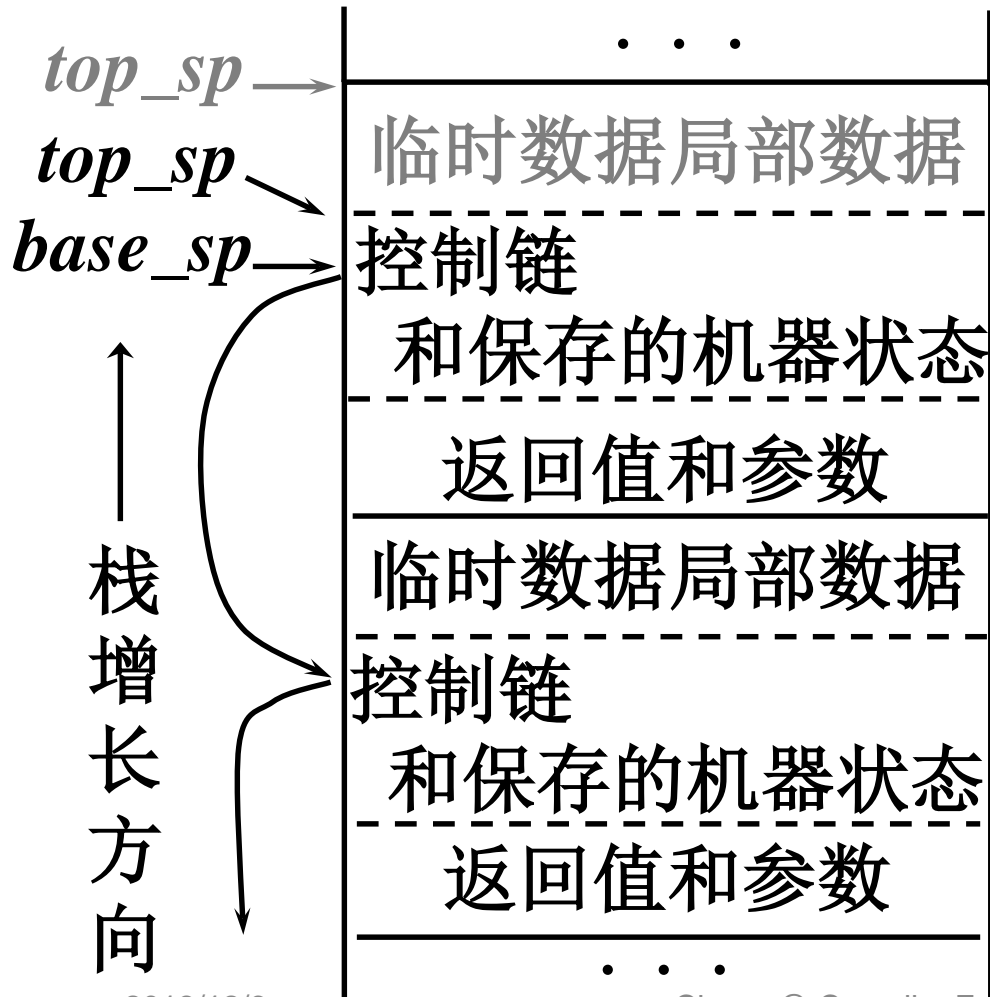


(1) q把返回值置入邻近p的活动记录的地方

引申：参数个数可变场合难以确定存放返回值的位置，因此通常用寄存器传递返回值



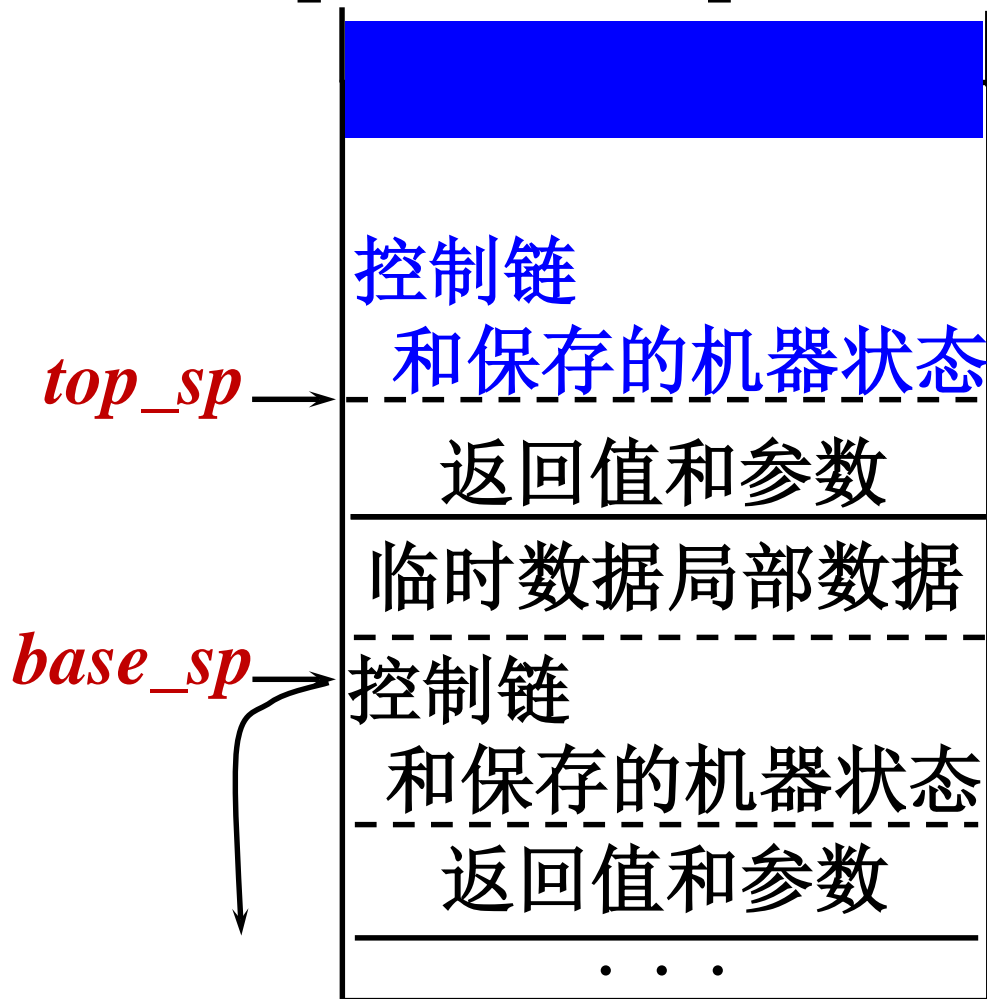
□ 过程p调用过程q的返回序列



(2) q对应调用序列的步骤(4), 增加 top_sp 的值



□ 过程p调用过程q的返回序列

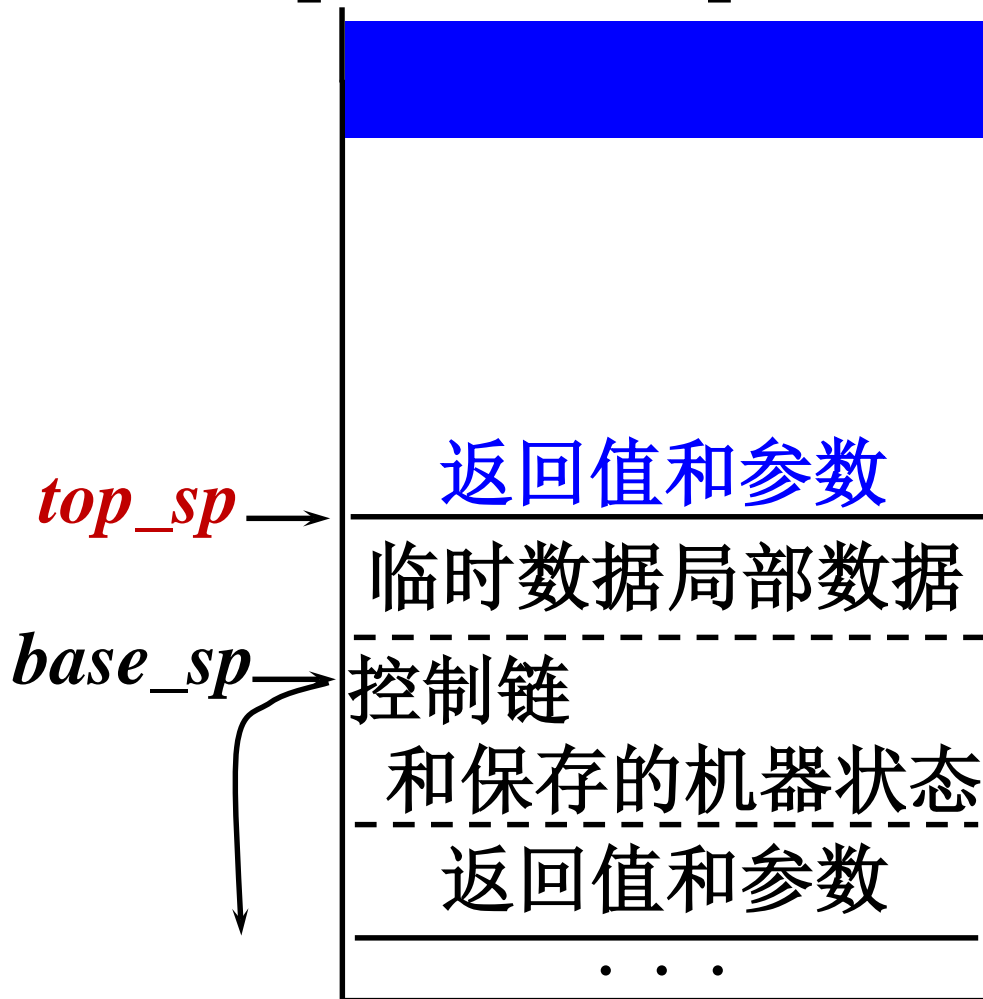


(3) q恢复寄存器(包括
`base_sp`)和机器状态,
返回p

控制权转到p



□ 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



有C程序如下：

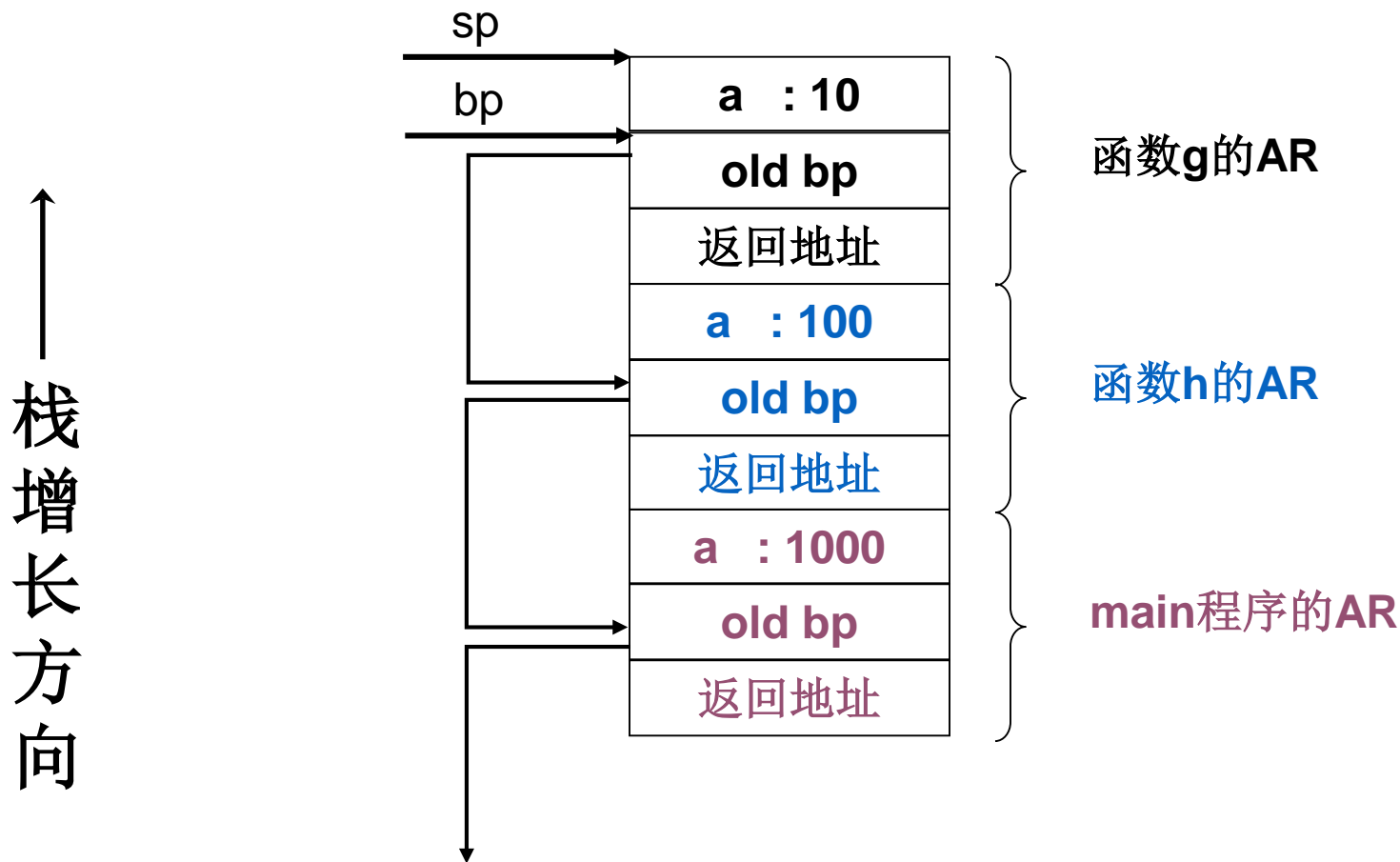
```
void g() { int a ; a = 10 ; }
```

```
void h() { int a ; a = 100; g(); }
```

```
main()
```

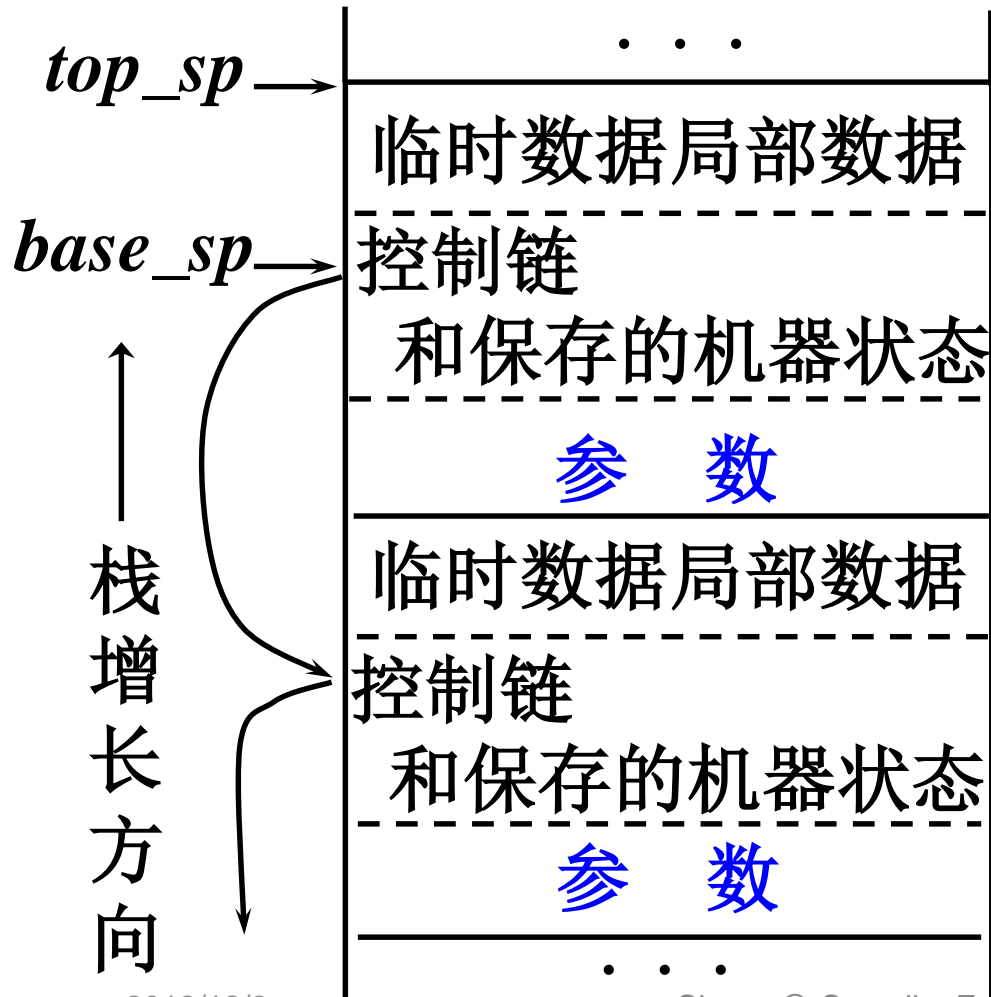
```
{ int a = 1000; h(); }
```

□过程g被调用时，活动记录栈的（大致）内容





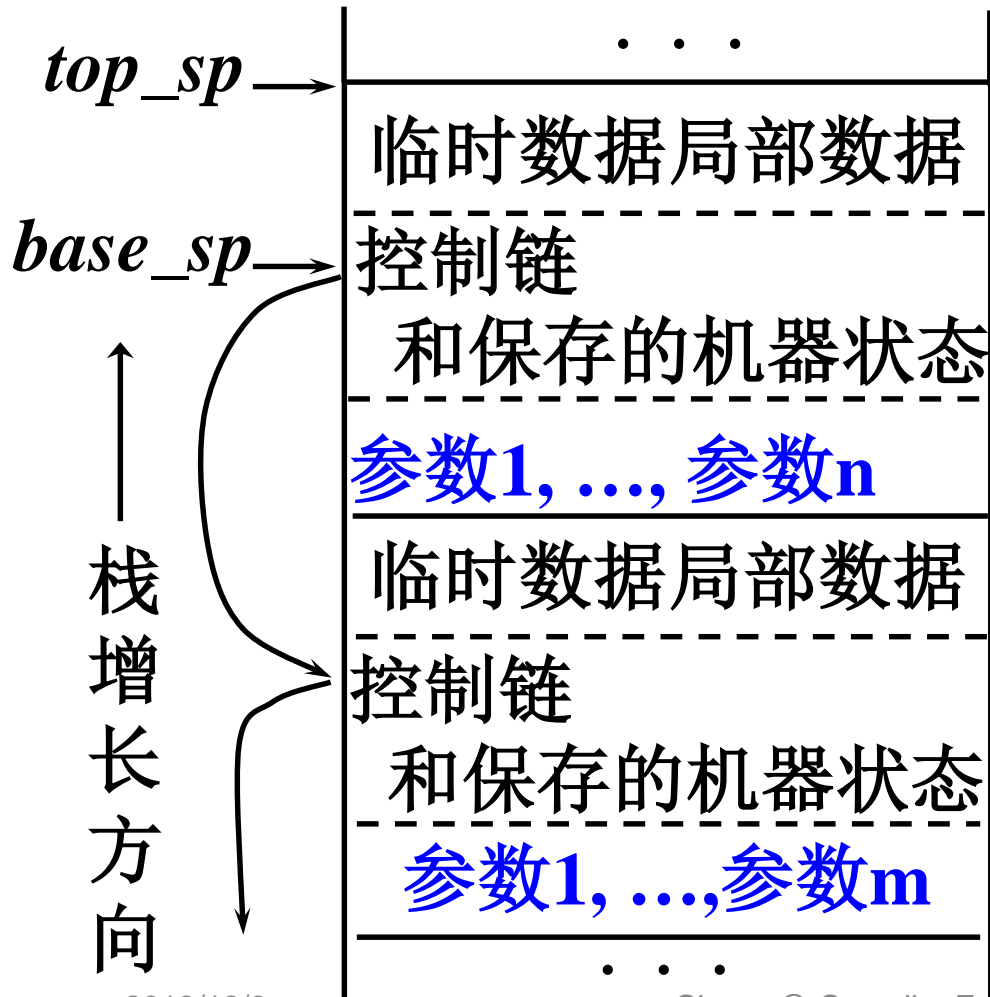
□过程的参数个数可变的情况



(1) 函数返回值改成用寄存器传递



□过程的参数个数可变的情况

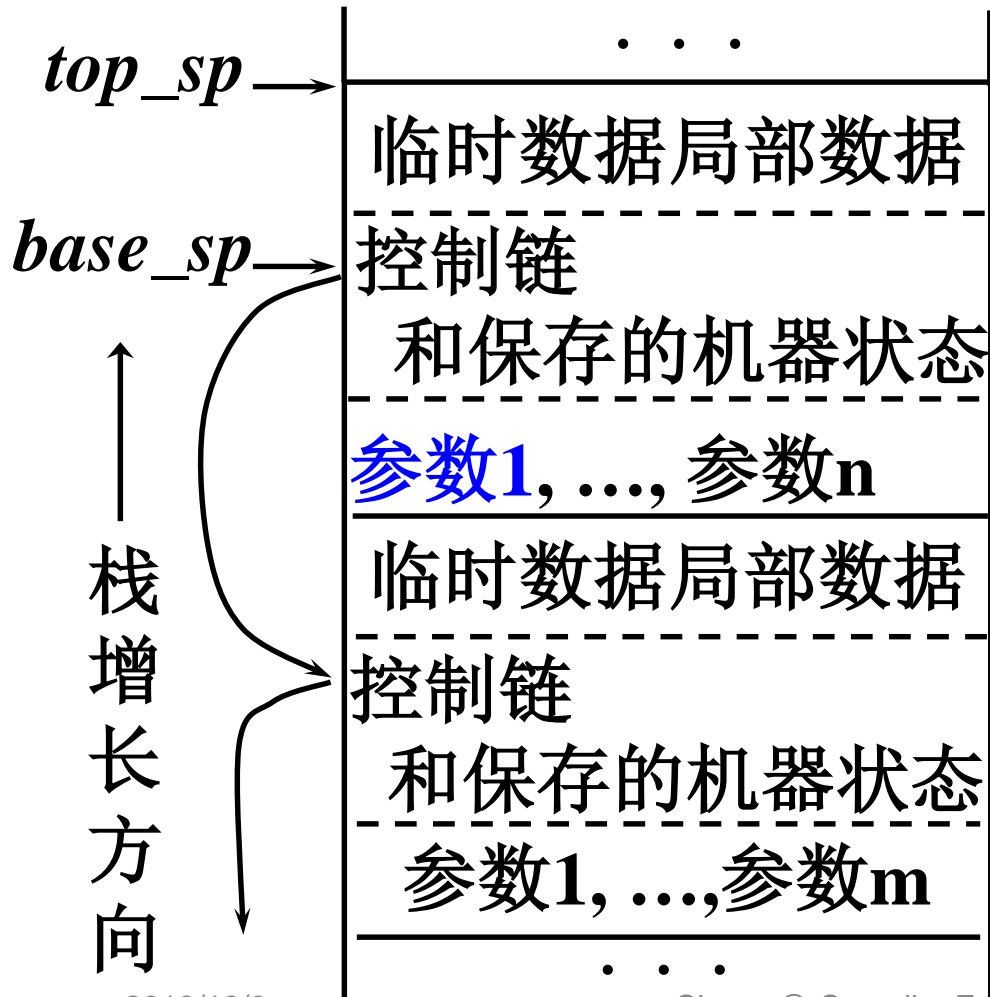


(2) 编译器产生将实参表达式逆序计算并将结果进栈的代码

自上而下依次是参数 1, ..., 参数n



□过程的参数个数可变的情况

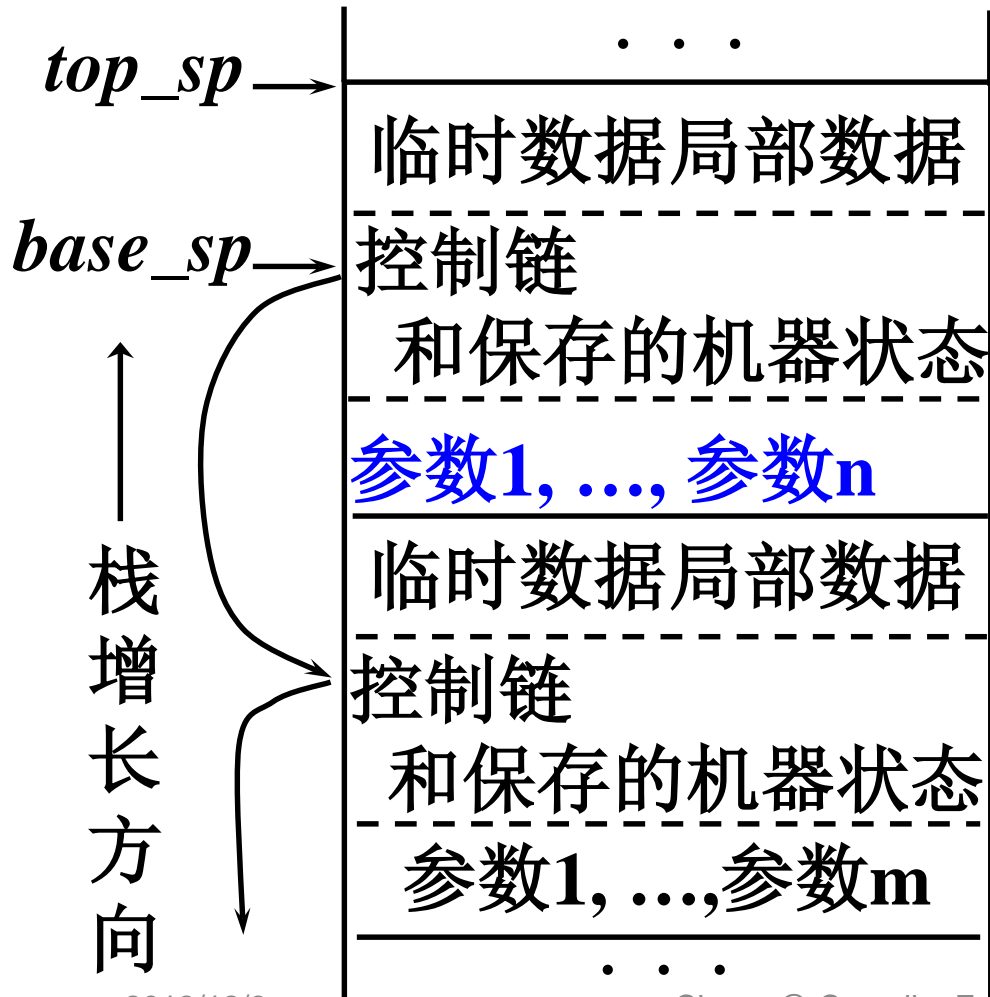


(3) 被调用函数能准确地知道第一个参数的位置

But why?



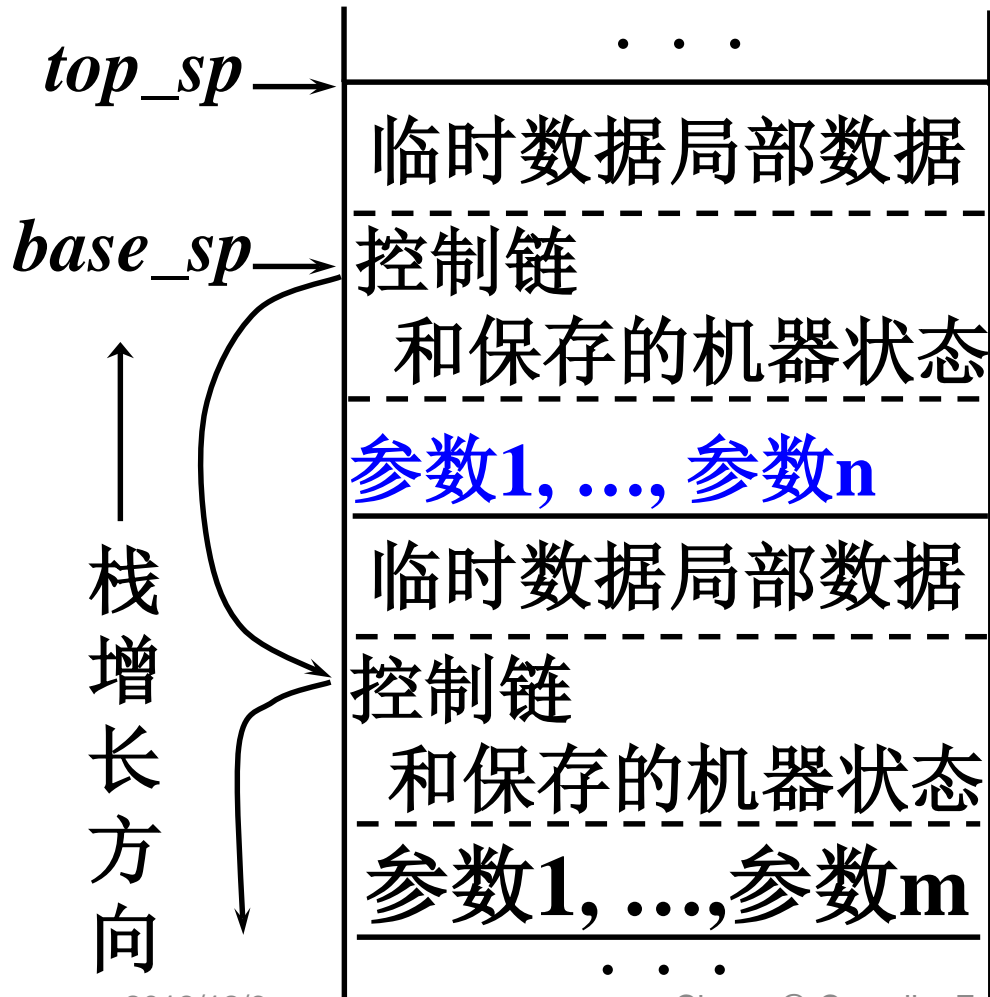
□过程的参数个数可变的情况



(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



□过程的参数个数可变的情况



C语言的printf函数就是按此方式，用C语言编写的

下面语句的输出？
`printf(“%d, %d, %d\n”);`



有参函数的活动记录



```
void func( int a , int b )  
{  
    int c , d;  
    c = a;  
    d = b;  
}
```



```
.file "ar.c"  
.text  
.globl func  
.type func,@function  
func:  
    pushl %ebp  
    movl %esp, %ebp  
    subl $8, %esp  
    movl 8(%ebp), %eax  
    movl %eax, -4(%ebp)  
    movl 12(%ebp), %eax  
    movl %eax, -8(%ebp)  
    leave  
    ret
```



```
void func( int a , int b )  
{  
    int c , d;  
    c = a;  
    d = b;  
}
```



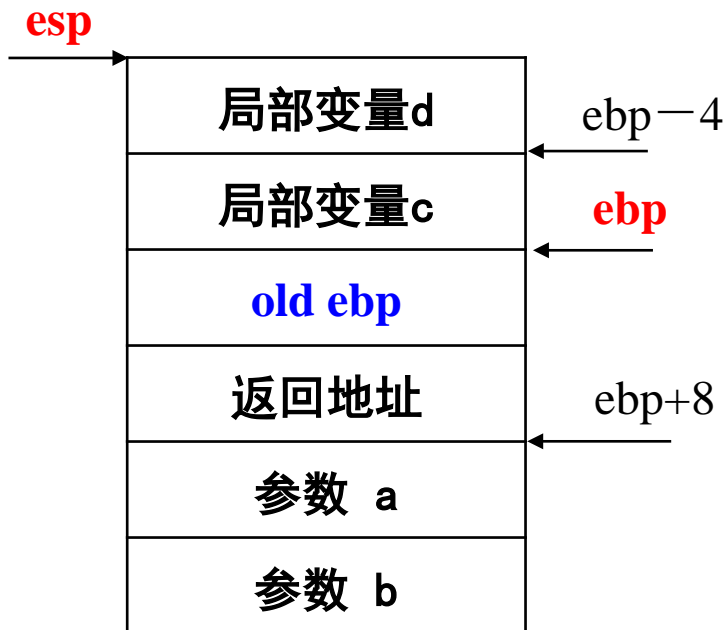
```
.file "ar.c"  
.text  
.globl func  
.type func,@function  
func:  
    pushl %ebp //老基地址压栈  
    movl %esp, %ebp //基地址指针=栈顶指针  
    subl $8, %esp  
    movl 8(%ebp), %eax  
    movl %eax, -4(%ebp)  
    movl 12(%ebp), %eax  
    movl %eax, -8(%ebp)  
    leave  
    ret
```



有参函数的活动记录



```
void func( int a , int b )
{
    int c , d;
    c = a;
    d = b;
}
```



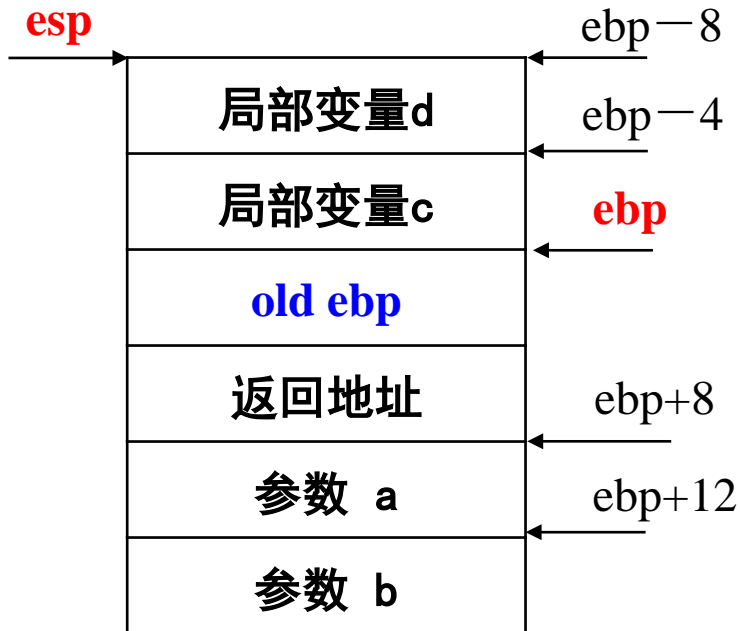
```
.file "ar.c"
.text
.globl func
.type func,@function
func:
    pushl %ebp //老基地址压栈
    movl %esp, %ebp //基地址指针=栈顶指针
    subl $8, %esp //分配c,d局部变量空间
    movl 8(%ebp), %eax //将a值放进寄存器
    movl %eax, -4(%ebp) //将a值赋给c
    movl 12(%ebp), %eax
    movl %eax, -8(%ebp)
    leave
    ret
```



有参函数的活动记录



```
void func( int a , int b )
{
    int c , d;
    c = a;
    d = b;
}
```



```
.file "ar.c"
.text
.globl func
.type func,@function
func:
    pushl %ebp //老基地址压栈
    movl %esp, %ebp //基地址指针=栈顶指针
    subl $8, %esp //分配c,d局部变量空间
    movl 8(%ebp), %eax //将a值放进寄存器
    movl %eax, -4(%ebp) //将a值赋给c
    movl 12(%ebp), %eax //将b值放进寄存器
    movl %eax, -8(%ebp) //将b值赋给d
    leave
    ret
```

有如下C程序：

```
main()
```

```
{
```

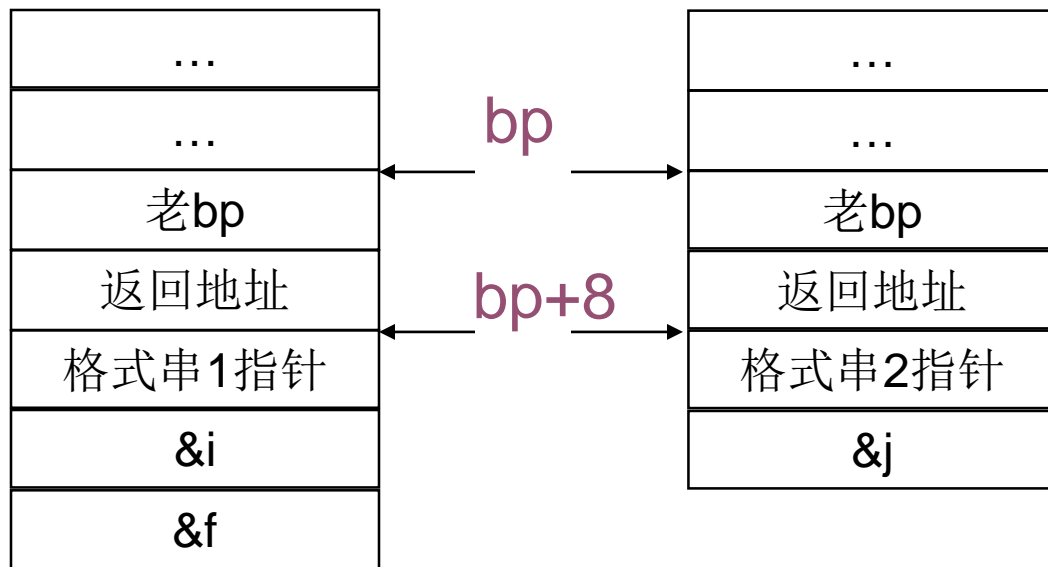
```
int i ; float f; int j ;
```

```
scanf(“%d%f”, &i, &f); //第一次调用时3个参数
```

```
scanf(“%d”, &j); //第二次调用时 2个参数
```

```
return 0;
```

```
}
```



scanf的第一次调用时AR

scanf的第二次调用时AR

由于C语言采用**逆序**传递参数，格式串参数将被放在AR中的“固定”位置，即**bp+8**。而由此参数即可确定待输入值的参数（变量）的个数。从而适应参数个数变化的情况。



□ 栈上可变长数据

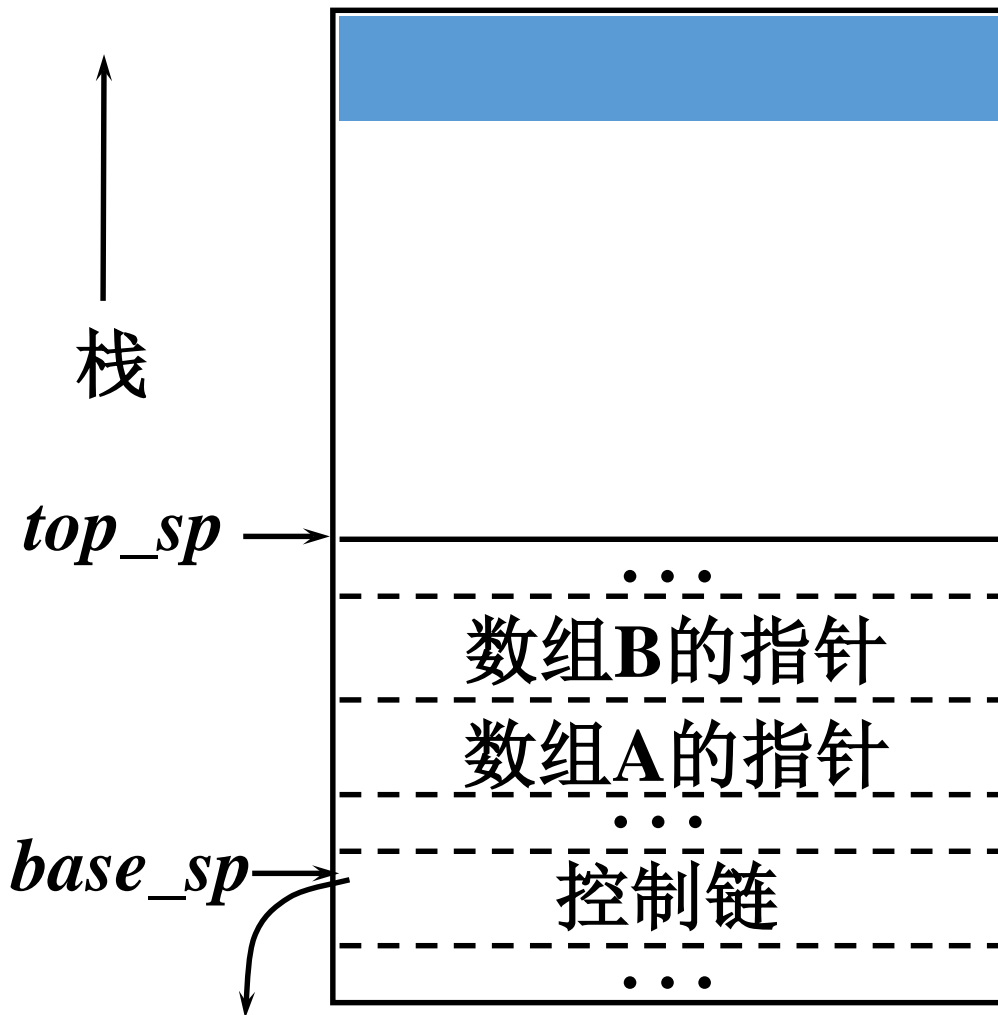
- ❖ 活动记录的长度在编译时不能确定的情况
- ❖ 但仅仅为改活动运行过程使用
- ❖ 例：局部数组的大小要等到过程激活时才能确定

□ 如何在栈上布局可变长的数组？

- ❖ 先分配存放数组指针的单元，对数组的访问通过指针间接访问
- ❖ 运行时，这些指针指向分配在栈顶的数组存储空间



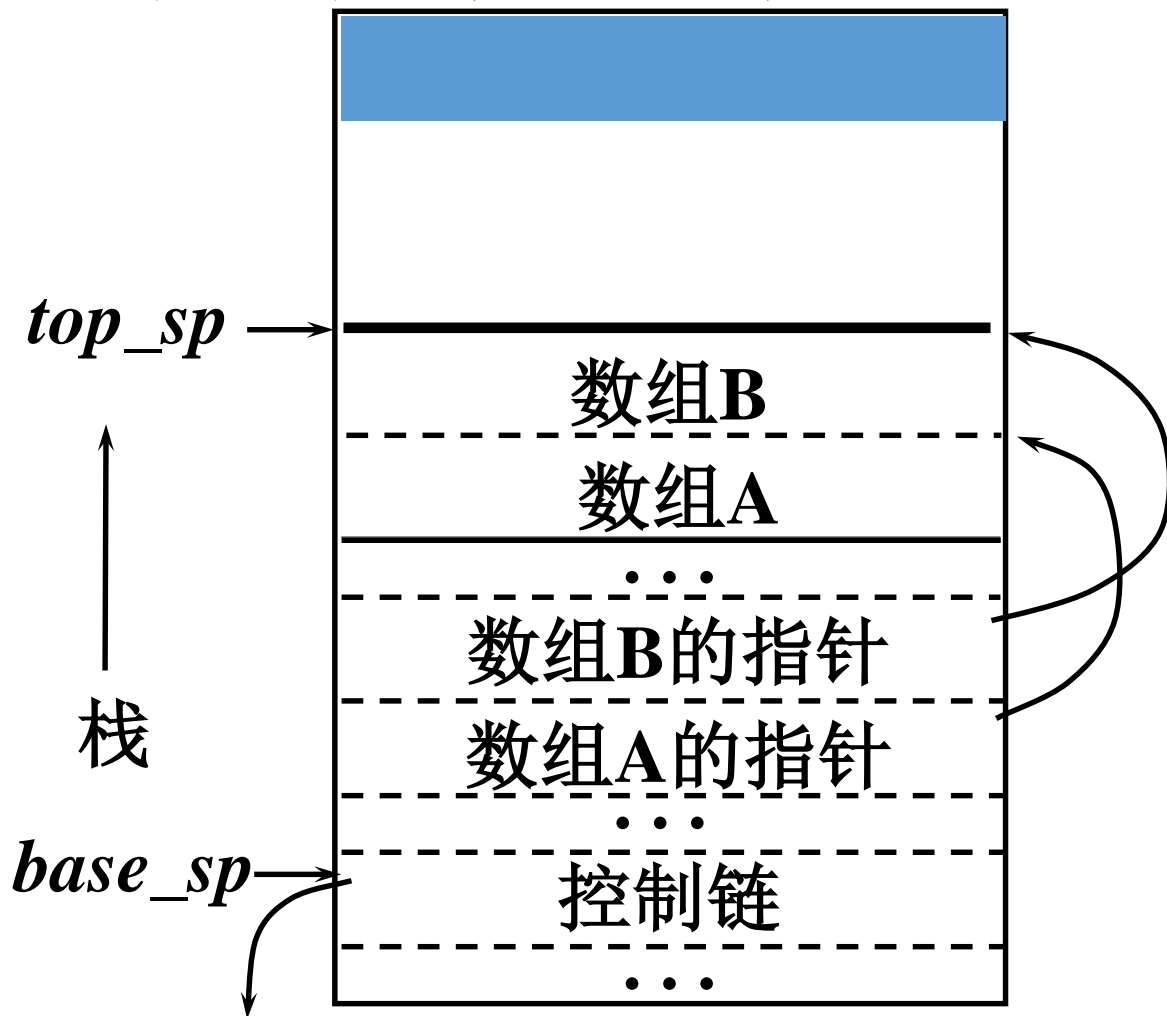
□访问动态分配的数组



(1) 编译时，在活动记录中为这样的数组分配存放数组指针的单元



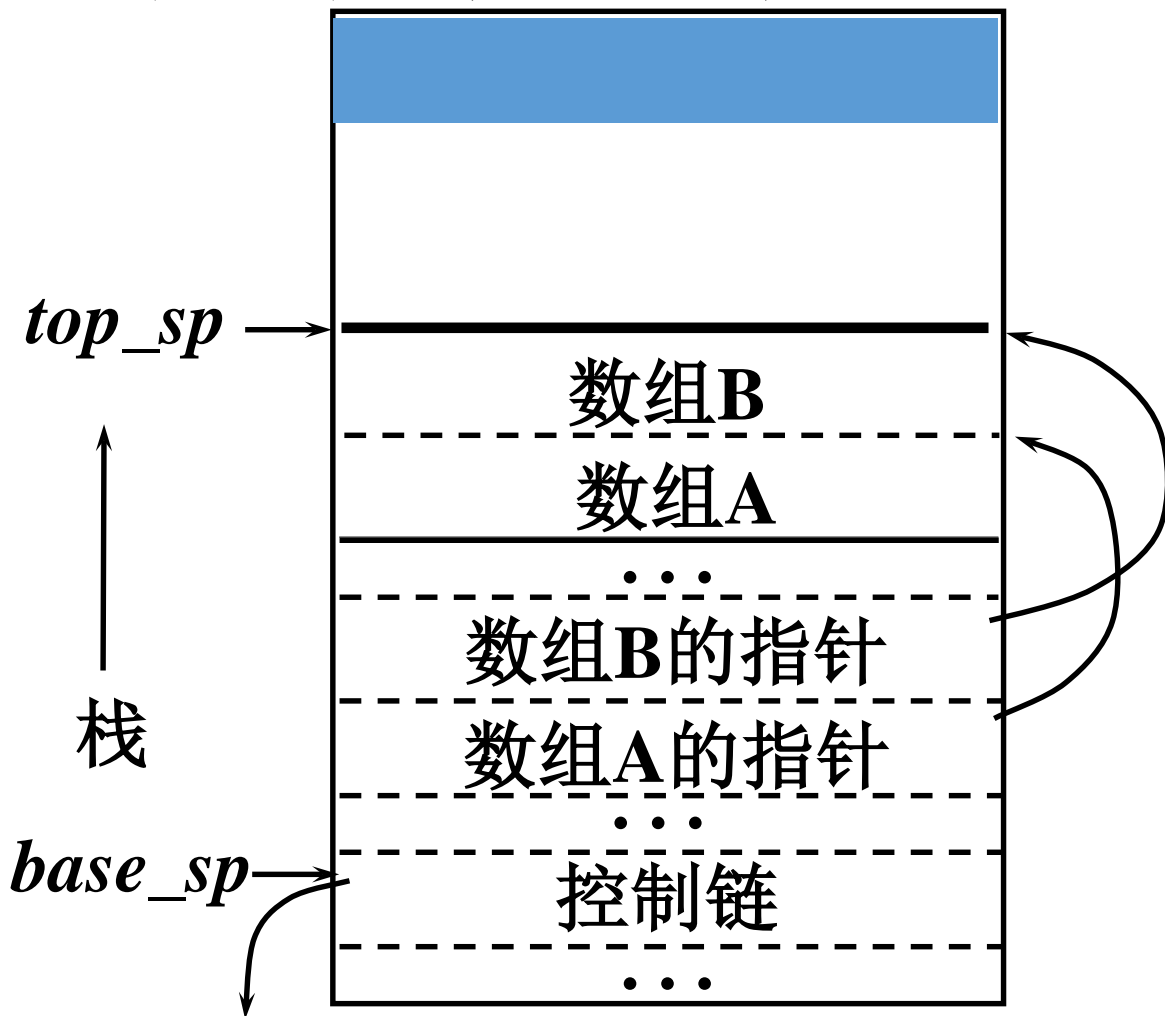
□访问动态分配的数组



(2) 运行时，这些指针指向分配在栈顶的数组存储空间



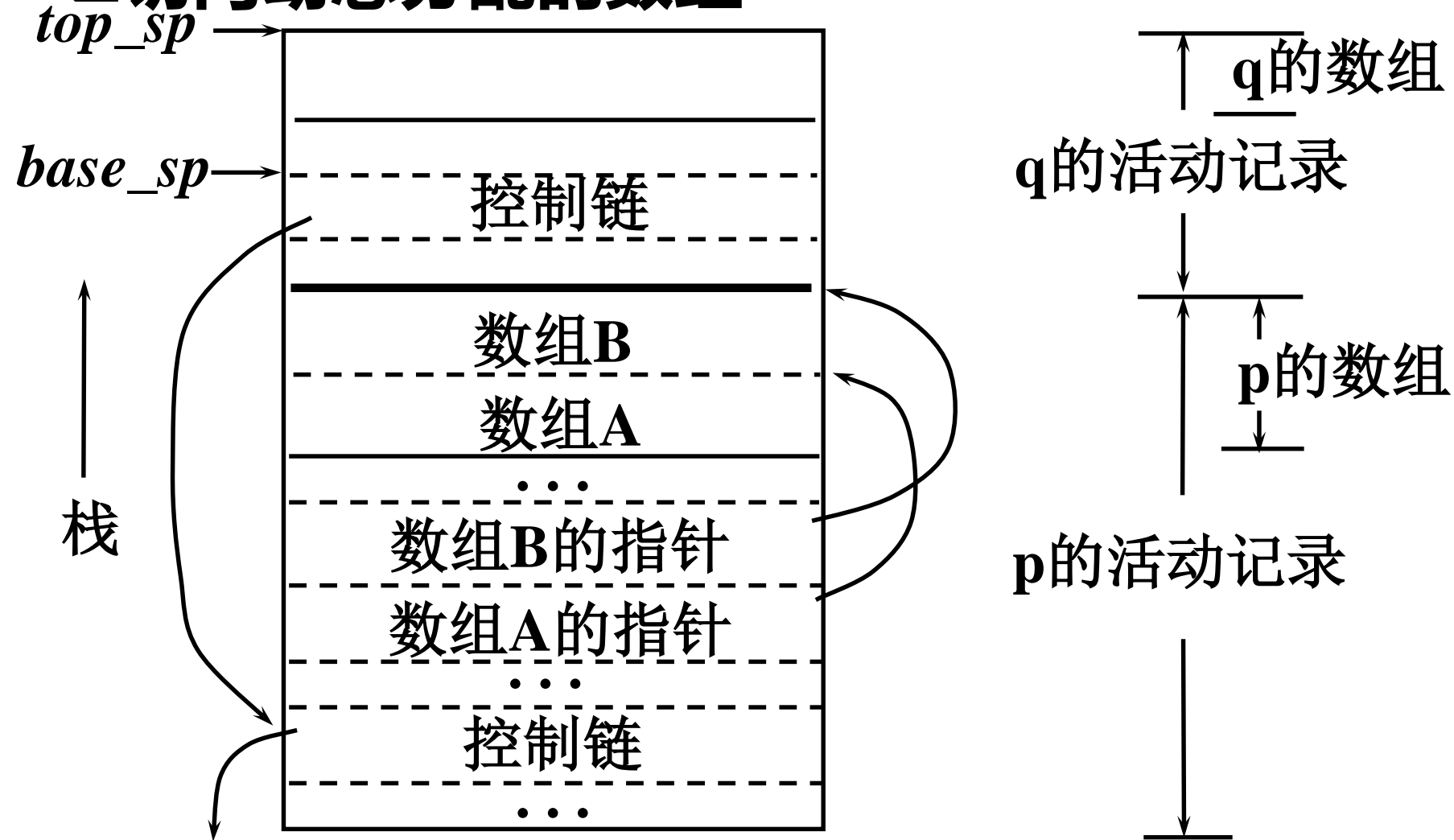
□访问动态分配的数组



(3) 运行时，对数组A和B的访问都要通过相应指针来间接访问



访问动态分配的数组





□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;  
} | }
```



- 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)



□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

□ 有过程嵌套的静态作用域

sort

readarray

exchange

quicksort

partition



□ 有过程嵌套的静态作用域

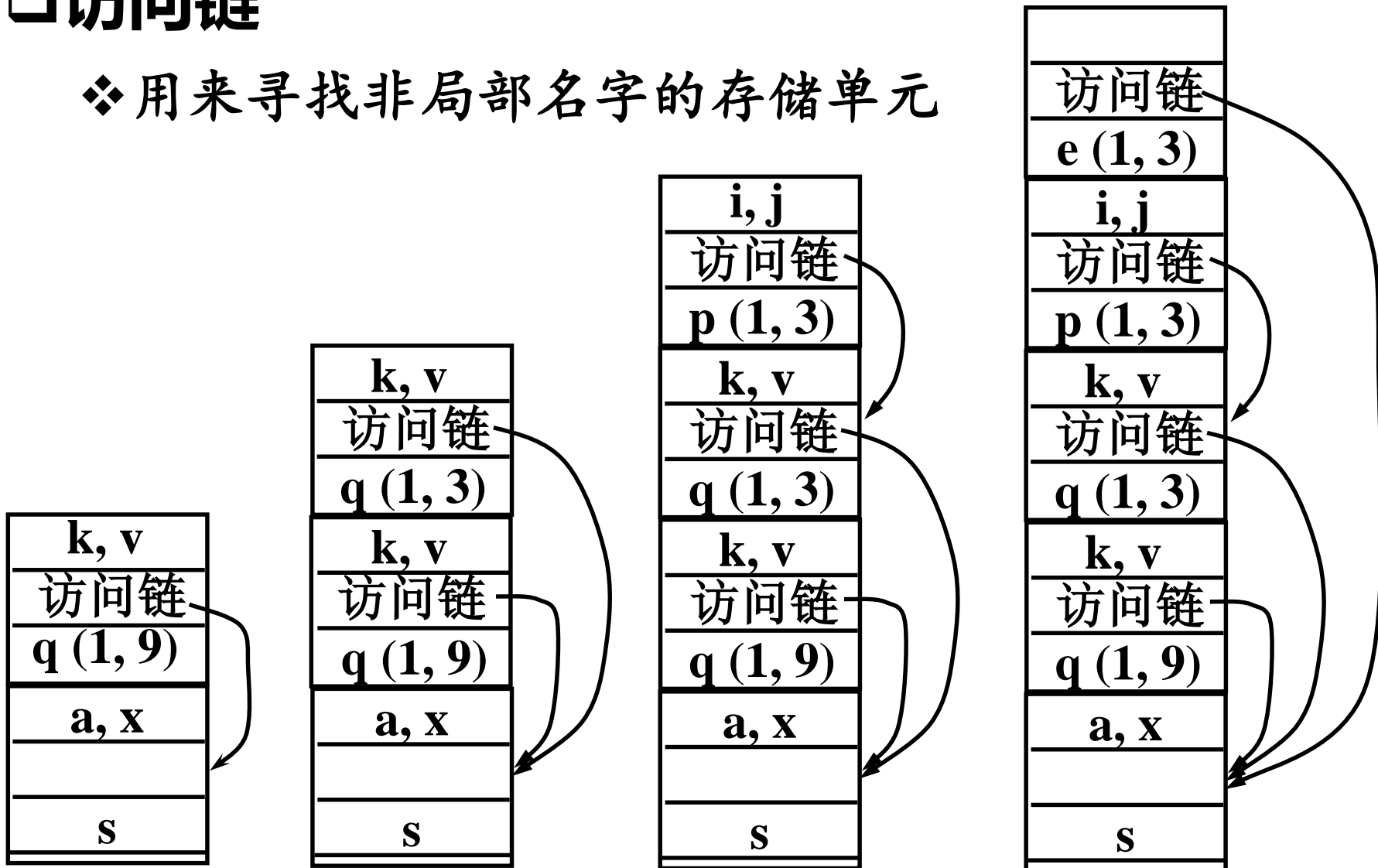
❖ 过程嵌套深度

❖ 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

<code>sort</code>	1
<code> readarray</code>	2
<code> exchange</code>	2
<code> quicksort</code>	2
<code>partition</code>	3

访问链

❖ 用来寻找非局部名字的存储单元





□两个关键问题需要解决：

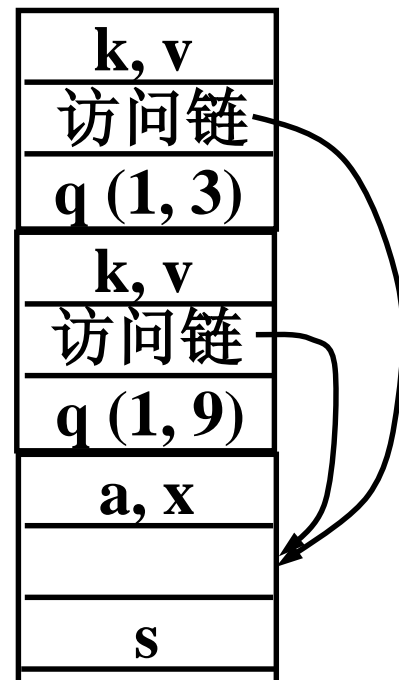
- ❖通过访问链访问非局部引用
- ❖访问链的建立



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



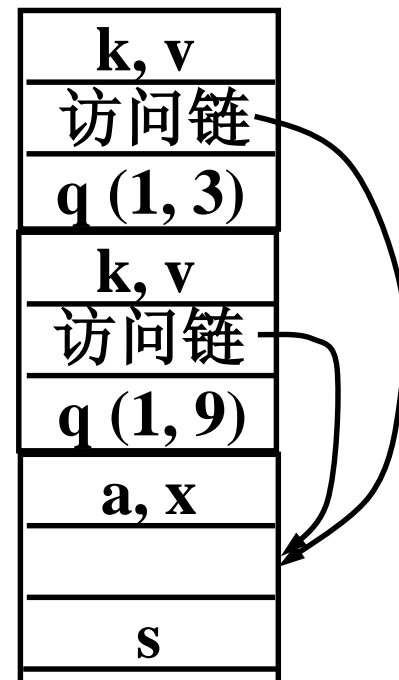


访问非局部名字的存储单元

❖ 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

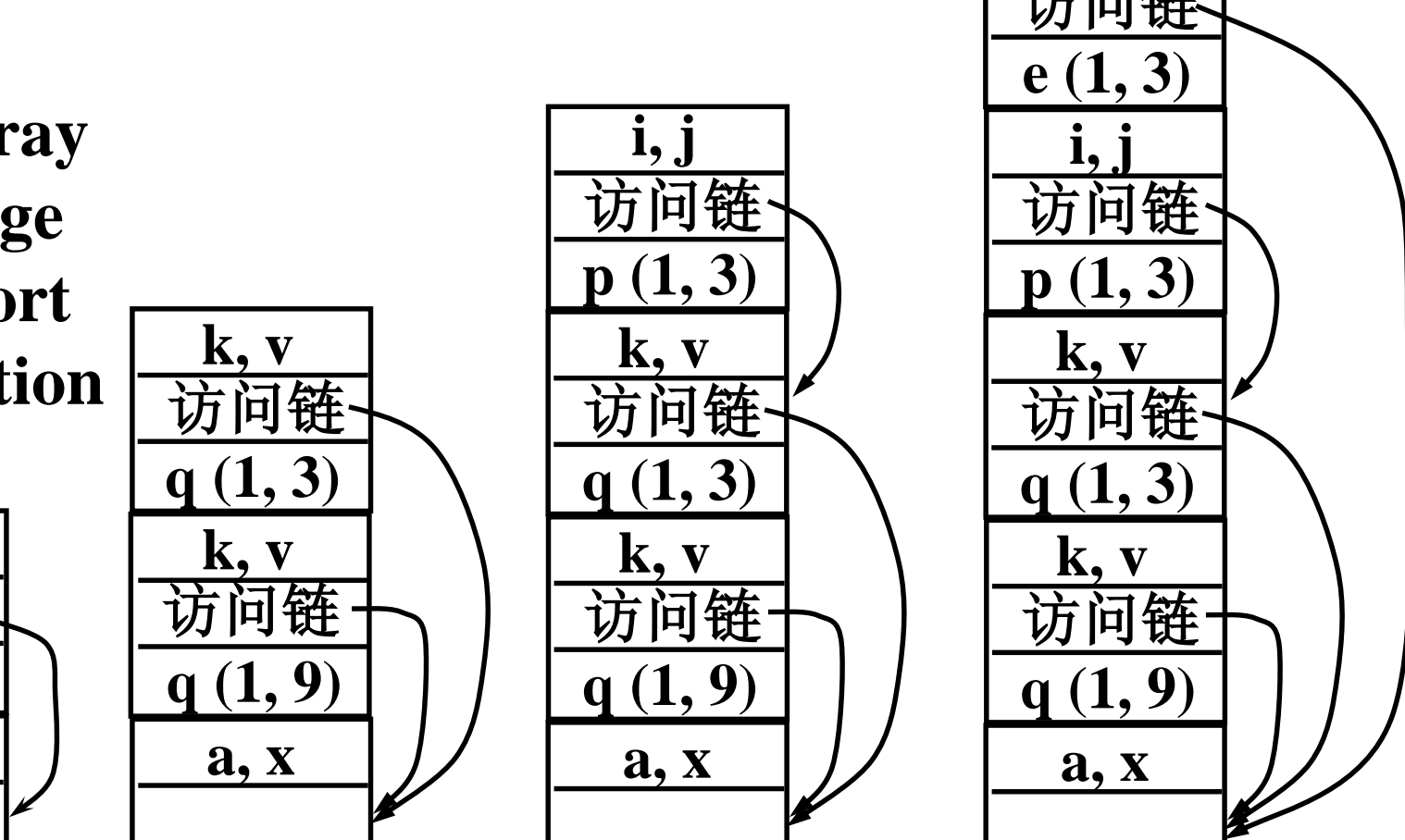
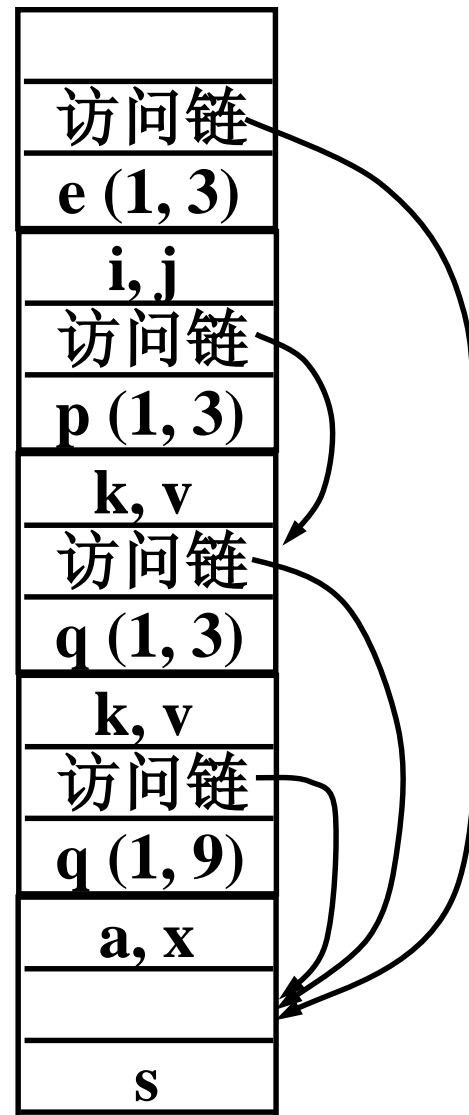
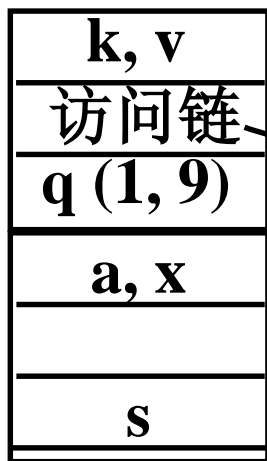
- 从栈顶的活动记录开始，追踪访问链 $n_p - n_a$ 次
- 到达 a 的声明所在过程的活动记录
- 访问链的追踪用间接操作就可完成

sort	1
readarray	2
exchange	2
quicksort	2
partition	3



访问非局部名字的存储单元

sort
readarray
exchange
quicksort
partition



□ 过程 p 对变量 a 访问时， a 的地址由下面的二元组表示：

$(n_p - n_a, a \text{ 在活动记录中的偏移})$

□建立访问链(过程调用序列代码的一部分)

❖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



□建立访问链

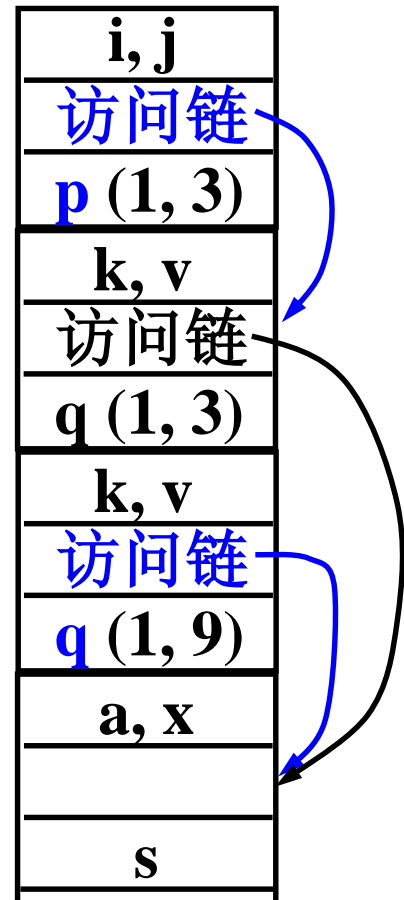
❖ 假定嵌套深度为 n_p 的过程 p 调用嵌套深度为 n_x 的过程 x

(1) $n_p < n_x$ 的情况

❖ 这时 x 肯定就声明在 p 中 (嵌套)

❖ 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

❖ sort 调用 quicksort、quicksort 调用 partition



□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
readarray	2
exchange	2
quicksort	2
partition	3

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



□建立访问链

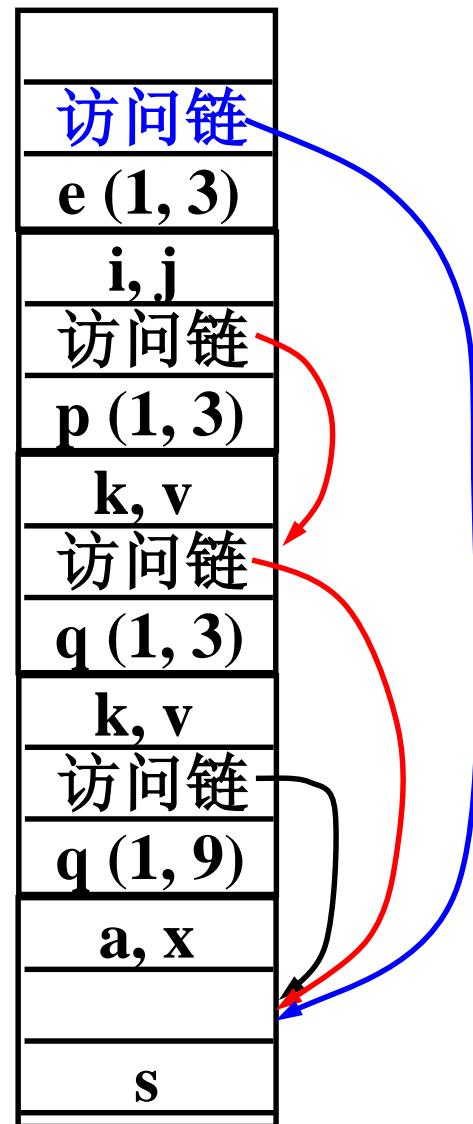
❖ 假定嵌套深度为 n_p 的过程 p 调用嵌套深度为 n_x 的过程 x

(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

❖ partition 调用 exchange





中国科学技术大学
University of Science and Technology of China



《编译原理与技术》

运行时存储空间的组织和管理

TBA