

HANDLING MULTIPLE SCOPES

Many languages allow procedure declarations to nest. Java now allows classes to nest.

Procedure nesting can be very useful, allowing a subroutine to directly access another routine's locals and parameters.

Run-time data structures are complicated because multiple frames, corresponding to nested procedure declarations, may need to be accessed.

To see the difficulties, assume that routines *can* nest in Java or C:

```
int p(int a){
    int q(int b){
        if (b < 0)
            return q(-b);
        else
            return a+b;
    }
    return q(-10);
}
```

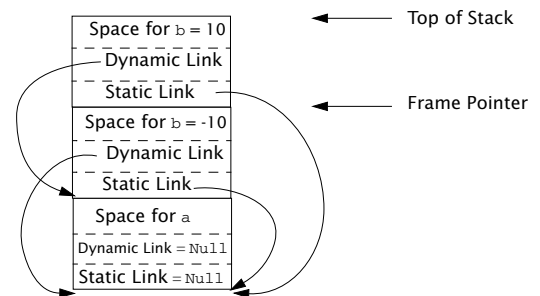
When **q** executes, it may access not only its own frame, but also that of **p**, in which it is nested.

If the depth of nesting is unlimited, so is the number of frames that must be accessible. In practice, the level of nesting actually seen is modest—usually no greater than two or three.

STATIC LINKS

Two approaches are commonly used to support access to multiple frames. One approach generalizes the idea of dynamic links introduced earlier. Along with a dynamic link, we'll also include a *static link* in the frame's control information area. The static link points to the frame of the procedure that statically encloses the current procedure. If a procedure is not nested within any other procedure, its static link is **null**.

The following illustrates static links:



As usual, dynamic links always point to the next frame down in the stack. Static links always point down, but they may skip past many frames. They always point to the most recent frame of the routine that statically encloses the current routine.

In our example, the static links of both of q 's frames point to p , since it is p that encloses q 's definition.

In evaluating the expression $a+b$ that q returns, b , being local to q , is accessed directly through the frame pointer. Variable a is local to p , but also visible to q because q nests within p . a is accessed by extracting q 's static link, then using that address (plus the appropriate offset) to access a .

Displays

An alternative to using static links to access frames of enclosing routines is the use of a *display*.

A display generalizes our use of a frame pointer. Rather than maintaining a single register, we maintain a *set of registers* which comprise the display.

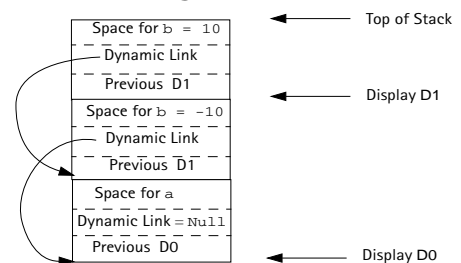
If procedure definitions nest n deep (this can be easily determined by examining a program's AST), we need $n+1$ display registers.

Each procedure definition is tagged with a nesting level. Procedures not nested within any other routine are at level 0. Procedures nested within only one routine are at level 1, etc.

Frames for routines at level 0 are always accessed using display register D0. Those at level 1 are always accessed using register D1, etc.

Whenever a procedure r is executing, we have direct access to r 's frame plus the frames of all routines that enclose r . Each of these routines must be at a different nesting level, and hence will use a different display register.

The following illustrates the use of display registers:



Since q is at nesting level 1, its frame is pointed to by D1. All of q 's local variables, including b , are at a fixed offset relative to D1.

Since p is at nesting level 0, its frame and local variables are accessed via D0. Each frame's control information area contains a slot for the previous value of the frame's display register. A display register is saved when a call

begins and restored when the call ends. A dynamic link is still needed, because the previous display values doesn't always point to the caller's frame.

Not all compiler writers agree on whether static links or displays are better to use. Displays allow direct access to all frames, and thus make access to all visible variables very efficient. However, if nesting is deep, several valuable registers may need to be reserved. Static links are very flexible, allowing unlimited nesting of procedures. However, access to non-local procedure variables can be slowed by the need to extract and follow static links.

HEAP MANAGEMENT

A very flexible storage allocation mechanism is *heap allocation*.

Any number of data objects can be allocated and freed in a memory pool, called a *heap*.

Heap allocation is enormously popular. Almost all non-trivial Java and C programs use **new** or **malloc**.

HEAP ALLOCATION

A request for heap space may be *explicit* or *implicit*.

An explicit request involves a call to a routine like **new** or **malloc**. An explicit pointer to the newly allocated space is returned.

Some languages allow the creation of data objects of unknown size. In Java, the **+** operator is overloaded to represent string catenation.

The expression **str1 + str2** creates a new string representing the catenation of strings **str1** and **str2**. There is no compile-time bound on the sizes of **str1** and **str2**, so heap space must be implicitly allocated to hold the newly created string.

Whether allocation is explicit or implicit, a *heap allocator* is needed. This routine takes a size parameter and examines unused heap space to find space that satisfies the request.

A *heap block* is returned. This block must be big enough to satisfy the space request, but it may well be bigger.

Heaps blocks contain a *header* field that contains the size of the block as well as bookkeeping information.

The complexity of heap allocation depends in large measure on how *deallocation* is done.

Initially, the heap is one large block of unallocated memory. Memory requests can be satisfied by simply modifying an "end of

heap” pointer, very much as a stack is pushed by modifying a stack pointer.

Things get more involved when previously allocated heap objects are deallocated and reused.

Deallocated objects are stored for future reuse on a *free space list*.

When a request for n bytes of heap space is received, the heap allocator must search the free space list for a block of sufficient size. There are many search strategies that might be used:

- **Best Fit**

The free space list is searched for the free block that matches most closely the requested size. This minimizes wasted heap space, the search may be quite slow.

- **First Fit**

The first free heap block of sufficient size is used. Unused space within the block is split off and linked as a smaller free space block. This approach is fast, but may “clutter” the beginning of the free space list with a number of blocks too small to satisfy most requests.

- **Next Fit**

This is a variant of first fit in which succeeding searches of the free space list begin at the position where the last search ended. The idea is to “cycle through” the entire free space list rather than always revisiting free blocks at the head of the list.

- **Segregated Free Space Lists**

There is no reason why we must have only *one* free space list. An alternative is to have several, indexed by the size of the free blocks they contain.