Lecture 9: Some practical aspects of compiling C

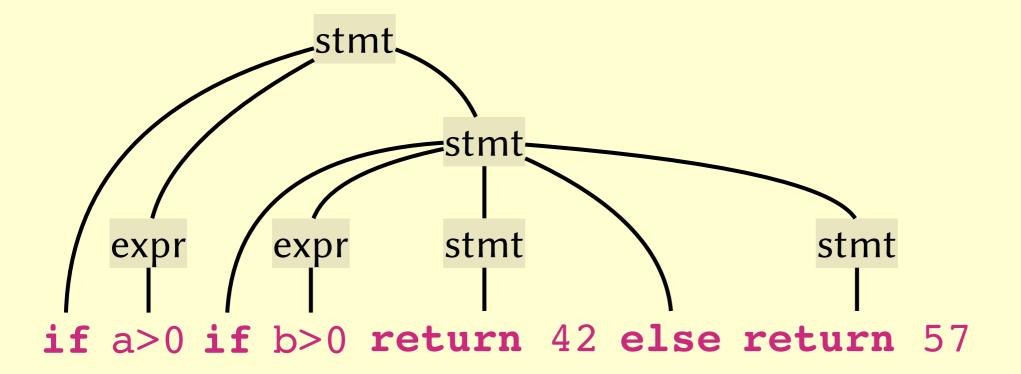
John Wickerson

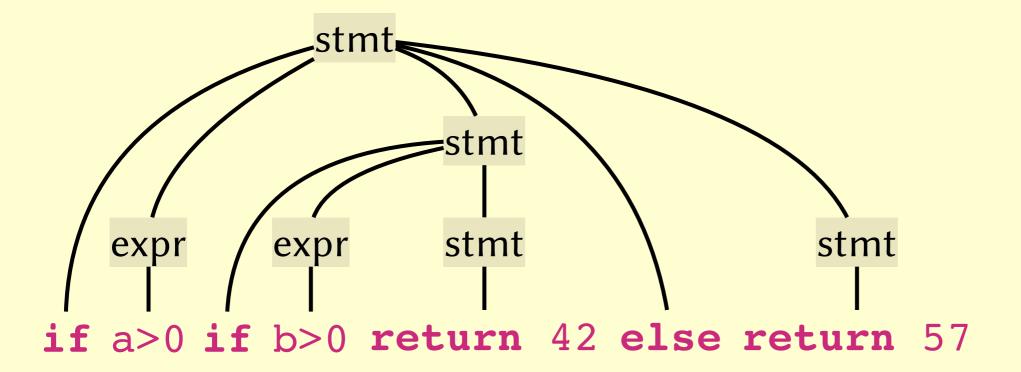
Lecture outline

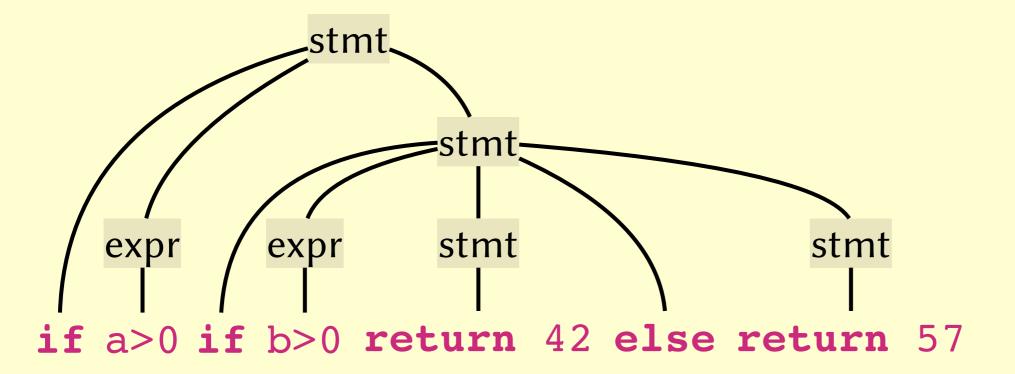
- Parsing C: the *dangling-else* problem.
- Parsing C: the *lexer hack*.
- Generating RISC-V assembly for function calls.

```
if (a > 0)
   if (b > 0)
    return 42;
else
   return 57;
```

```
if (a > 0)
  if (b > 0)
    return 42;
else
  return 57;
```







Lecture outline

- Parsing C: the *dangling-else* problem.
- Parsing C: the *lexer hack*.
- Generating RISC-V assembly for function calls.

C is not context-free!

```
int a, b;
```

C is not context-free!

typedef int a;

a * b;

The lexer hack

• When the parser sees a typedef instruction, e.g.

```
typedef int a;
it adds a to a list of 'type variables'.
```

- When the lexer sees a token that looks like an identifier, it looks up the matched string in this list. If it is there, it returns a **TYPEID** token, and otherwise it returns an **ID** token.
- Thus the parser can distinguish between
 TYPEID * ID; // declaring a pointer to a typedef'd type and

```
ID * ID; // a multiplication expression
```

Lecture outline

- Parsing C: the *dangling-else* problem.
- Parsing C: the *lexer hack*.
- Generating RISC-V assembly for function calls.

Function calls

```
int f (int a, int b,
       int c, int d) {
                                   t = 6;
  int t=6;
                                   u = 7;
  int u=7;
                                   return 42;
  return 42;
                                main:
                                   param 1;
int main () {
                                   param 2;
  f(1,2,3,4);
                                   param 3;
                                   param 4;
                                   call f, 4;
```

```
f:
  local t, 4;
  local u, 4;
```

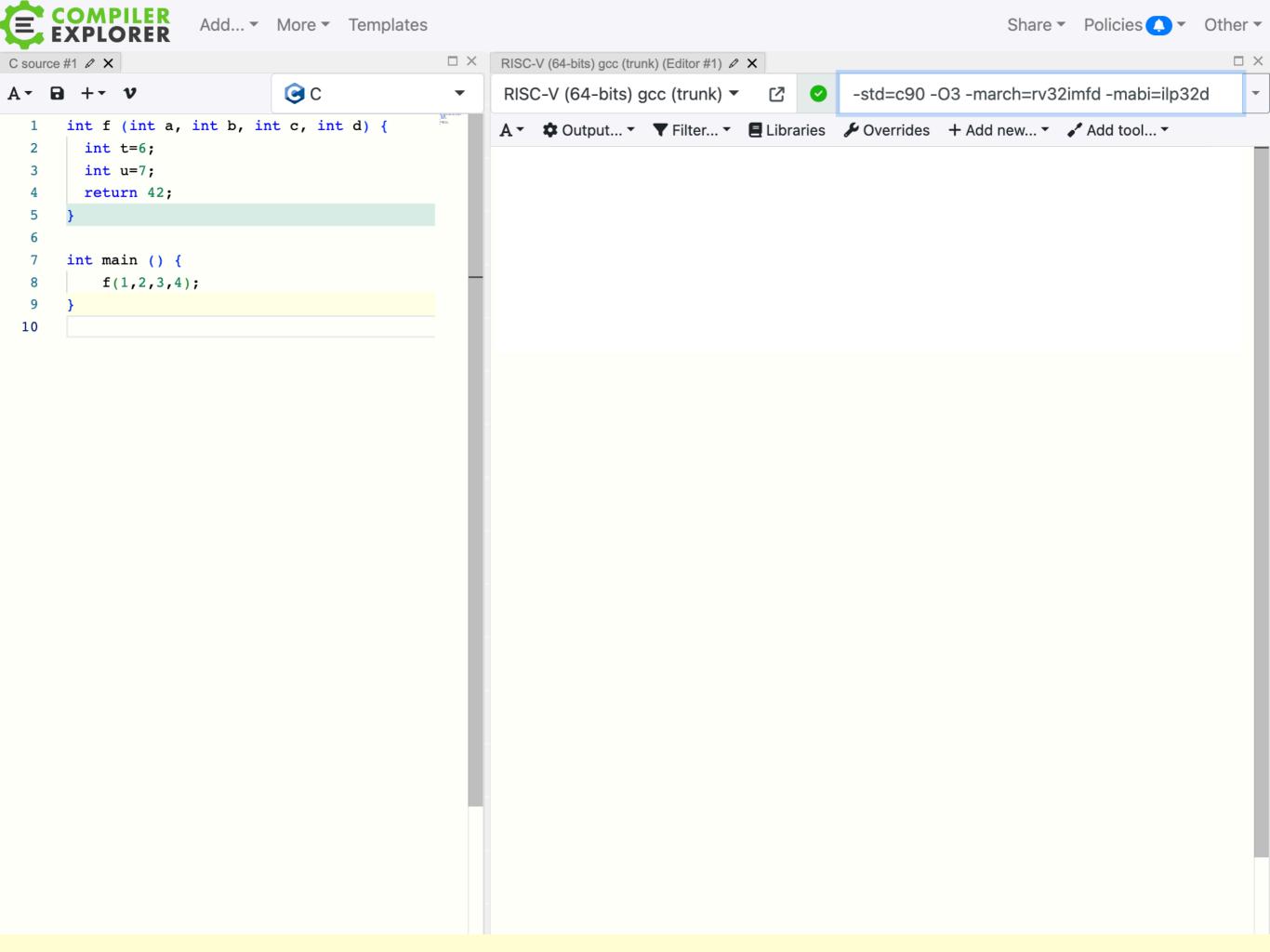
Function calls

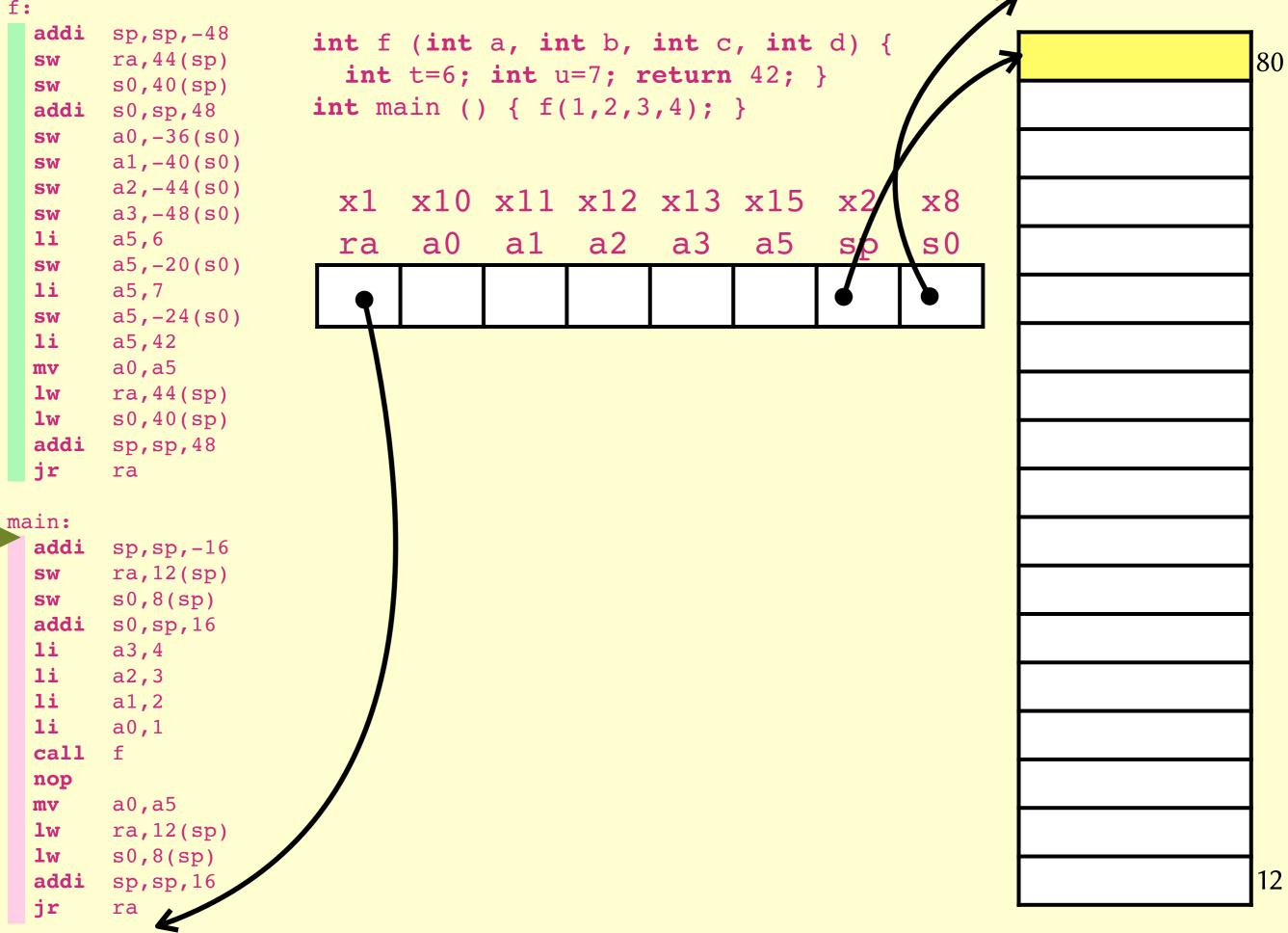
```
f:
  local t, 4;
  local u, 4;
  t = 6;
  u = 7;
  return 42;
main:
  param 1;
  param 2;
  param 3;
  param 4;
  call f, 4;
```

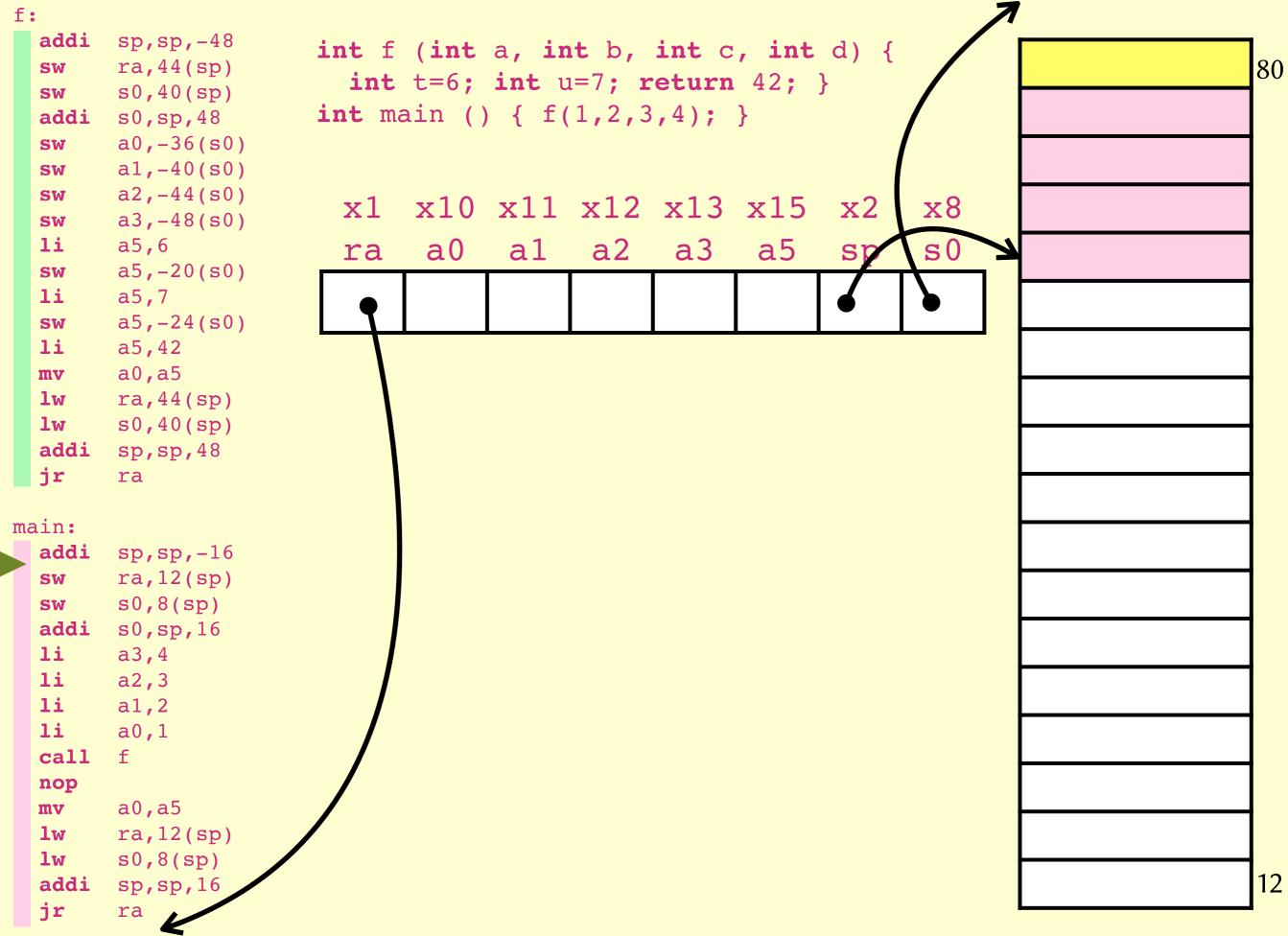
```
1
                   stack
           2
                   frame
                    for
           3
                  main
           4
          42
 ret val
ret addr
                   stack
                   frame
     t
                   for f
     u
```

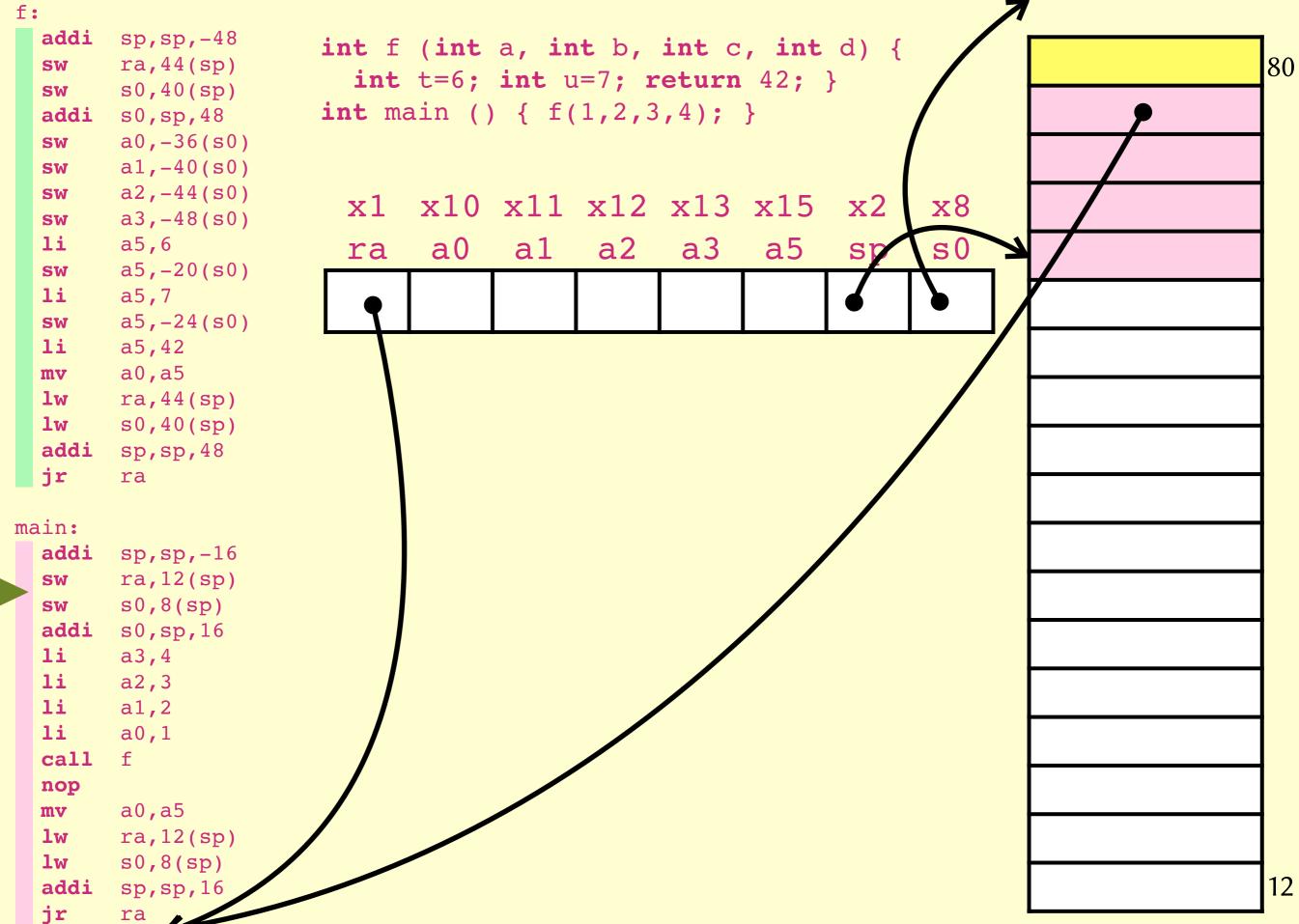
ABI

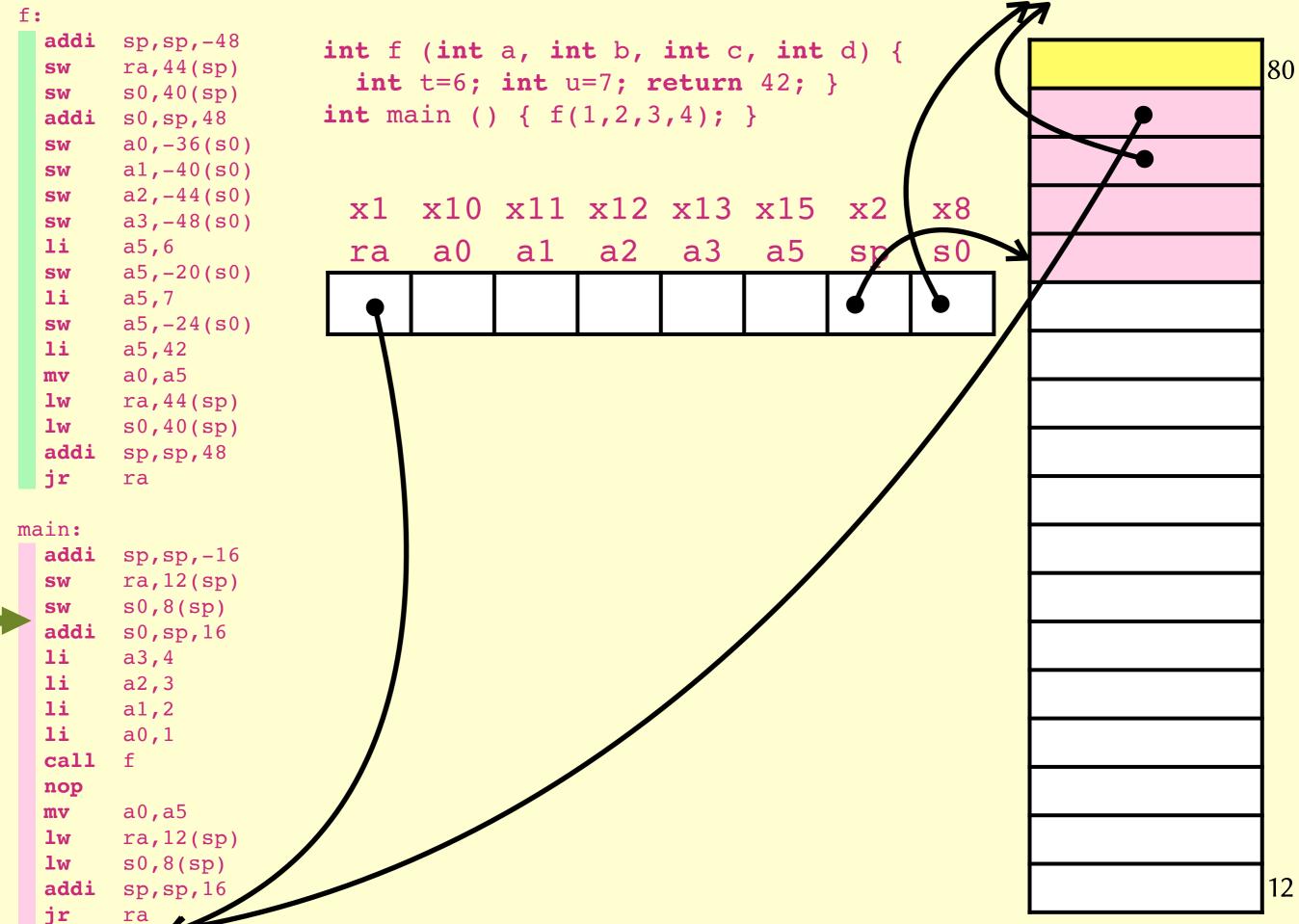
- An Application Binary Interface defines:
 - how parameters are passed to functions,
 - which registers need to be preserved by a function call,
 - how the fields of a **struct** are laid out,
 - and so on.
- Here's an illustration of the ABI used by GCC when targeting RISC-V...

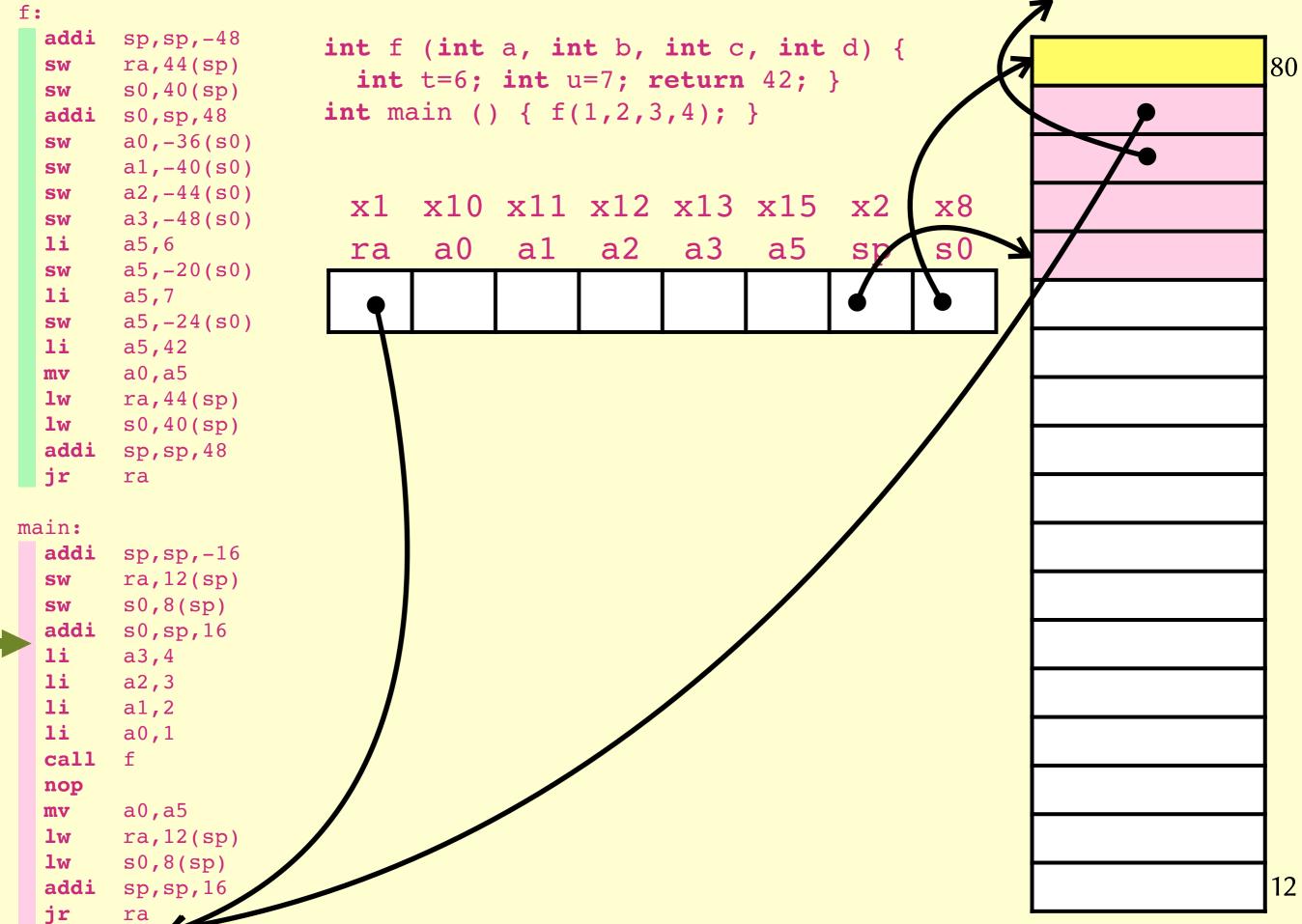


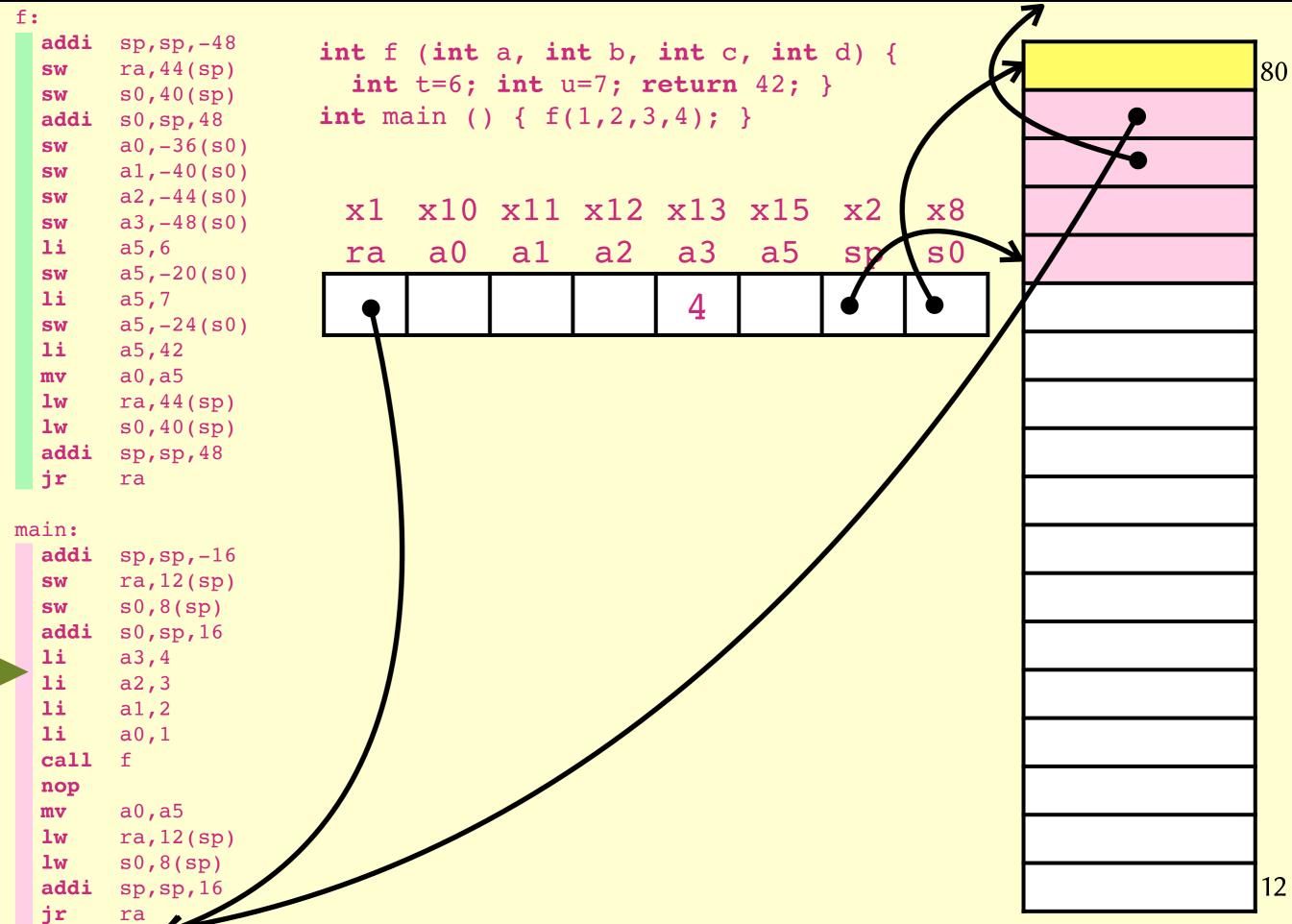


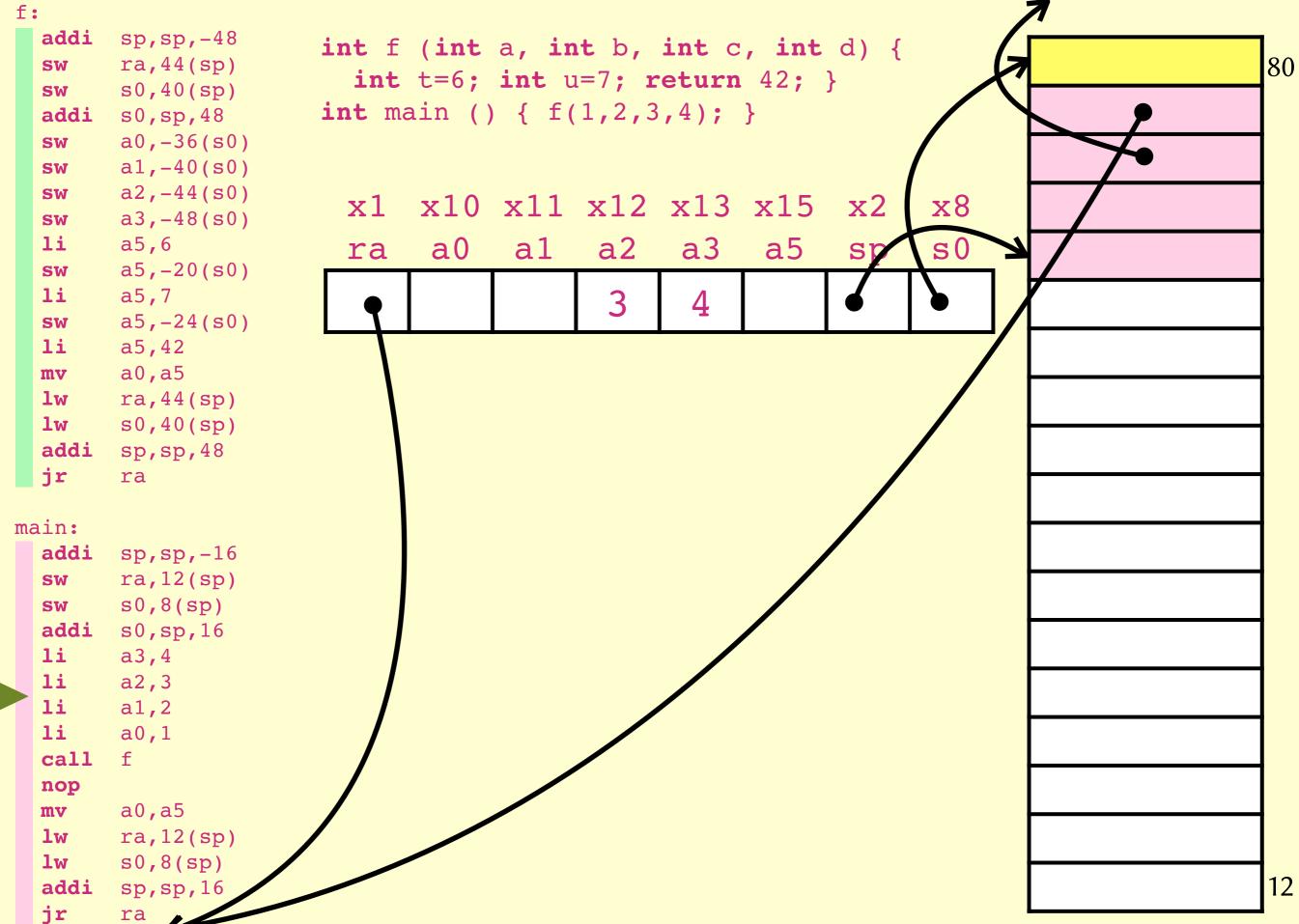


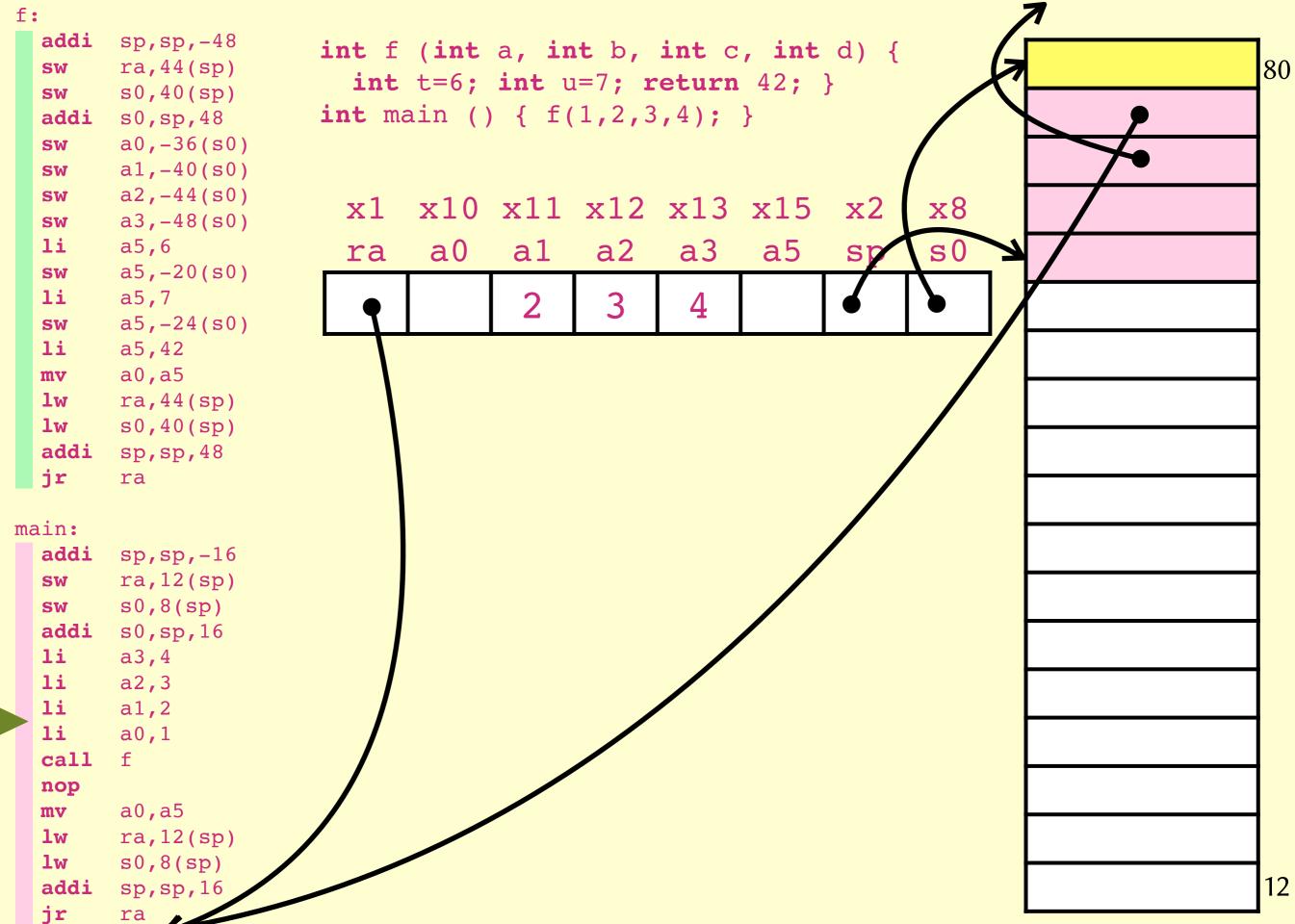


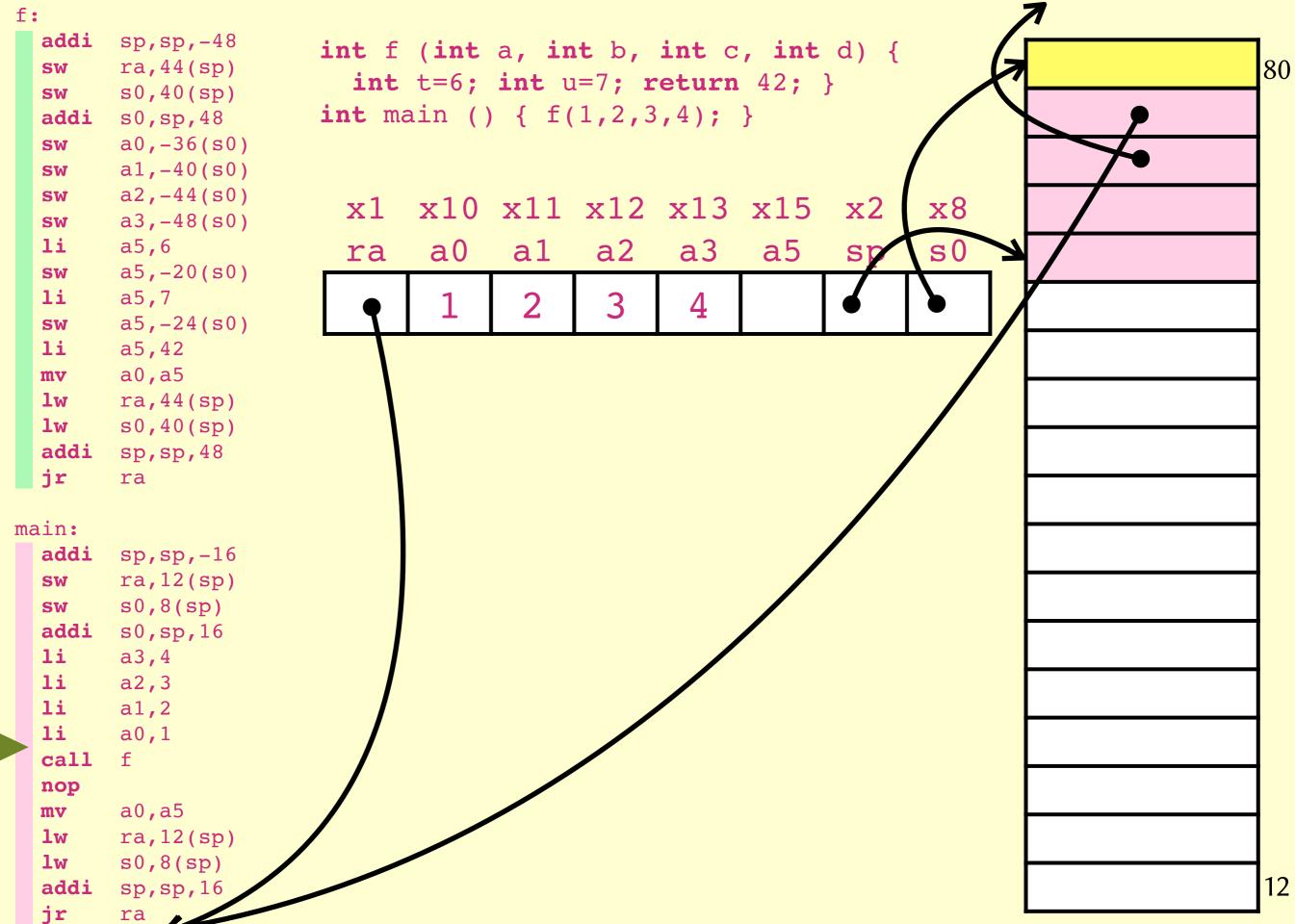


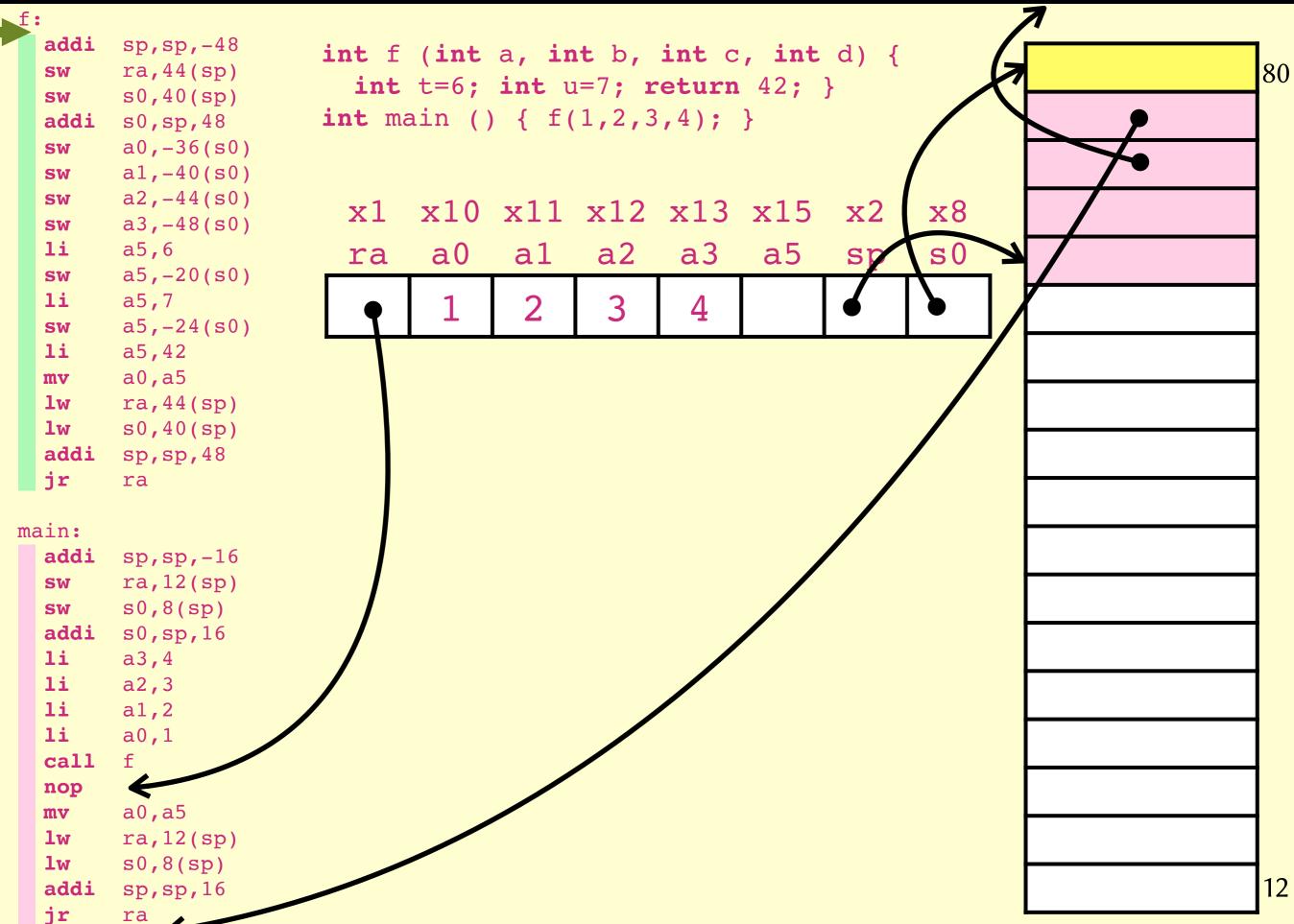


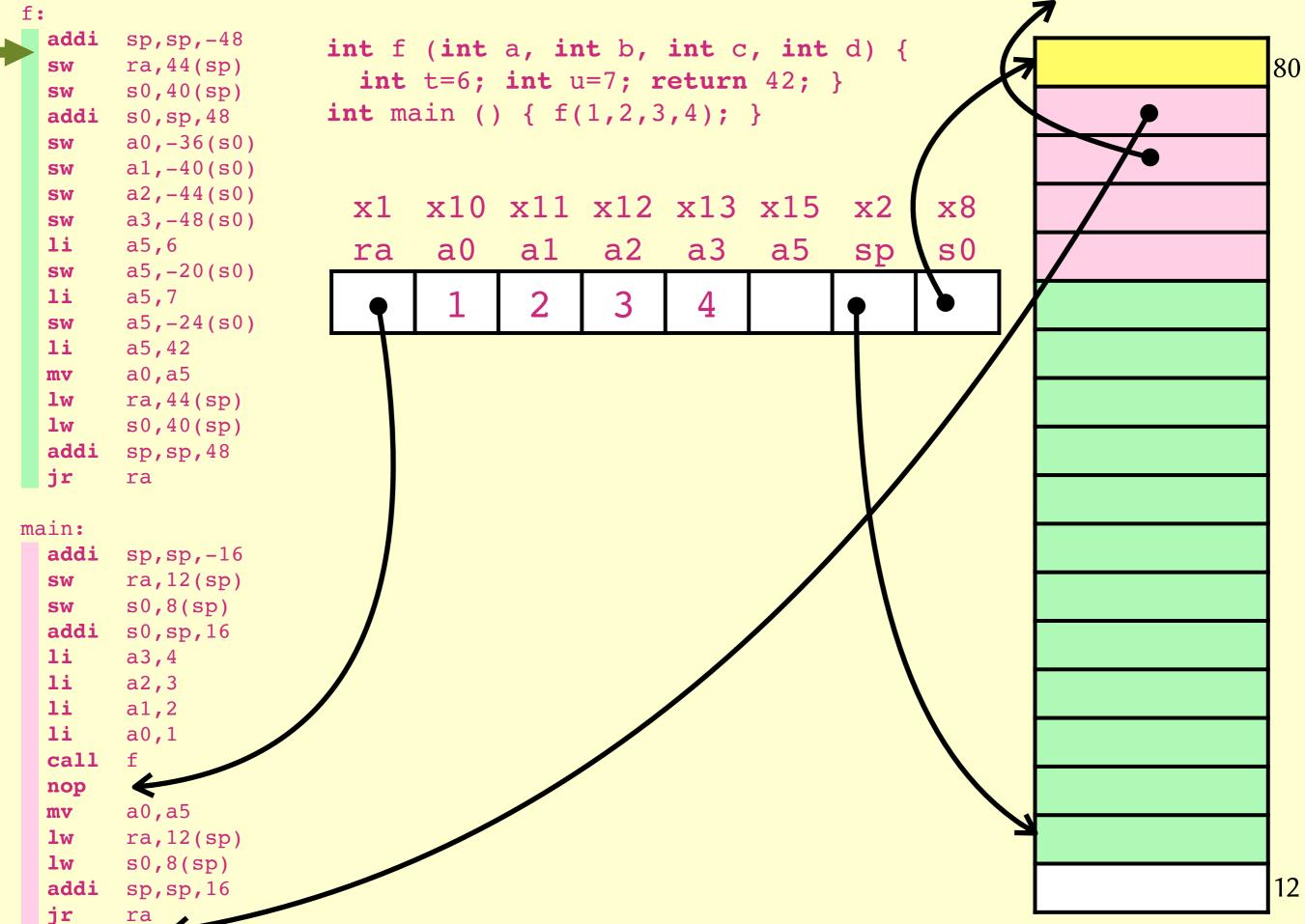


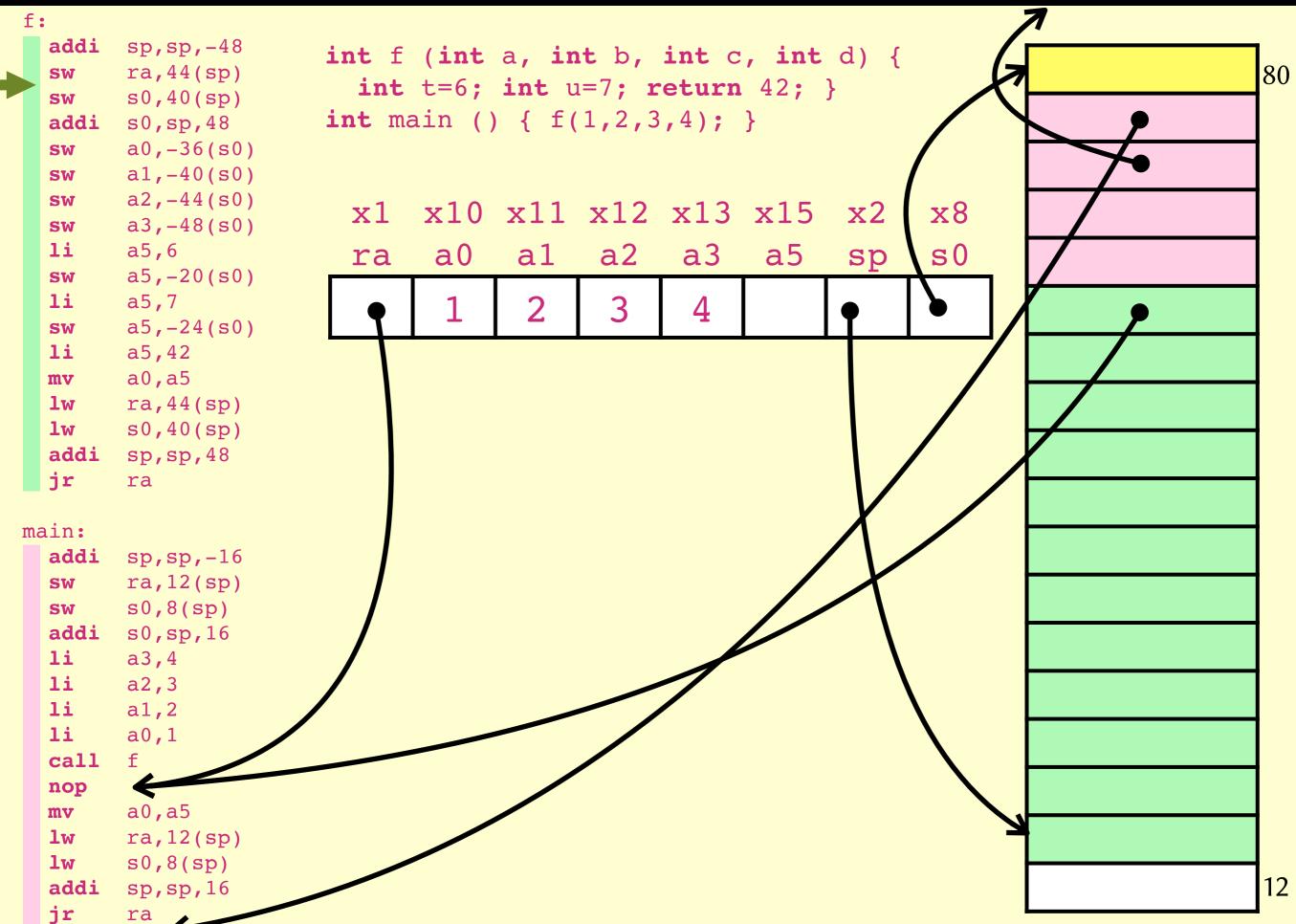


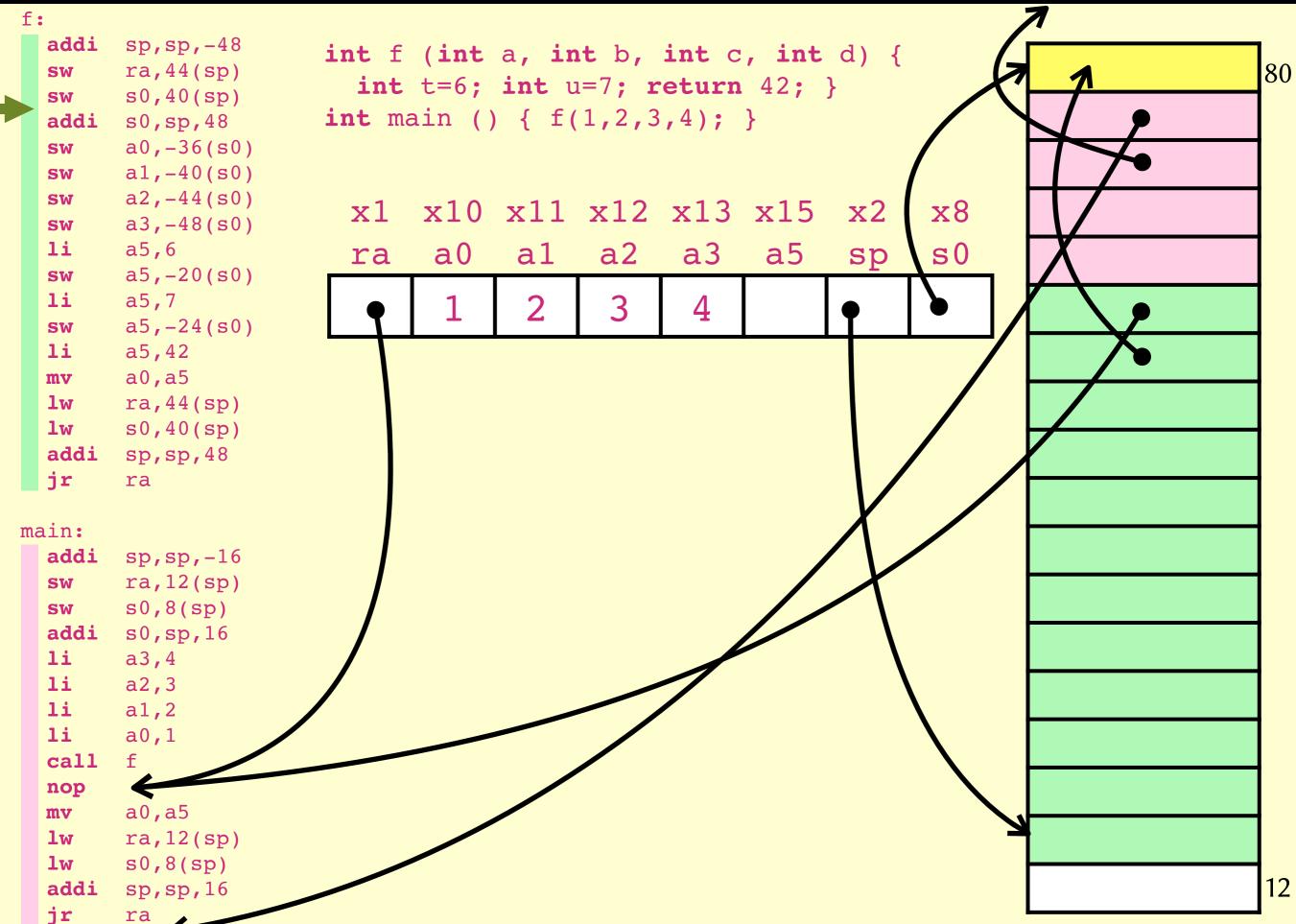


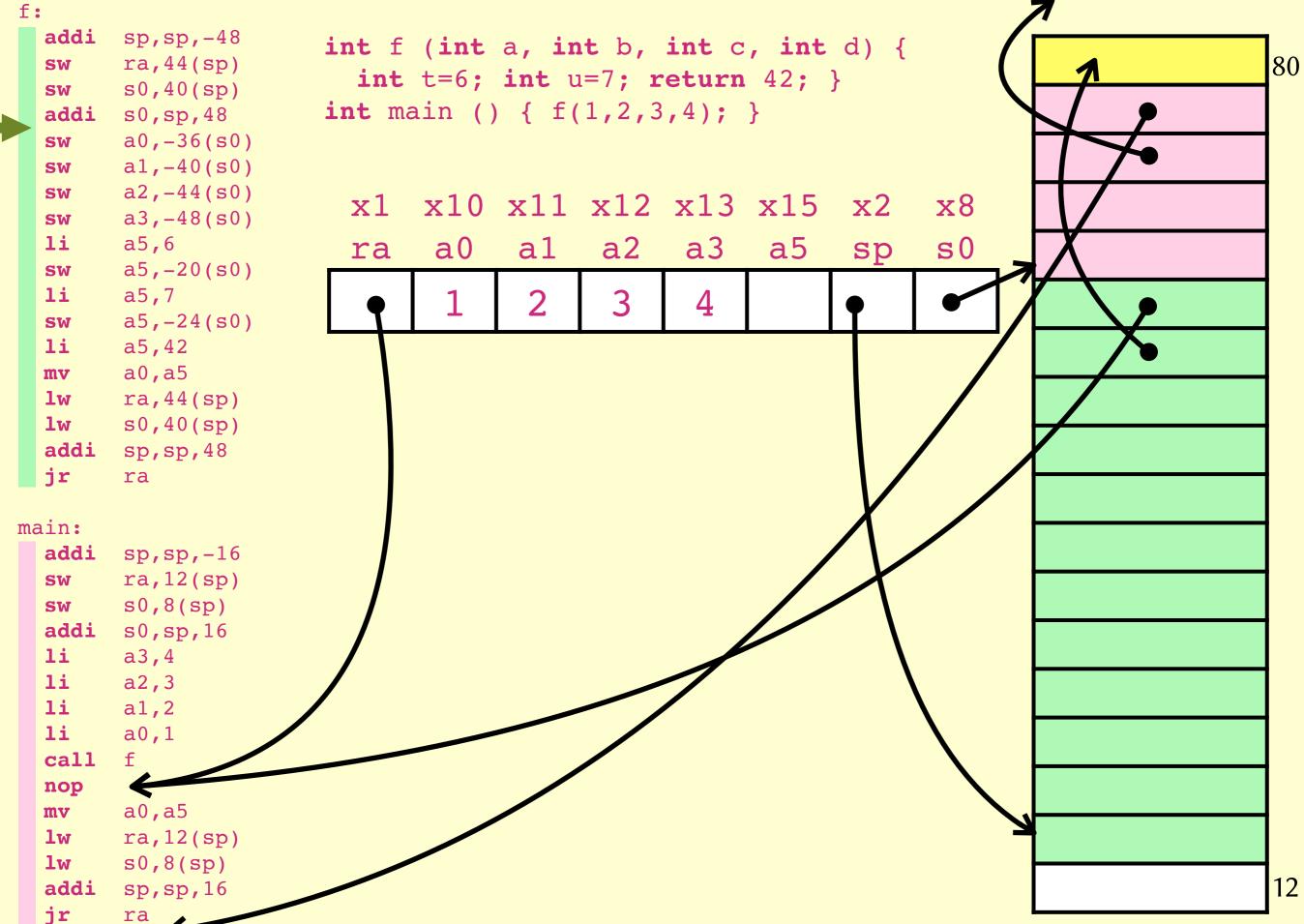


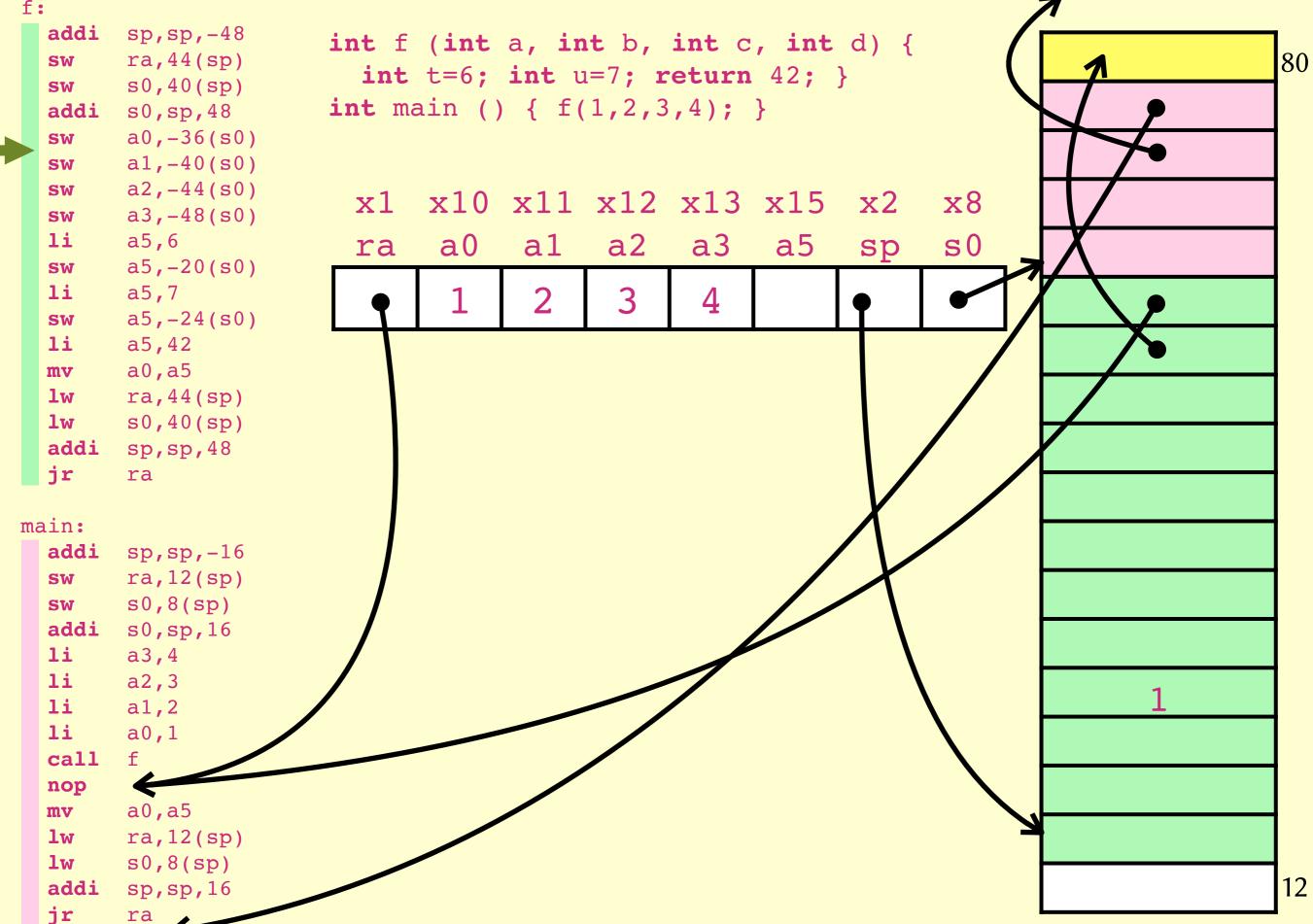


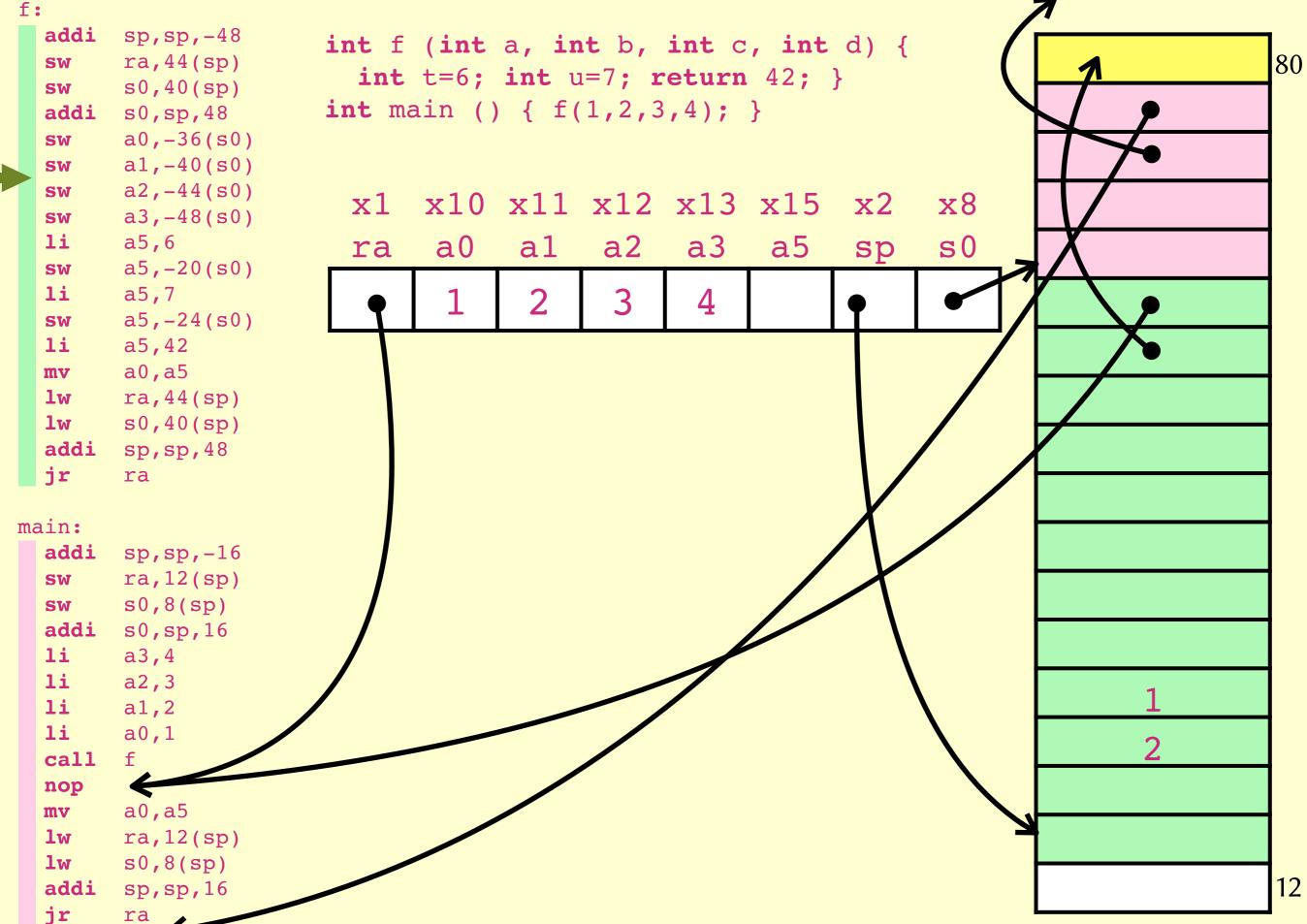


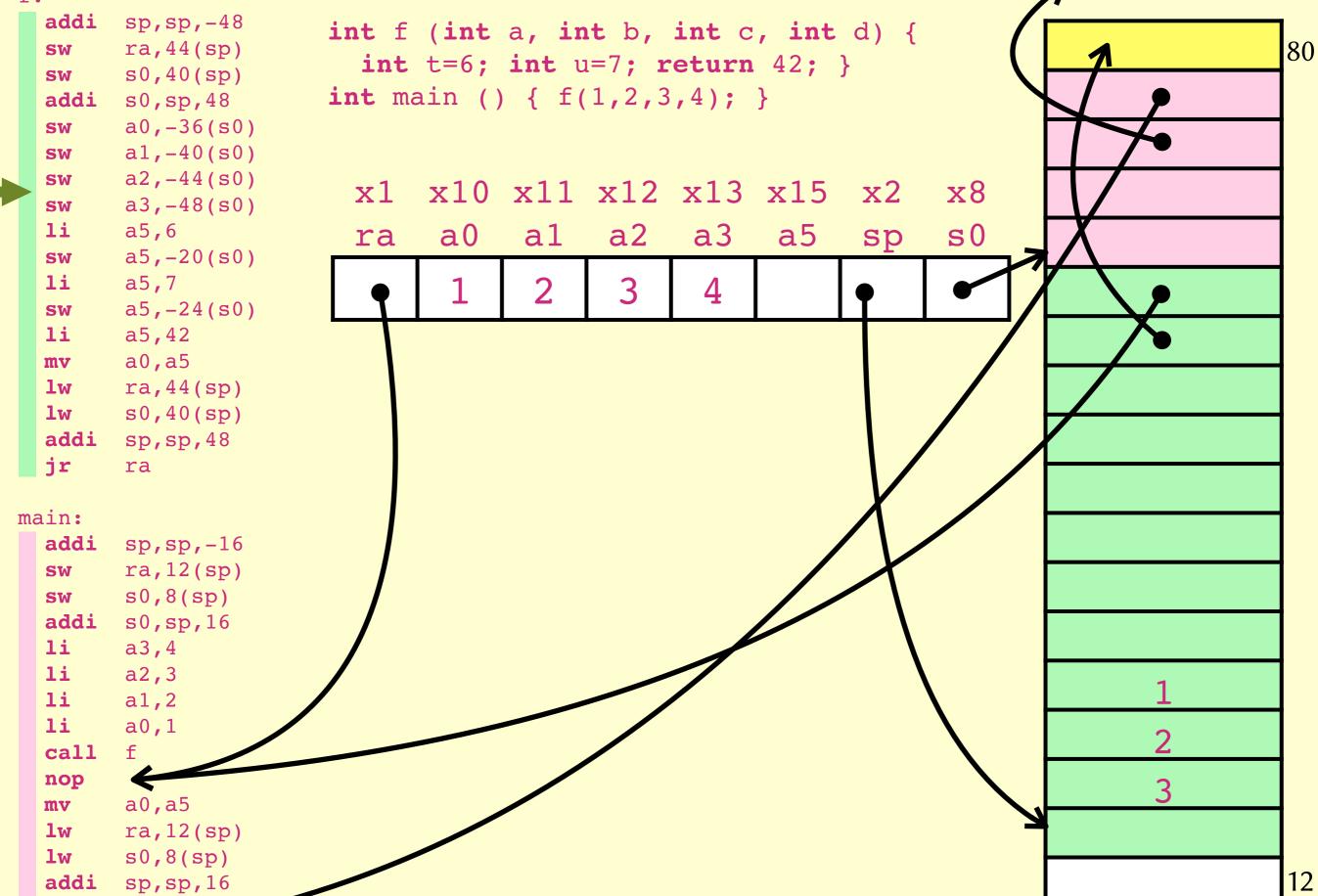


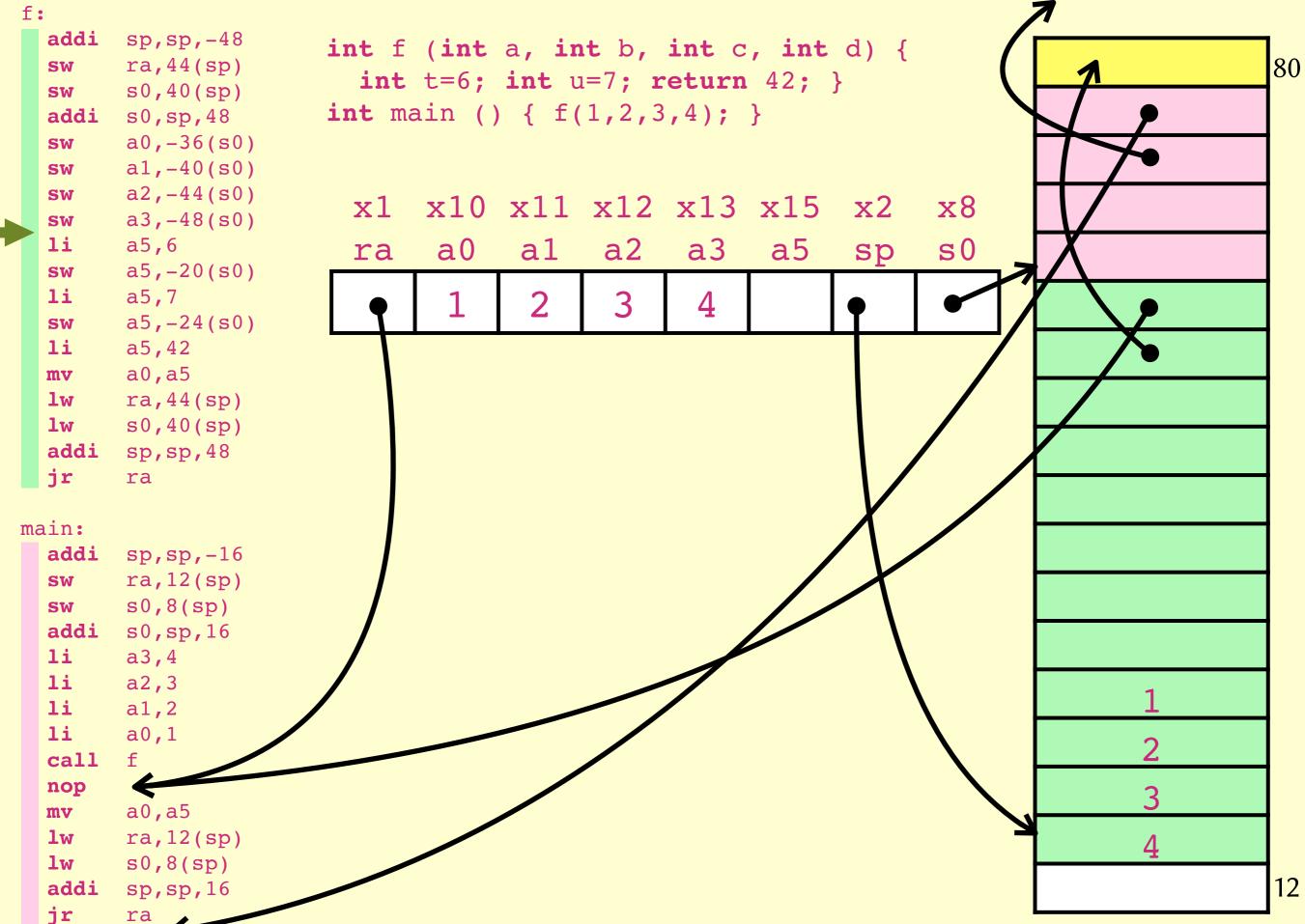


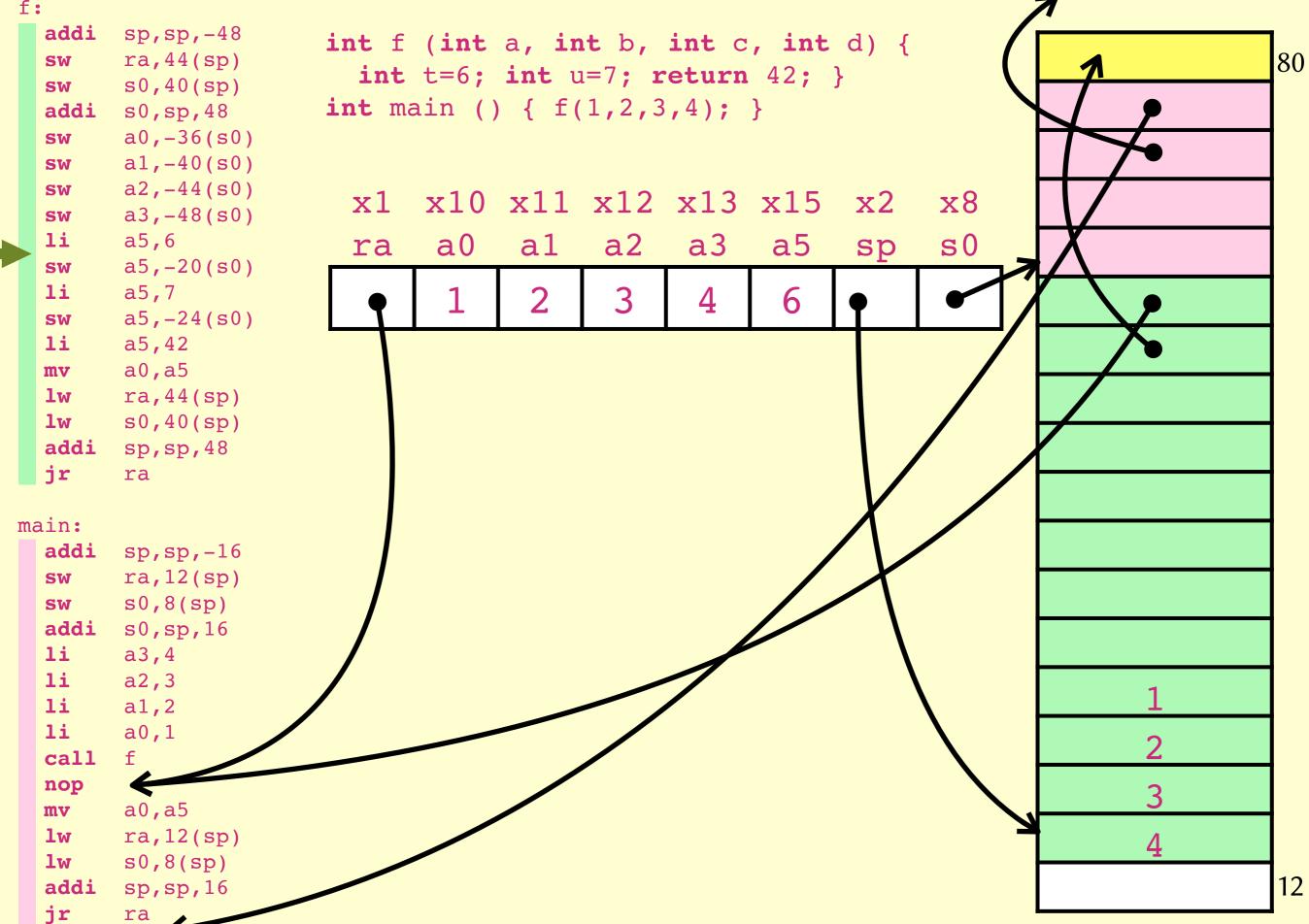


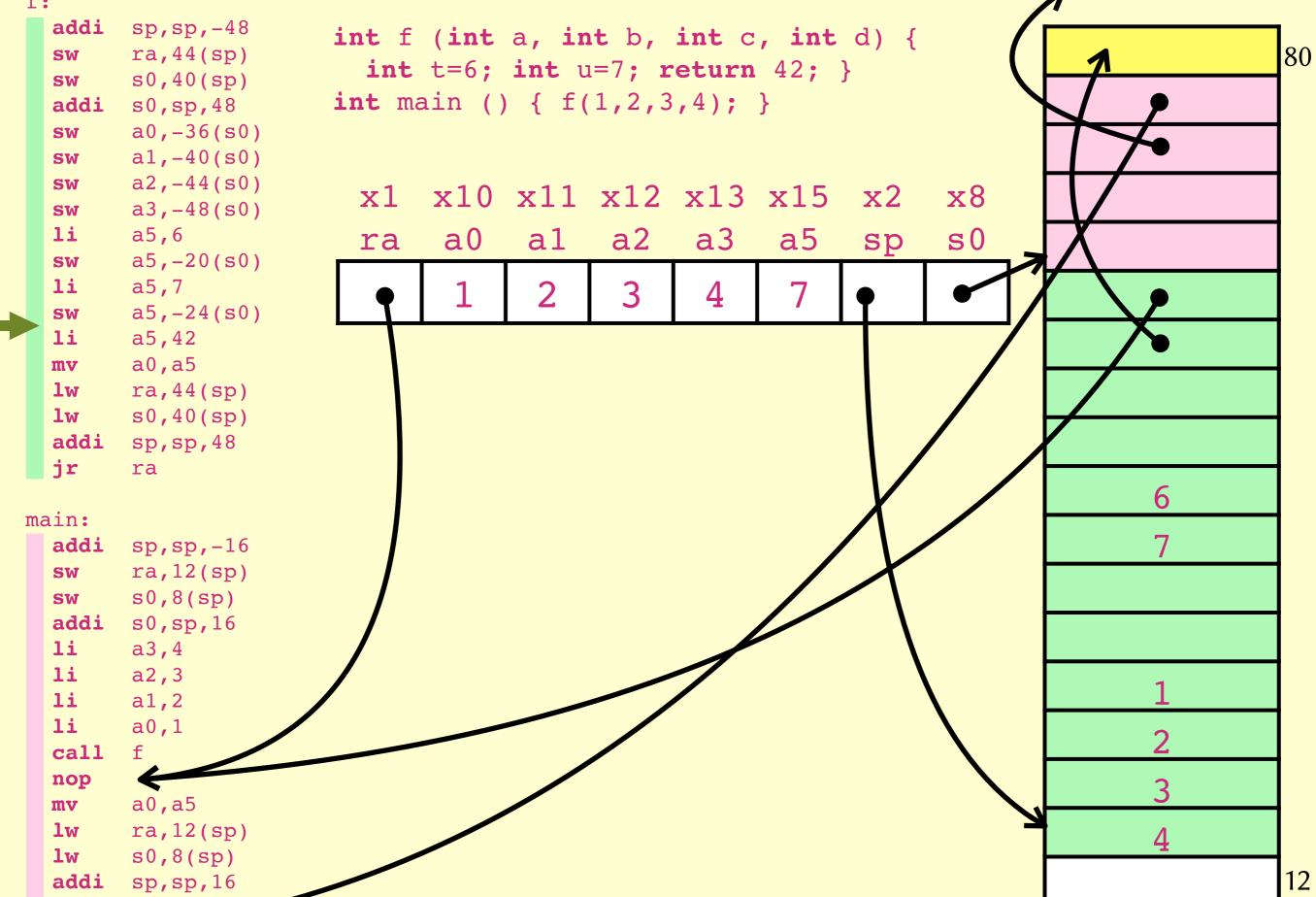


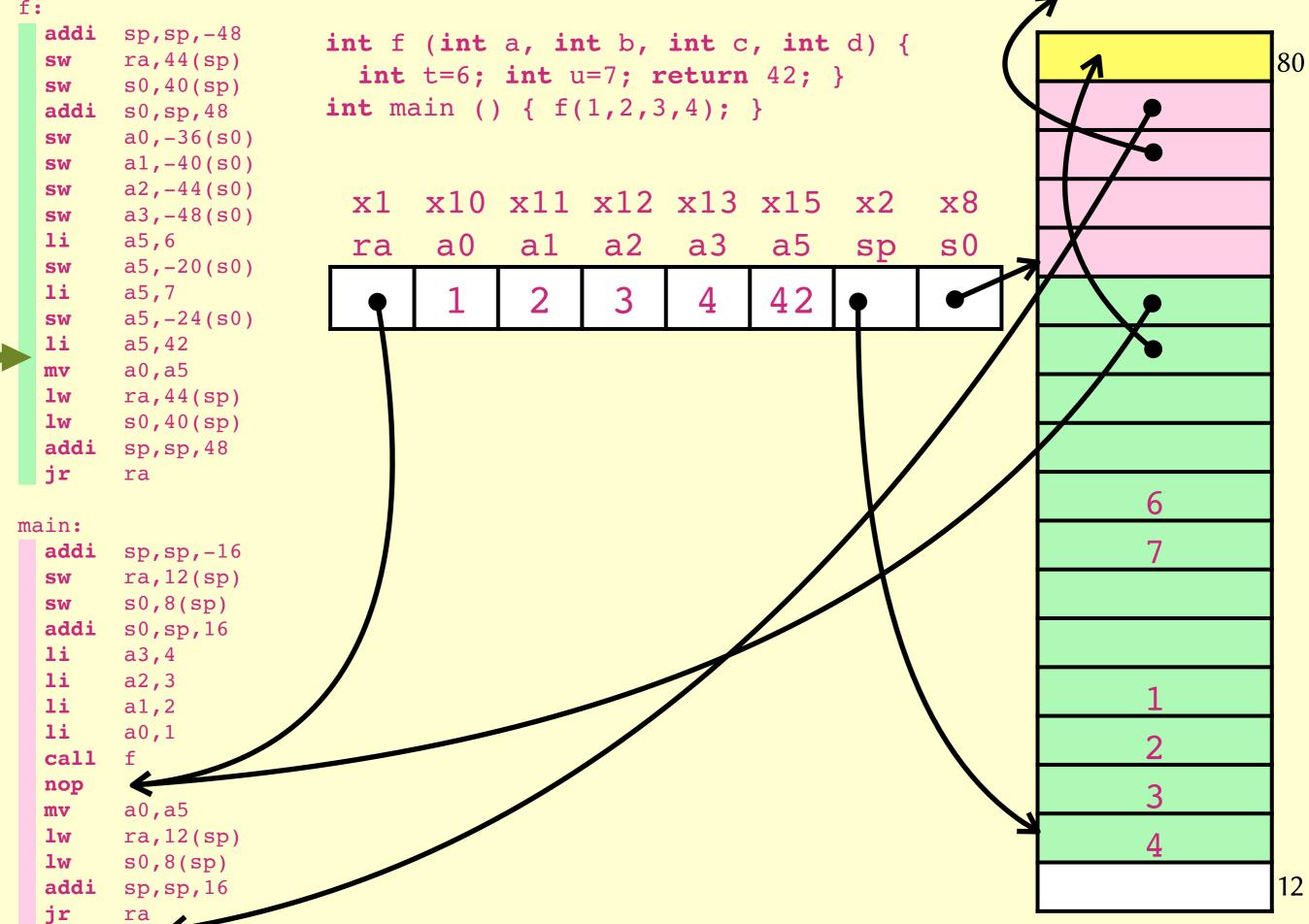












1i a0,1 call nop a0, a5 mv lw ra, 12 (sp)

s0,8(sp)

sp, sp, 16

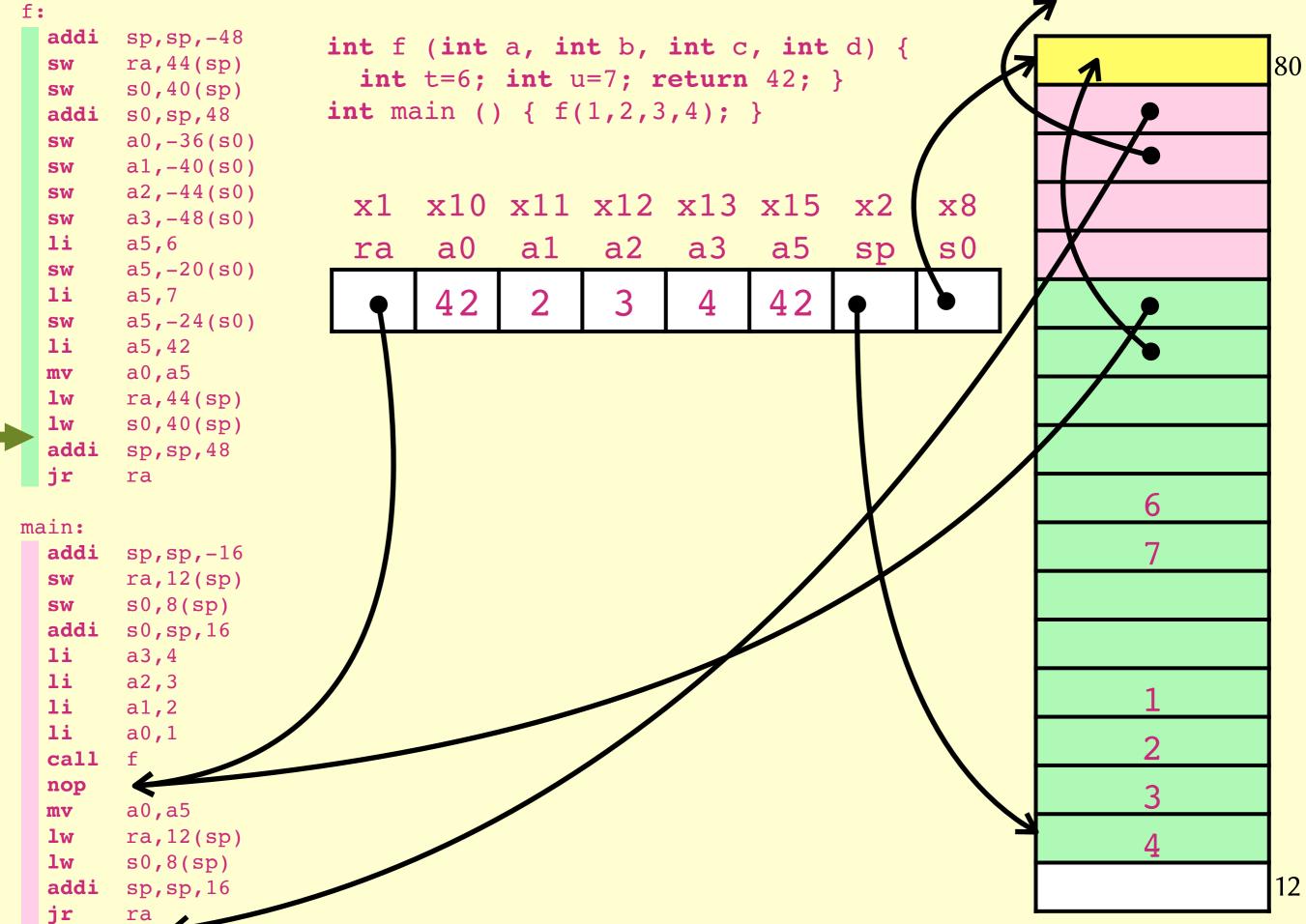
ra

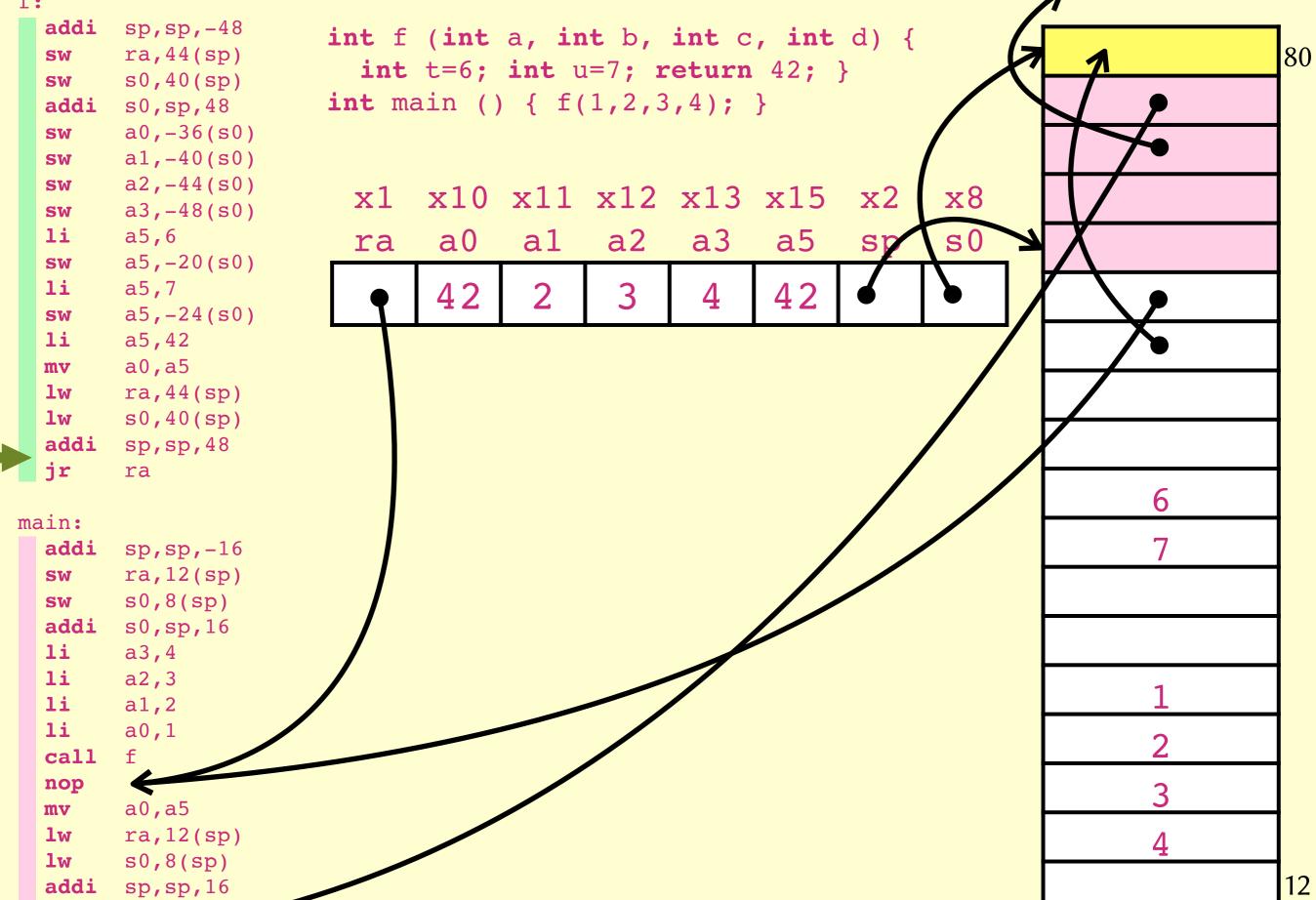
lw

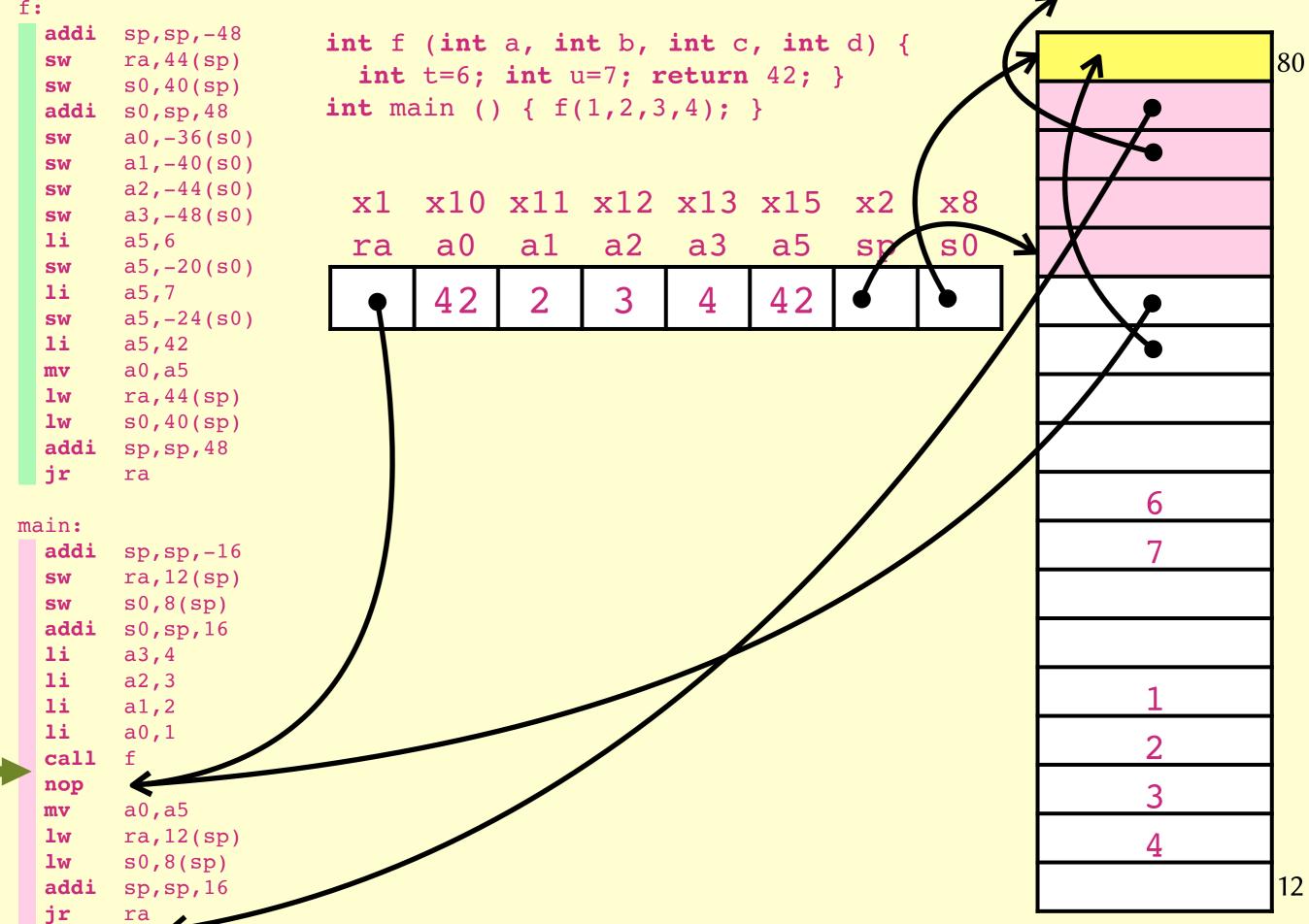
jr

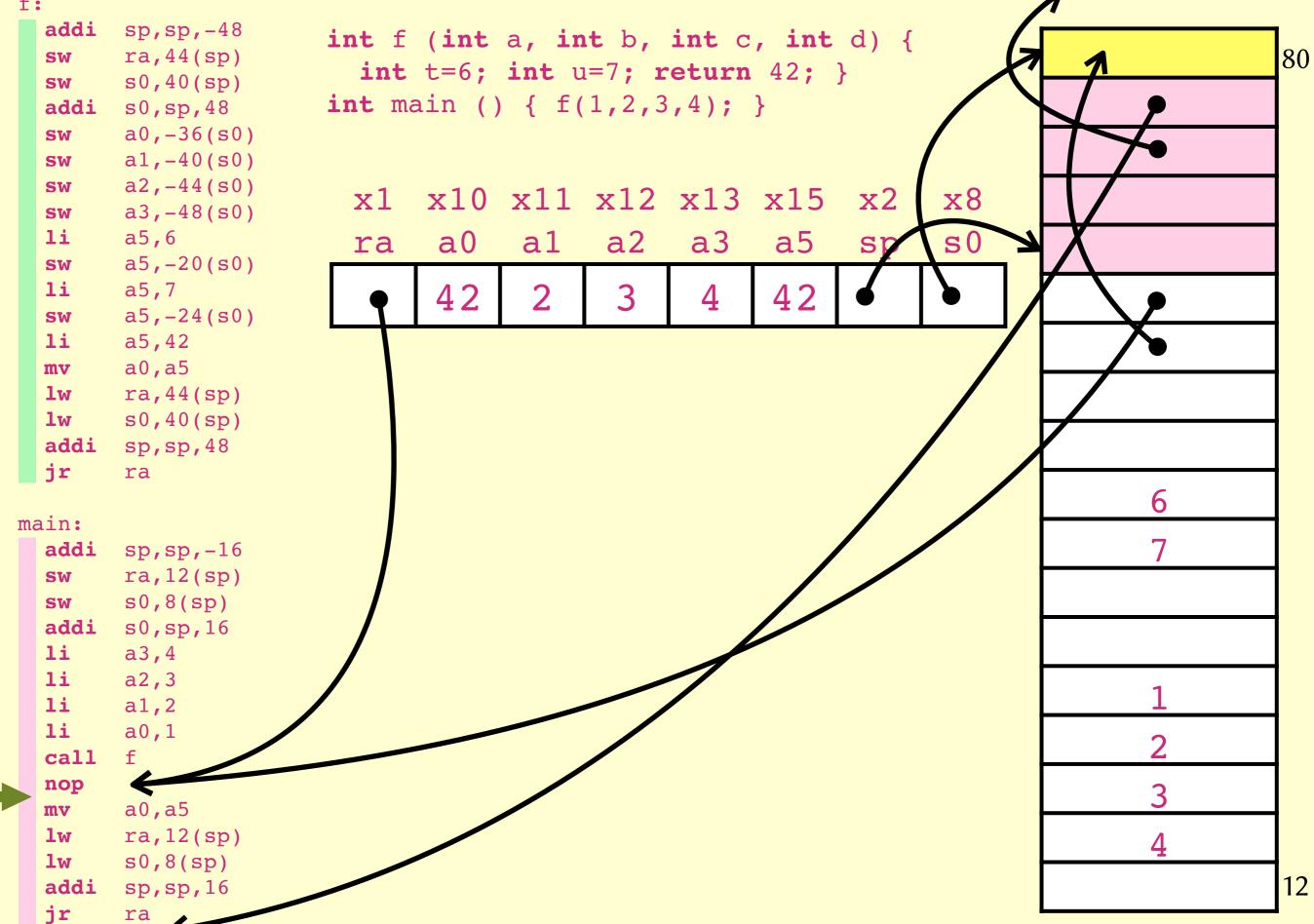
addi

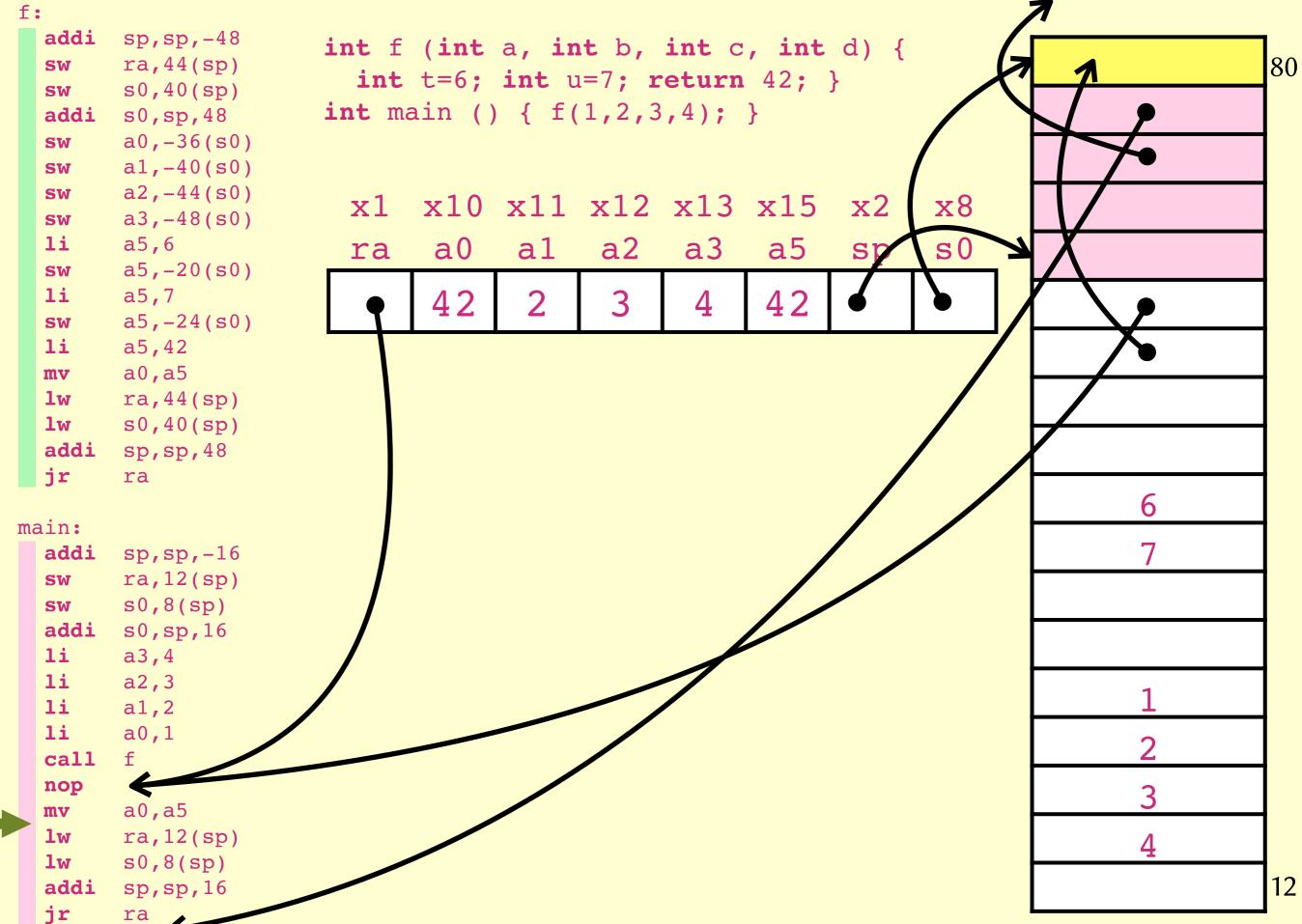
12

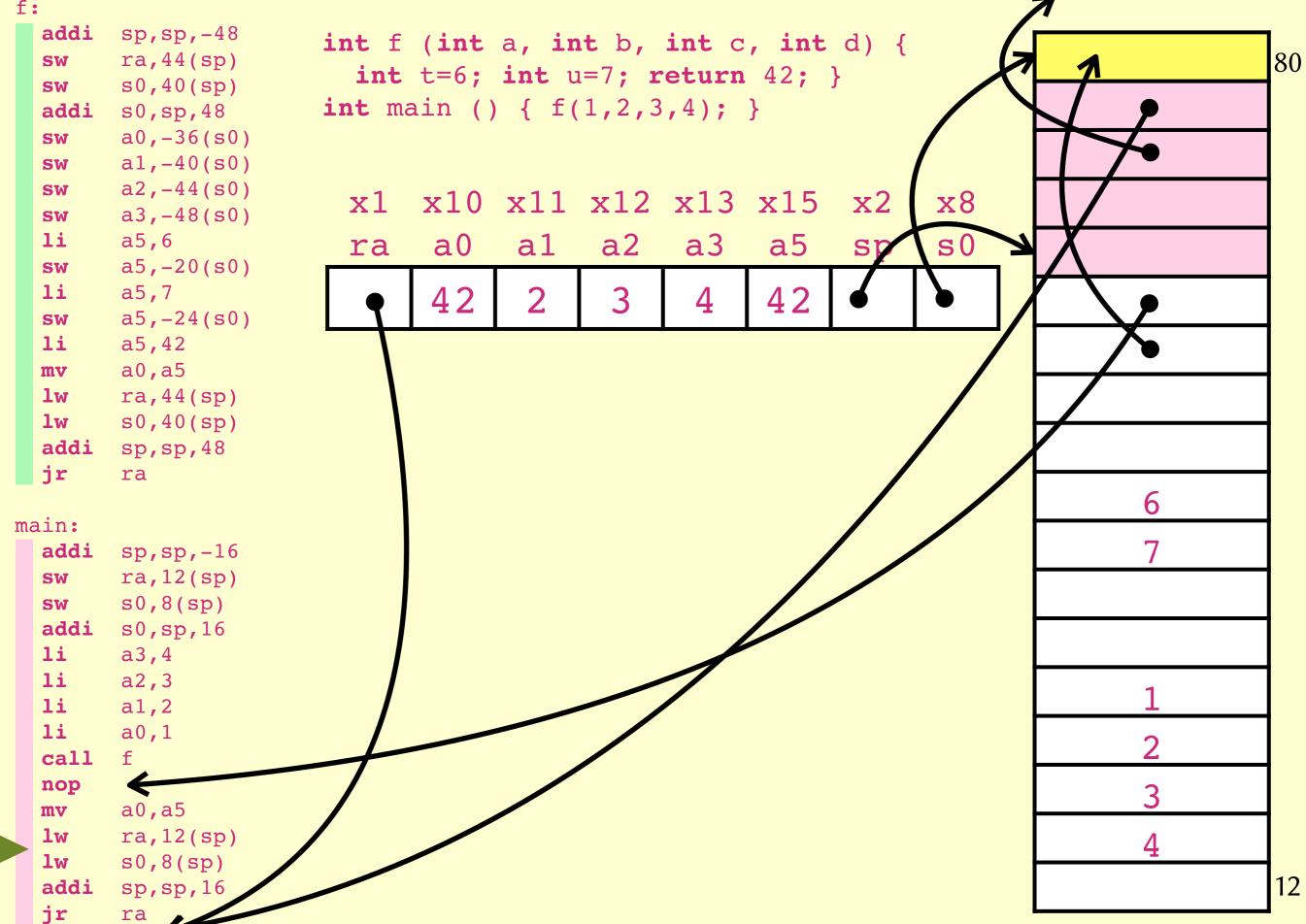


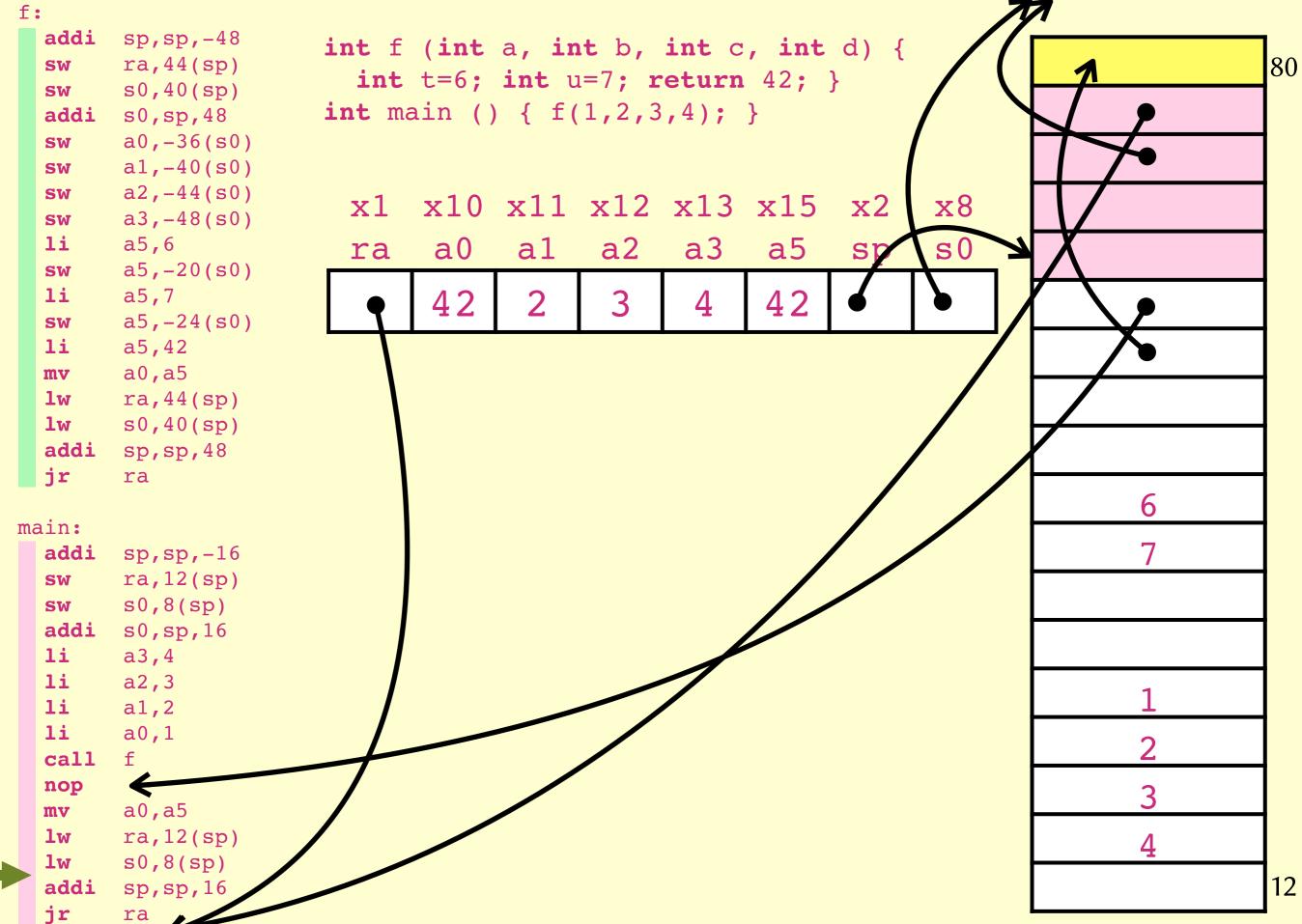


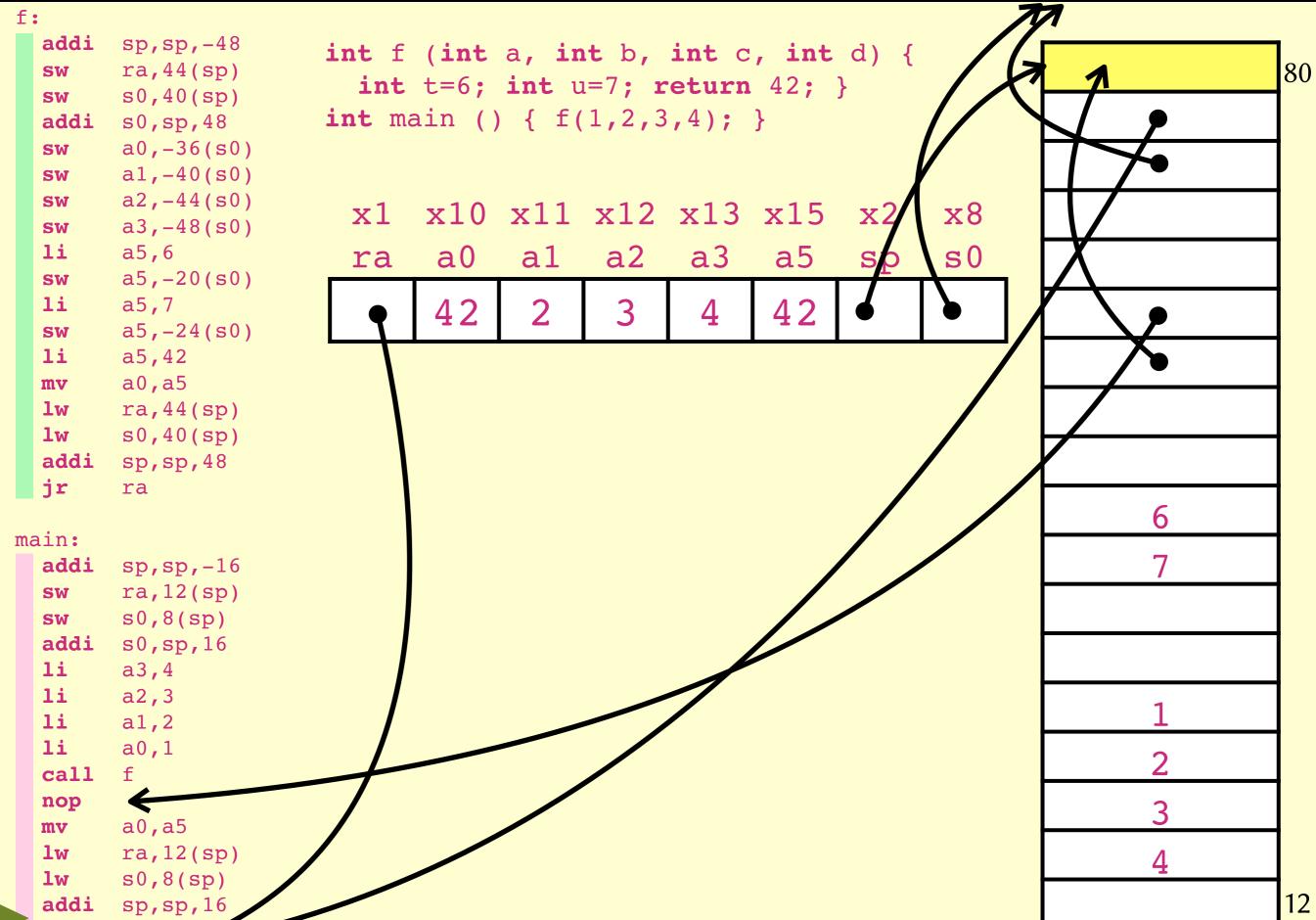


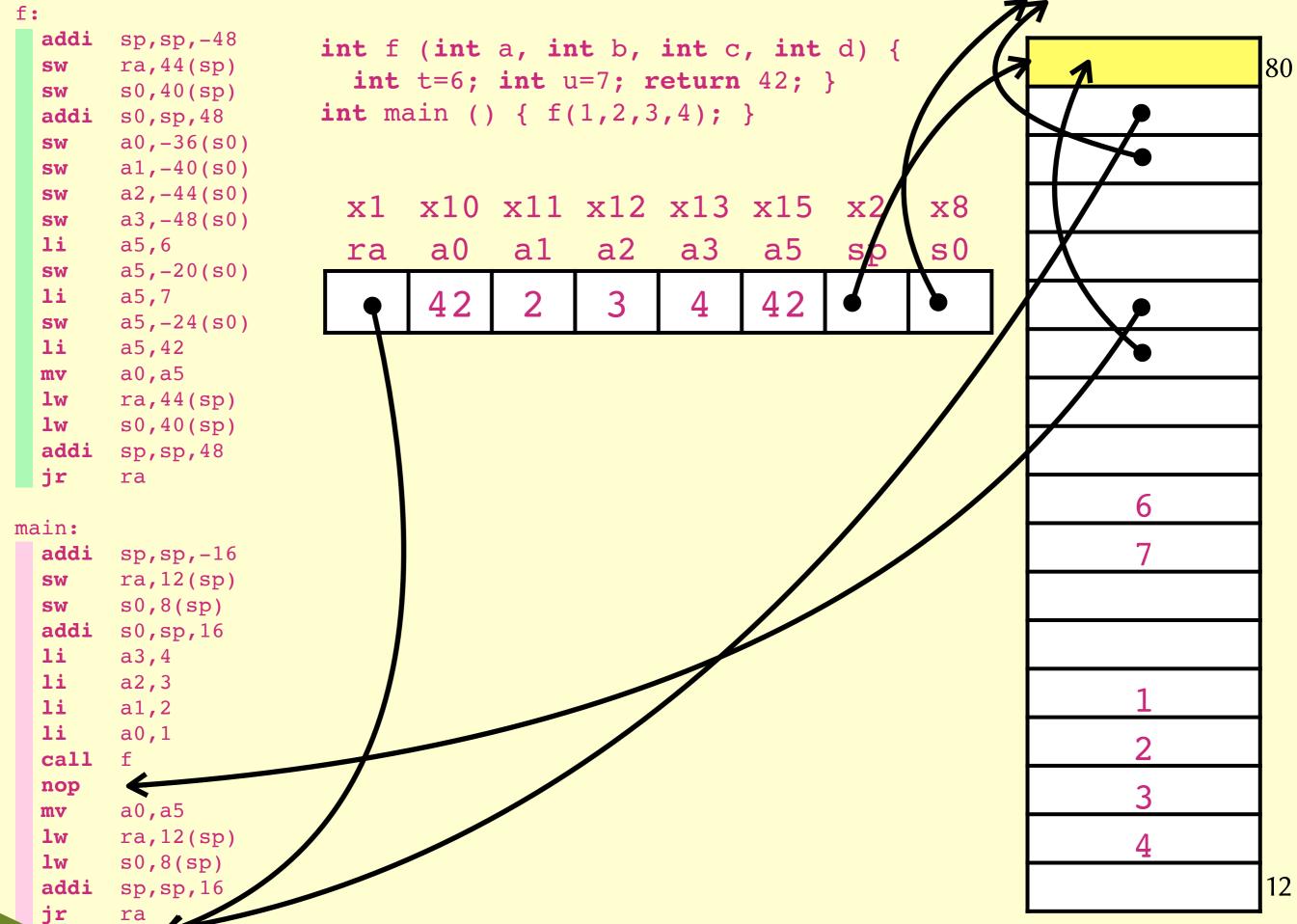












Lecture outline

- Parsing C: the *dangling-else* problem.
- Parsing C: the *lexer hack*.
- Generating RISC-V assembly for function calls.