

Script generated by TTT

Title: Seidl: Virtual_machines (15.05.2012)

Date: Tue May 15 14:01:15 CEST 2012

Duration: 91:09 min

Pages: 45

Example: `int a[10], *b;` with $\rho = \{a \mapsto 7, b \mapsto 17\}$.

For the statement: `*a = 5;` we obtain:

```

codeL (*a) ρ      = codeR a ρ = codeL a ρ = loadc 7
code (*a = 5;) ρ  = loadc 5
                  loadc 7
                  store
                  pop
    
```

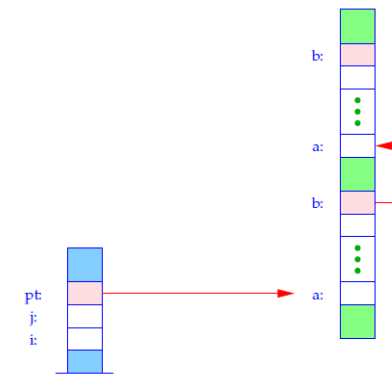
As an exercise translate:

$s_1 \equiv b = (&a) + 2;$ and $s_2 \equiv *(b + 3) = 5;$

$code_R q \rho = loadc q$ q constant

$code_R (e_1 = e_2) \rho = code_R e_2 \rho$
 $code_L e_1 \rho$
 store

$code_R e \rho = code_L e \rho$
 load otherwise



Be $\rho = \{i \mapsto 1, j \mapsto 2, pt \mapsto 3, a \mapsto 0, b \mapsto 7\}$. Then:

$code_L((pt \rightarrow b) \rightarrow a)[i+1] \rho$
 $= code_R((pt \rightarrow b) \rightarrow a) \rho = code_R((pt \rightarrow b) \rightarrow a) \rho$
 $code_R(i+1) \rho \quad loada \ 1$
 $loadc \ 1 \quad loadc \ 1$
 $mul \quad add$
 $add \quad loadc \ 1$
 mul
 add

Dereferencing of Pointers:

The application of the operator $*$ to the expression e returns the contents of the storage cell, whose address is the R-value of e :

$$code_L(*e) \rho = code_R e \rho$$

Example: Given the declarations

```
struct t { int a[7]; struct t *b; };
int i, j;
struct t *pt;
```

and the expression $((pt \rightarrow b) \rightarrow a)[i+1]$

Because of $e \rightarrow a \equiv (*e).a$ holds:

$$code_L(e \rightarrow a) \rho = code_R e \rho$$

$$loadc(\rho a)$$

$$add$$

Be $\rho = \{i \mapsto 1, j \mapsto 2, pt \mapsto 3, a \mapsto 0, b \mapsto 7\}$. Then:

$code_L((pt \rightarrow b) \rightarrow a)[i+1] \rho$
 $= code_R((pt \rightarrow b) \rightarrow a) \rho = code_R((pt \rightarrow b) \rightarrow a) \rho$
 $code_R(i+1) \rho \quad loada \ 1$
 $loadc \ 1 \quad loadc \ 1$
 $mul \quad add$
 $add \quad loadc \ 1$
 mul
 add

For arrays, their R-value equals their L-value. Therefore:

$$code_R((pt \rightarrow b) \rightarrow a) \rho = code_R(pt \rightarrow b) \rho =$$

loada 3
loadc 7
add
load
loadc 0
add

In total, we obtain the instruction sequence:

loada 3	load	loada 1	loadc 1
loadc 7	loadc 0	loadc 1	mul
add	add	add	add

Be $\rho = \{i \mapsto 1, j \mapsto 2, pt \mapsto 3, a \mapsto 0, b \mapsto 7\}$. Then:

$$\begin{aligned}
 \text{code}_L((pt \rightarrow b) \rightarrow a)[i+1] \rho & \\
 = \text{code}_R((pt \rightarrow b) \rightarrow a) \rho & = \text{code}_R((pt \rightarrow b) \rightarrow a) \rho \\
 \text{code}_R(i+1) \rho & \quad \text{loada 1} \\
 \text{loadc 1} & \quad \text{loadc 1} \\
 \text{mul} & \quad \text{add} \\
 \text{add} & \quad \text{loadc 1} \\
 & \quad \text{mul} \\
 & \quad \text{add}
 \end{aligned}$$

For arrays, their R-value equals their L-value. Therefore:

$$\begin{aligned}
 \text{code}_R((pt \rightarrow b) \rightarrow a) \rho & = \text{code}_R(pt \rightarrow b) \rho & = \text{loada 3} \\
 & \quad \text{loadc 0} & \quad \text{loadc 7} \\
 & \quad \text{add} & \quad \text{add} \\
 & & \quad \text{load} \\
 & & \quad \text{loadc 0} \\
 & & \quad \text{add}
 \end{aligned}$$

In total, we obtain the instruction sequence:

loada 3	load	loada 1	loadc 1
loadc 7	loadc 0	loadc 1	mul
add	add	add	add

7 Conclusion

We tabulate the cases of the translation of expressions:

$$\begin{aligned}
 \text{code}_L(e_1[e_2]) \rho & = \text{code}_R e_1 \rho \\
 & \quad \text{code}_R e_2 \rho \\
 & \quad \text{loadc } |t| \\
 & \quad \text{mul} \\
 & \quad \text{add} \quad \text{if } e_1 \text{ has type } t^* \text{ or } t[] \\
 \\
 \text{code}_L(e.a) \rho & = \text{code}_L e \rho \\
 & \quad \text{loadc } (\rho a) \\
 & \quad \text{add}
 \end{aligned}$$

$$\begin{aligned}
 \text{code}_L(*e) \rho & = \text{code}_R e \rho \\
 \\
 \text{code}_L x \rho & = \text{loadc } (\rho x) \\
 \\
 \text{code}_R(\&e) \rho & = \text{code}_L e \rho \\
 \\
 \text{code}_R e \rho & = \text{code}_L e \rho \quad \text{if } e \text{ is an array} \\
 \\
 \text{code}_R(e_1 \square e_2) \rho & = \text{code}_R e_1 \rho \\
 & \quad \text{code}_R e_2 \rho \\
 & \quad \text{op} \quad \text{op instruction for operator '}\square\text{' }
 \end{aligned}$$

$$\text{code}_L (*e) \rho = \text{code}_R e \rho$$

$$\text{code}_L x \rho = \text{loadc} (\rho x)$$

$$\text{code}_R (&e) \rho = \text{code}_L e \rho$$

$$\text{code}_R e \rho = \text{code}_L e \rho \quad \text{if } e \text{ is an array}$$

$$\text{code}_R (e_1 \square e_2) \rho = \text{code}_R e_1 \rho$$

$$\text{code}_R e_2 \rho$$

op instruction for operator ' \square '

62

$$\text{code}_R q \rho = \text{loadc } q \quad q \text{ constant}$$

$$\text{code}_R (e_1 = e_2) \rho = \text{code}_R e_2 \rho$$

$$\text{code}_L e_1 \rho$$

store

$$\text{code}_R e \rho = \text{code}_L e \rho$$

load otherwise

63

Example: `int a[10], *b;` with $\rho = \{a \mapsto 7, b \mapsto 17\}$.

For the statement: `*a = 5;` we obtain:

LOJ

$$\text{code}_L (*a) \rho = \text{code}_R a \rho = \text{code}_L a \rho = \text{loadc } 7$$

$$\text{code} (*a = 5;) \rho = \text{loadc } 5$$

loadc 7

store

pop

As an exercise translate:

$$s_1 \equiv b = (&a) + 2; \quad \text{and} \quad s_2 \equiv *(b + 3) = 5;$$

64

$$\text{code} (s_1 s_2) \rho = \text{loadc } 7 \quad \text{loadc } 5$$

$$\text{loadc } 2 \quad \text{loadc } 17$$

$$\text{loadc } 10 \quad // \text{ size of int}[10] \quad \text{load}$$

$$\text{mul} \quad // \text{ scaling} \quad \text{loadc } 3$$

$$\text{add} \quad \text{loadc } 1 \quad // \text{ size of int}$$

$$\text{loadc } 17 \quad \text{mul} \quad // \text{ scaling}$$

$$\text{store} \quad \text{add}$$

$$\text{pop} \quad // \text{ end of } s_1 \quad \text{store}$$

$$\text{pop} \quad // \text{ end of } s_2$$

65

```

code (s1s2) ρ =   loadc 7           loadc 5
                  loadc 2           loadc 17
                  loadc 10 // size of int[10] load
                  mul // scaling      loadc 3
                  add                 loadc 1 // size of int
                  loadc 17            mul // scaling
                  store               add
                  pop // end of s1     store
                                      pop // end of s2

```

65

```

code (s1s2) ρ =   loadc 7           loadc 5
                  loadc 2           loadc 17
                  loadc 10 // size of int[10] load
                  mul // scaling      loadc 3
                  add                 loadc 1 // size of int
                  loadc 17            mul // scaling
                  store               add
                  pop // end of s1     store
                                      pop // end of s2

```

65

*int (*b) [10];*

Example: int a[10], *b; with $\rho = \{a \mapsto 7, b \mapsto 17\}$.

For the statement: `*a = 5;` we obtain:

```

codeL (*a) ρ = codeR a ρ = codeL a ρ = loadc 7
code (*a = 5;) ρ = loadc 5
                  loadc 7
                  store
                  pop

```

As an exercise translate:

$s_1 \equiv b = (&a) + 2;$ and $s_2 \equiv *(b + 3) = 5;$

64

```

code (s1s2) ρ =   loadc 7           loadc 5
                  loadc 2           loadc 17
                  loadc 10 // size of int[10] load
                  mul // scaling      loadc 3
                  add                 loadc 10 // size of int
                  loadc 17            mul // scaling
                  store               add
                  pop // end of s1     store
                                      pop // end of s2

```

65

Example: `int a[10], *b;` with $\rho = \{a \mapsto 7, b \mapsto 17\}$.

For the statement: `*a = 5;` we obtain:

```
codeL (*a) ρ = codeR a ρ = codeL a ρ = loadc 7
code (*a = 5;) ρ = loadc 5
                    loadc 7
                    store
                    pop
```

As an exercise translate:

$s_1 \equiv b = (&a) + 2;$ and $s_2 \equiv *(b+3) = 5;$

64

```
code (s1s2) ρ = loadc 7           loadc 5
                  loadc 2         loadc 17
                  loadc 10 // size of int[10] load
                  mul // scaling      loadc 3
                  add                loadc 1 // size of int
                  loadc 17           mul // scaling
                  store              add
                  pop // end of s1  store
                                      pop // end of s2
```

65

Example: `int a[10], *b;` with $\rho = \{a \mapsto 7, b \mapsto 17\}$.

For the statement: `*a = 5;` we obtain:

```
codeL (*a) ρ = codeR a ρ = codeL a ρ = loadc 7
code (*a = 5;) ρ = loadc 5
                    loadc 7
                    store
                    pop
```

As an exercise translate:

$s_1 \equiv b = (&a) + 2;$ and $s_2 \equiv *(b+3) = 5;$

64

Example: `int a[10], *b;` with $\rho = \{a \mapsto 7, b \mapsto 17\}$.

For the statement: `*a = 5;` we obtain:

```
codeL (*a) ρ = codeR a ρ = codeL a ρ = loadc 7
code (*a = 5;) ρ = loadc 5
                    loadc 7
                    store
                    pop
```

As an exercise translate:

$s_1 \equiv b = (&a) + 2;$ and $s_2 \equiv *(b+3) = 5;$

64

```

code (s1s2) ρ =   loadc 7           loadc 5
                  loadc 2           loadc 17
                  loadc 10 // size of int[10]   load
                  mul // scaling           loadc 3
                  add                       loadc 10 // size of int [1]
                  loadc 17                 mul // scaling
                  store                      add
                  pop // end of s1           store
                                           pop // end of s2

```

65

Example: `int a[10], *b;` with $\rho = \{a \mapsto 7, b \mapsto 17\}$.

For the statement: `*a = 5;` we obtain:

```

code_L (*a) ρ = code_R a ρ = code_L a ρ = loadc 7
code (*a = 5;) ρ = loadc 5
                  loadc 7
                  store
                  pop

```

As an exercise translate:

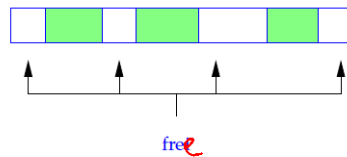
$s_1 \equiv b = (&a) + 2;$ and $s_2 \equiv *(b + 3) = 5;$

64

8 Freeing Occupied Storage

Problems:

- The freed storage area is still referenced by other pointers (dangling references).
- After several deallocations, the storage could look like this (fragmentation):



66

Potential Solutions:

- Trust the programmer. Manage freed storage in a particular data structure (free list) \implies `malloc` or `free` may become expensive.

- Do nothing, i.e.:

```

code free (e); ρ = code_R e ρ
                  pop

```

\implies simple and (in general) efficient.

- Use an automatic, potentially "conservative" Garbage-Collection, which occasionally collects certainly inaccessible heap space.

67

Potential Solutions:

- Trust the programmer. Manage freed storage in a particular data structure (free list) \implies `malloc` or `free` may become expensive.
- Do nothing, i.e.:

```
code free (e);  $\rho$  = codeR e  $\rho$ 
                    pop
```

\implies simple and (in general) efficient.

- Use an automatic, potentially "conservative" **Garbage-Collection**, which occasionally collects **certainly** inaccessible heap space.

67

$g(f)$

9 Functions

The definition of a function consists of

- a **name**, by which it can be called,
- a specification of the **formal parameters**;
- maybe a **result type**;
- a **statement part**, the **body**.

For **C** holds:

```
codeR f  $\rho$  = loadc _f = starting address of the code for f
```

\implies Function names must also be managed in the address environment!

68

Example:

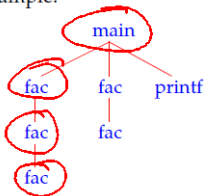
```
int fac (int x) {
  if (x <= 0) return 1;
  else return x * fac(x-1);
}
```

```
main() {
  int n;
  n = fac(2) + fac(1);
  printf ("%d", n);
}
```

At any time during the execution, several **instances** of one function may exist, i.e., may have started, but not finished execution.

An instance is created by a call to the function.

The recursion tree in the example:



69

We conclude:

The **formal parameters** and **local variables** of the different **instances** of the same function must be kept separate.

Idea:

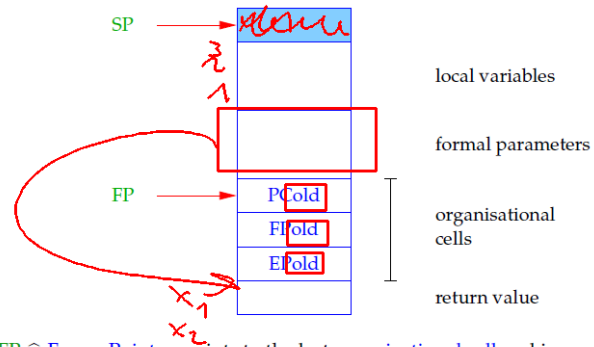
Allocate a special storage area for each instance of a function.

In sequential programming languages these storage areas can be managed on a stack. They are therefore called **Stack Frames**.

70

$f(x_1, x_2, \dots)$

9.1 Storage Organization for Functions



FP $\hat{=}$ Frame Pointer; points to the last organizational cell and is used to address the formal parameters and the local variables.

The caller must be able to continue execution in its frame after the return from a function. Therefore, at a function call the following values have to be saved into organizational cells:

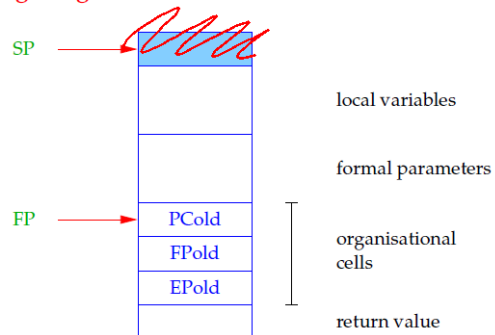
- the FP
- the continuation address after the call and
- the actual EP.

Simplification: The return value fits into one storage cell.

Translation tasks for functions:

- Generate code for the body!
- Generate code for calls!

9.1 Storage Organization for Functions



FP $\hat{=}$ Frame Pointer; points to the last organizational cell and is used to address the formal parameters and the local variables.

The caller must be able to continue execution in its frame after the return from a function. Therefore, at a function call the following values have to be saved into organizational cells:

- the FP
- the continuation address after the call and
- the actual EP.

Simplification: The return value fits into one storage cell.

Translation tasks for functions:

- Generate code for the body!
- Generate code for calls!

9.2 Computing the Address Environment

We have to distinguish two different kinds of variables:

1. **globals**, which are defined externally to the functions;
2. **locals/automatic** (including formal parameters), which are defined internally to the functions.



The address environment ρ associates pairs $(tag, a) \in \{G, L\} \times N_0$ with their names.

Note:

- There exist more refined notions of visibility of (the defining occurrences of) variables, namely **nested blocks**.
- The translation of different program parts in general uses different address environments!

73

9.2 Computing the Address Environment

We have to distinguish two different kinds of variables:

1. **globals**, which are defined externally to the functions;
2. **locals/automatic** (including formal parameters), which are defined internally to the functions.



The address environment ρ associates pairs $(tag, a) \in \{G, L\} \times N_0$ with their names.

Note:

- There exist more refined notions of visibility of (the defining occurrences of) variables, namely **nested blocks**.
- The translation of different program parts in general uses different address environments!

73

Example (1):

```

[0] int i;
    struct list {
        int info;
        struct list * next;
    } * l;
[1] int ith(struct list * x, int i) {
    if (i ≤ 1) return x → info;
    else return ith(x → next, i - 1);
}

[2] main() {
    int k;
    scanf("%d", &i);
    scanlist(&l);
    printf("\n\t%d\n", ith(l, i));
}

```

address environment at [0]

$\rho_0 :$	i	\mapsto	$(G, 1)$
	l	\mapsto	$(G, 2)$
	ith	\mapsto	$(G, _ith)$
	$main$	\mapsto	$(G, _main)$
	...		

74

Example (2):

```

[0] int i;
    struct list {
        int info;
        struct list * next;
    } * l;
[1] int ith(struct list * x, int i) {
    if (i ≤ 1) return x → info;
    else return ith(x → next, i - 1);
}

[2] main() {
    int k;
    scanf("%d", &i);
    scanlist(&l);
    printf("\n\t%d\n", ith(l, i));
}

```

[1] inside of ith:

$\rho_1 :$	i	\mapsto	$(L, 2)$
	x	\mapsto	$(L, 1)$
	l	\mapsto	$(G, 2)$
	ith	\mapsto	$(G, _ith)$
	$main$	\mapsto	$(G, _main)$
	...		

75

Example (3):

```

0 int i;
  struct list {
    int info;
    struct list * next;
  } * l;

1 int ith (struct list * x, int i) {
  if (i ≤ 1) return x → info;
  else return ith (x → next, i - 1);
}

2 main () {
  int k;
  scanf ("%d", &i);
  scanlist (&l);
  printf ("\n\t%d\n", ith (l,i));
}

```

9.3 Calling/Entering and Leaving Functions

Be f the actual function, the **Caller**, and let f call the function g , the **Callee**.

The code for a function call has to be distributed among the Caller and the Callee:

The distribution depends on **who** has **which** information.

Actions upon **calling/entering** g :

- | | | |
|--|---------|--------------------|
| 1. Saving FP, EP | } mark | } available in f |
| 2. Computing the actual parameters | | |
| 3. Determining the start address of g | | |
| 4. Setting the new FP | | |
| 5. Saving PC and
jump to the beginning of g | } call | } available in g |
| 6. Setting the new EP | } enter | |
| 7. Allocating the local variables | } alloc | |

Actions upon **leaving** g :

- | | |
|---|----------|
| 1. Restoring the registers FP, EP, SP | } return |
| 2. Returning to the code of f , i.e. restoring the PC | |

Altogether we generate for a call:

```

codeR g(e1, ..., en) ρ = mark
                        codeR e1 ρ
                        ...
                        codeR em ρ
                        codeR g ρ
                        call n

```



where n = space for the actual parameters

Note:

- Expressions occurring as actual parameters will be evaluated to their **R-value** \implies **Call-by-Value**-parameter passing.
- Function g can also be an **expression**, whose **R-value** is the start address of the function to be called ...

Actions upon calling/entering g :

- | | | | | | | |
|--|---|-------|---|------------------|---|------------------|
| 1. Saving FP, EP | } | mark | } | available in f | | |
| 2. Computing the actual parameters | | | | | | |
| 3. Determining the start address of g | | | | | | |
| 4. Setting the new FP | } | call | | | | |
| 5. Saving PC and
jump to the beginning of g | | | | | | |
| 6. Setting the new EP | } | enter | | | } | available in g |
| 7. Allocating the local variables | | | | | | |

Actions upon leaving g :

- | | | |
|---|---|--------|
| 1. Restoring the registers FP, EP, SP | } | return |
| 2. Returning to the code of f , i.e. restoring the PC | | |

Altogether we generate for a call:

```

codeR g(e1, ..., en) ρ = mark
                        codeR e1 ρ
                        ...
                        codeR em ρ
                        codeR g ρ
                        call n
    
```

where n = space for the actual parameters

Note:

- Expressions occurring as actual parameters will be evaluated to their R-value \implies Call-by-Value-parameter passing.
- Function g can also be an expression, whose R-value is the start address of the function to be called ...