2IC30: Computer systems
Translating Higher Languages

Jan Friso Groote



Technische Universiteit Eindhoven University of Technology

Where innovation starts

A program in machine code (PP2 processor)

Another Example

```
;
; INPUT: R0 = N , precondition: 0 <= N
; OUTPUT: R1 = Fib(N) , R2 = Fib(N+1)
; USES: R3
;</pre>
```

Machine code

```
0010001000000000000010010000000000000101000000000000000000000000000000101001001100100000001001001001000000110001110100001010001000000000000001000000011111111010000100010111110001
```



```
; Rob Hoogerwoord.
                                                                   In PP2 assembly
;
     INPUT: R0 = N
                         , precondition: 0 <= N
    OUTPUT:
             R1 = Fib(N), R2 = Fib(N+1)
      USES: R3
;
;
;
                                   ; invariant: m = N - R0 and
                LOAD
                      R1
                LOAD
                      R2
                                     R1 = Fib(m) and R2 = Fib(m+1)
                CMP
                      R0
                                     test RO, to prepare the condition codes
                                     enter the "loop"
                           +5
                BRA
                LOAD
                      R3
                          R1
                                     R2 = Fib(m+1) and R3 = Fib(m)
                LOAD
                      R1
                          R2
                                     R1 = Fib(m+1) and R2 = Fib(m+1) and R3 = Fib(m)
                ADD
                      R2
                          R3
                                     R1 = Fib(m+1) and R2 = Fib(m+2)
                TRCS
                                      abort when unsigned overflow, i.e., carry set
                                     R0 := R0 - 1, i.e.: m := m + 1, hence:
                      R0
                           1
                SUB
                                     R1 = Fib(m) and R2 = Fib(m+1)
                          -6
                                   ; if R0 = 0 then R1 = Fib(N)
                BNE
                RTS
```

@END



Assembly Language (1)

Problem

- Manually counting branch distances still is tedious and error prone
- Manually calculating memory addresses is tedious and error prone

Solution

• Use names to identify the targets of branch instructions: *labels*



Assembly Language (1)

In ARM assembly

```
; Rob Hoogerwoord;
                        , precondition: 0 <= N
     INPUT: RO = N
    OUTPUT: R1 = Fib(N), R2 = Fib(N+1)
      USES: R3
   Fibonacci: MOV
                     R1,#0
                                 ; invariant: m = N - RO and
                     R2,#1
                                 ; R1 = Fib(m) and R2 = Fib(m+1)
                MOV
                                 ; test RO, to prepare the condition codes
                     RO,#0
                CMP
                     Fibon while; enter the "loop"
                     R3,R1
    Fibon do:
               MOV
                                 ; R2 = Fib(m+1) and R3 = Fib(m)
                                 ; R1 = Fib(m+1) and R2 = Fib(m+1) and R3 = Fib(m)
                MOV
                     R1,R2
                ADDS R2,R2,R3
                                 ; R1 = Fib(m+1) and R2 = Fib(m+2)
                                 ; software interrupt when (unsigned) overflow
                SWIOV
                SUBS R0, #1
                                 ; R0 := R0 - 1, i.e.: m := m + 1, hence:
                                 ; R1 = Fib(m) and R2 = Fib(m+1)
  Fibon while: BNE Fibon do
                                 ; if R0 = 0 then R1 = Fib(N)
Fibonacci end : MOV PC,LR
                                 ; return to invoking program
```



Memory Addressing (0)

- Suppose an integer variable is located at word 13
- Loading this integer into register R1 (say):

```
MOV R0, #13
LDR R1, [R0]
```

- This is still error prone! Therefore: *always* use names!
- Calling the variable x (say), it can be *bound* to its location in the data segment by means of an "EQU" pseudo-instruction:

```
.equ x 13; now x equals 13
.set x 13; equivalent to .equ x 13
x = 13; also equivalent to .equ x 13
```

• The same instruction to load x into R1:

```
MOV R0, #x
LDR R1, [ R0 ]
```

Does not work for arbitrary *x*. Why?

• Advantages: less error prone and better readable



Memory Addressing (0)

• Calling the variable x (say), it can be *bound* to its location in the data segment by means of an "EQU" pseudo-instruction:

```
.equ x 134657463; now the address of x equals 134657463 .set x 134657463 x = 134657463
```

• The same instruction to load x into R1 (leads to an error by the assembler):

```
MOV R0, #x
LDR R1, [ R0 ]
```

• The same instruction to load x into R1:

```
LDR R0, =x
LDR R1, [ R0 ]
```



Example of a concrete assembly program.

```
.global main
.data
.equ constant 45
var_x: .int
                    10
var_y: .byte 'A', 0x31, 32, 0x33, 34, $35
string: .asciz "This is a tekst.\n"
main:
          LDR R0, =var_x
          LDR R3, [R0]
          STR R3, [SP, #-1]!
          BL subroutine
          STR R4, [SP], #1
          STR R3, [R4, #constant]
          B main
```

.end

Example of the use of the heap (0)

Swap values of variables x and y on the heap.

```
INPUT: x, y
  OUTPUT: y is old value of x, x is old value of y: x=y', y=x'
global main:
main:
             LDR GB, =global base
             LDR R0, [GB, #x]
                                    ; R0:=x
             LDR R1, [GB,#y]
                                    ; R1:=y
             STR R0, [GB,#y]
                                    ; y:=R0
                                                          Heap
             STR R1, [GB, #x]
                                    ; x:=R1
.data
   .equ x 0
            ; x resides at GB+0
    .equ y 1
                         ; y resides at GB+1
    .equ global base ....; definition of the global base
                                                           GB \rightarrow
```



Memory

Example of the use of the heap (1)

Reset an array a[0,...,n-1].

```
INPUT: R0, a; Precondition R0=n
   OUTPUT: Postcondition a[i]=0 for all 0<=i<n.
.global main
main:
                                        ; R1:=0 R1 represents i.
reset a:
               VOM
                     R1, #0
                     R3, =global base
               LDR
                                                                Heap
                     R3, #a
                                        ; R3:=GB+a, address where
               ADD
                                        ; array a starts.
while reset a: MOV
                     R2,R1
               SUBS R2, R0
                                        ; R2 := (R1-R0) R2 = i-n
                      endwhile reset
                                        ; if i=n, stop
               BEQ
                     R4,#0
                                        ; R4 := 0
               MOV
                                                                 GB \rightarrow
               STORE R4 [R3,R1]
                                        ; a[i]:=0
                     R1,#1
                                        ; i:=i+1
               ADD
                     while reset a
                                        ; loop back
               В
endwhile reset: ...
.data
                           ; base address of a
   .equ a ...
   .equ global base ... ; definition of global base
```

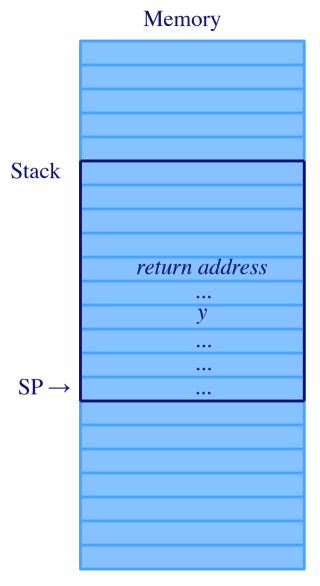
Memory a[n-1]a[2]a[1]a[0]



Memory Addressing (1)

Local variables: used only temporarily

- Local variables can be kept in registers, or ... on the *stack*!
- The stack grows *downwards* in memory; to *allocate*-- i.e. reserve -- 5 words on the stack, use: SUB SP 5.
- The addresses of these 5 words are SP, SP+1,..., SP+4.
- Better is to name them; if we use the word at position 3 as a local variable y (say), it can be *bound* to its location in the Stack data segment by means of an ".equ" pseudoinstruction: .equ y 3; now the address of y equals SP + y.





Stack addressing: allocating local variables

```
Example: A subroutine needs 6 words to store local variables.
procedure f()
var a, b, c, d, e, x: integer;
.....
end;
```

- ☐ Allocation: By *decreasing* SP by the amount −6− needed.
- □ Deallocation: By *increasing* SP by the amount –6– used.
- Allocation takes place at the *beginning* of the subroutine, *before* the *actual body* of the subroutine.
- Deallocation takes place at the *end* of the subroutine, *after* the *actual body* of the subroutine, but *before* the *return*.
- Every subroutine must leave the stack effectively unchanged!!!



Stack addressing: allocating local variables



Generation of programming languages

First generation: Machine code

Second generation: Assembly code

Third generation: Higher level languages, such as ALGOL 60,

Fortran, Cobol, ALGOL 68, PASCAL, C++,

Java.

Prof.dr. Frans Kruseman Aretz:

"With the third level languages we thought our troubles were over.

We could not be more wrong..."



Foto: TU/e



Generation of programming languages

First generation: Machine code

Second generation: Assembly code

Third generation: Higher level languages, such as ALGOL 60,

Fortran, Cobol, ALGOL 68, PASCAL, C++,

Java.

Fourth generation: Not so clear. Domain specific languages???

SQL, HTML?







Large(r) Projects: Program Design

- *Separate* program design and construction of the Assembly Language program: *firstly*, design and write the program in Java, C++, ...; *secondly*, translate that program into Assembly Language.
- Why???



Large(r) Projects: Program Design

- *Separate* program design and construction of the Assembly Language program: *firstly*, design and write the program in Java, Pascal, ..., or in a DSL (Domain Specific Language) or just plain mathematics; *secondly*, translate that program into Assembly Language.
- Why???
- Humans are very bad in handling complexity. Always work as abstract as as possible. Use mathematics, or any domain specific notation....



Example with recursion:

n! with a recursive procedure, n > 0:

```
procedure fac(var n:integer);
var h:integer;
begin
 if (n≠1) then begin
   h:=n;
   n := n-1;
   fac(n);
   n:=h*n;
 end
```



In assembler:

CMP: subtract without storing the results.

MULS: multiply arguments.

```
procedure fac(var n:integer):
var h:integer;
                            fac main: CMP R0 1
                                       BEQ fac exit
begin
                                       STOR R0 [--SP]
 if (n≠1) then begin
                                       SUB R0 1
   h:=n;
                                       BRS fac main
   n:=n-1;
                            fac ret: LOAD R1 [SP++]
   fac(n);
                                       MULS R0 R1
   n:=h*n;
                            fac exit: RTS
 end
end;
```

Note: parameter n is passed in register R0.

Local variable h is stored on the stack.



Questions?





Example: Translating a Java function into Assembly (1)

```
int PowerSum (int X, int N)
// requires: 0 \le N
// result: PowerSum(X,N) = (\Sigma i : 0 \le i \le N : X^i)
\{ \text{ int } x, y, z, n ; 
   z = 0; y = 1; x = X; n = N;
   while (n!=0)
   { while (n\%2 == 0)
       \{ y = y * (1+x) ; x = x * x ; n = n/2 ; \}
       z = z + y; y = y * x; n = n - 1;
   return z;
```



Example: Translating a Java function into Assembly (2)

- (0) Decide how the *parameters* will be transferred: parameters *X* and *N* are *values* to be transferred *to* the subroutine, and the function's *result* is to be transferred back *from* the subroutine to the calling program:
 - These three (integer) values are passed via the *Stack*:
 - Just *before* the call of the subroutine, the calling program:
 - --reserves 1 word on the stack for the function's *result*;
 - --then, it pushes X onto the stack;
 - --next, it pushes N onto the stack;
 - --finally, it calls the subroutine.
 - Directly *after* the call of the subroutine, the calling program:
 - --retrieves 1 the function's *result* from the reserved word for this purpose;
 - --then, it *cleans up* the stack;

```
return_addr ←SP

N

X

result
```

```
int PowerSum (int X, int N) 
// requires: 0 \le N 
// result: PowerSum(X,N) = (\Sigma i : 0 \le i \le N : X^i)
```

Example: Translating a Java function into Assembly (3)

; Calling sequence to invoke subroutine PowerSum:

```
call begin :
                SUB
                      SP 1
                                   ; allocate one word on the stack,
                                    for the result.
                LOAD
                      R0
                         X
                                   ; here we just write "X": may
                                     be more complicated.
                STOR
                      R0 [--SP]; push X onto the stack.
                LOAD
                      R0
                          N
                                   ; here we just write "N".
                      R0 [--SP]; push N onto the stack.
                 STOR
                BRS
                      PowerSum
                                   ; call subroutine PowerSum.
    call end :
                LOAD
                      R0 [SP+2]; retrieve PowerSum's result...
                 STOR
                      R0
                                   ; ... and store it somewhere,
                          \boldsymbol{z}
return_addr ←SP
                                     here just called "z"
   N
                       SP 3
                ADD
                                      clean up Stack: deallocate the 3
                                      words used
  result
                                      and continue...
```



Example: Translating a Java function into Assembly (4)

(1) When execution of the subroutine starts, the top of the stack has this

```
←SP
                                                                   \boldsymbol{\mathcal{X}}
          ] = "return address of the calling pro
RAM [SP
                                                                   \boldsymbol{n}
RAM[SP+1] =
                                                               return_addr
RAM[SP+2] = X
RAM[SP+3] = "word for the (anonymous) function
```

(2) Now, decide where *local variables* x, y, z, n will be allocated. Because the final value of z will be returned as the result, z may be identified with the result word at SP+3. Variables x, y and n are put on the stack:

```
RAM[SP+0]
RAM[SP+1]
RAM[SP+2]
RAM[SP+3+3]
```

Compensate for two local variables on the stack.

```
int PowerSum (int X, int N)
// requires: 0 \le N
   result: PowerSum(X, N) = (\Sigma i : 0 \le i \le N)
   int x, y, z, n;
   z = 0; y = 1; x = X; n = N;
   while (n!=0)
```

Example: Translating a Java function into Assembly (1)

```
int PowerSum (int X, int N)
// requires: 0 \le N
// result: PowerSum(X,N) = (\Sigma i : 0 \le i \le N : X^i)
\{ \text{ int } x, y, z, n ; 
   z = 0; y = 1; x = X; n = N;
   while (n!=0)
   { while (n\%2 == 0)
       \{ y = y*(1+x); x = x*x; n = n/2; \}
      z = z + y; y = y * x; n = n - 1;
   return z;
```



Example: Translating a Java function into Assembly (5)

```
.equ local vars
                           z = 0; y = 1; x = X; n = N;
.equ PowerSum N
                  1+
                           while (n!=0)
                  2+]
.equ PowerSum X
                           { while (n\%2 == 0)
.equ PowerSum z
                  3+1
.equ Powersum x
                               \{ y = y * (1+x) ; x = x * x ; n = n/2 ; \}
.equ PowerSum y
                               z = z + y; y = y * x; n = n - 1;
.equ PowerSum n
                         SP local vars
                                                ; reserve room on the stack
PowerSum begin:
                  SUB
                  T<sub>i</sub>OAD
                         R0
                                                ; initialize variables
                             [SP+PowerSum z]
                  STOR
                         R0
                                                z \leftarrow 0
                                                ; R0 ← 1
                  LOAD
                         R0
    RAM[SP+0]
                  STOR
                         R0
                             [SP+PowerSum y]; y \leftarrow 1
    RAM[SP+1]
    RAM[SP+2]
                                                ; R0 \leftarrow X
                         R0
                  LOAD
                             [SP+PowerSum X]
    RAM[SP+6]
                  STOR
                         R0
                             [SP+PowerSum x]
                                                x \leftarrow X
                  LOAD
                         R0
                             [SP+PowerSum N]
                                                : R0 \leftarrow N
                  STOR
                         R0
                             [SP+PowerSum n]
                                                n \leftarrow N
PowerSum while0: BEQ
                         PowerSum od0
                                                ; exit outer repetition
                                                ; flags reflect value of n
  continued on next page...
```

Example: Translating a Java function into Assembly (6)

```
PowerSum while1: LOAD R0 [SP+PowerSum n]
                                              : R0 \leftarrow n
                                              : R0 ← n%2
                  AND
                        R0
                         PowerSum od1
                                              ; IF not n%2=0, THEN exit
                  BNE
                                              ; inner repetition
PowerSum dol:
                  LOAD R0 [SP+PowerSum x]
                                             ; R0 \leftarrow x
                  ADD
                        R0
                                             ; R0 \leftarrow x + 1
                  LOAD R1 [SP+PowerSum y]; R1 ← y
     RAM[SP+0]
                  MULS R1
                                             ; R1 \leftarrow y \times (x+1)
                           R0
     RAM[SP+1]
     RAM[SP+2]
                  STOR R1 [SP+PowerSum y]; y ← R1
     RAM[SP+6]
                  LOAD R0 [SP+PowerSum x] ; R0 \leftarrow x
                  LO
                           z = 0; y = 1; x = X; n = N;
                  MU]
                           while (n!=0)
                  ST
                           { while (n\%2 == 0)
                  LO
                  SH
                               \{ y = y * (1+x) ; x = x * x ; n = n/2 ; \}
                  ST
                              z = z + y; y = y * x; n = n - 1;
                  BRZ
                                              ; inner repetitionPowerSum woodsileit
```

Example: Translating a Java function into Assembly (6)

```
PowerSum while1: LOAI
                           z = 0; y = 1; x = X; n = N;
                  AND
                           while (n!=0)
                  BNE
                           { while (n\%2 == 0)
PowerSum do1:
                  LOAI
                               \{ y = y*(1+x); x = x*x; n = n/2; \}
                  ADD
                               z = z + y; y = y * x; n = n - 1;
                  LOAI
     RAM[SP+0]
                  MULS KI RU
     RAM[SP+1]
     RAM[SP+2]
                  STOR R1 [SP+PowerSum y]; y ← R1
     RAM[SP+6]
                  LOAD R0 [SP+PowerSum x] ; R0 \leftarrow x
                  MULS RO RO
                                            : R0 \leftarrow x \times x
                  STOR R0 [SP+PowerSum x]; x \leftarrow R0
                  LOAD R0 [SP+PowerSum n] ; R0 \leftarrow n
                  SHIFTRIGHT RO
                                            ; n \leftarrow n/2
                  STOR R0 [SP+PowerSum n]; n \leftarrow R0
                  BRA PowerSum while1
                                            ; jump back to beginning of
                                             ; inner repetitionPowerSum od1
```

Examp

PowerSum_while1:

```
 \begin{split} z &= 0 \; ; \; y = 1 \; ; \; x = X \; ; \; n = N \; ; \\ \text{while} \; (n! = 0) \\ \{ \; \; \text{while} \; (n\%2 == 0) \\ \{ \; \; y = y*(1 + x) \; ; \; x = x*x \; ; \; n = n/2 \; ; \; \} \\ z &= z + y \; ; \; y = y*x \; ; \; n = n-1 \; ; \end{split}
```

```
PowerSum dol:
PowerSum od1:
                  LOAD R0 [SP+PowerSum z] ; R0 \leftarrow z
                        R0 [SP+PowerSum y] ; R0 \leftarrow z + y
                   ADD
                   STOR R0 [SP+PowerSum z]; z \leftarrow R0
     RAM[SP+0]
                  LOAD R0 [SP+PowerSum y]; R0 ← y
     RAM[SP+1]
     RAM[SP+2]
                  LOAD R1 [SP+PowerSum x]; R1 \leftarrow x
     RAM[SP+6]
                   MULS RO
                            R1
                                               ; R0 \leftarrow y \times x
                   STOR R0 [SP+PowerSum y]; y ← R0
                   LOAD R0 [SP+PowerSum n]; R0 \leftarrow n
                                               : R0 \leftarrow n-1
                   SUB
                        R0 1
                   STOR R0 [SP+PowerSum n]; n \leftarrow R0
                        PowerSum while0; jump back to beginning of
                   BRA
                                               ; outer repetitie
```

Example: Translating a Java function into Assembly (8)

```
\{ y = y * (1+x) ; x = x * x ; n = n/2 ; \}
PowerSum while1: LOAD
                               z = z + y; y = y * x; n = n - 1;
                 AND
                 BNE
                           return z;
PowerSum do1:
                                             jump back to the inner loop
                       PowerSum_while1
                 BRA
                                            ; inner repetition
PowerSum od1:
                 BRA
                       PowerSum while0
                                           ; jump back to the
                                            ; outer repetition
                                            ; remove local vars
 PowerSum od0 : ADD SP local vars
                                            ; result value is already
                 RTS
                                            ; there: just RETURN
        RAM[SP+0]
        RAM[SP+1]
        RAM[SP+2]
        RAMISP+61
```

Questions?





Software interrupts or software exceptions.

- ■SWI #number. ; ARM processor.
- □TRAP ; PP2 processor.
- Call a special subroutine (in supervisor mode on the ARM).
- ☐ The effect of a software interrupt is:

```
LR :=PC-4 (points to the next instruction)
```

$$PC := 8$$
 (at address 8 there is a jump to the required

software handling code)

CPSR := APSR (save the status flags).

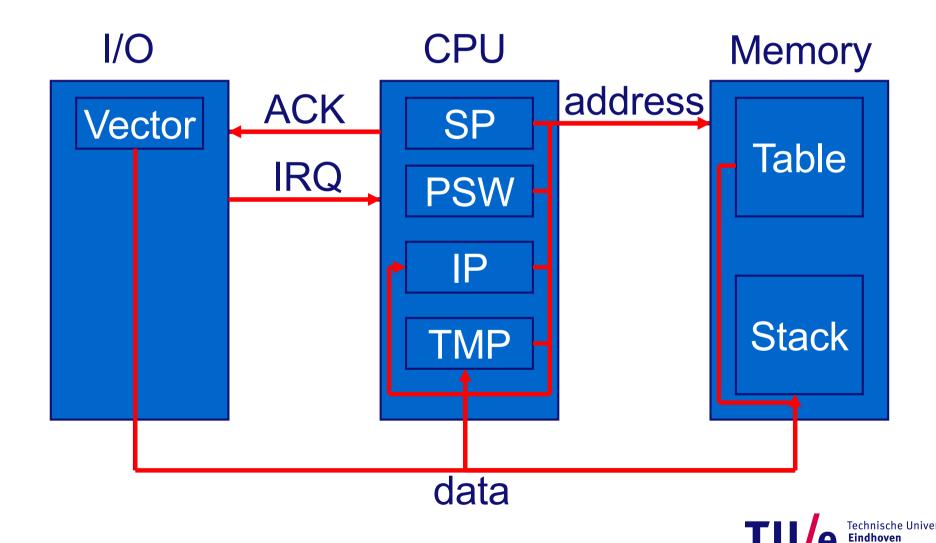


Hardware interrupts happen

- ☐ Device passes interrupt to CPU
- □ CPU finishes current instruction and sends an acknowledgement to the device
- Device provides additional information (e.g. IRQ-number, target address or table-index)
- □{CPU fetches information and temporarily stores this}
- □CPU pushes IP (and often PSW) onto the stack
- ☐ CPU calls interrupt handler (which one depends on the additional information)



Hardware: interrupt (example)



Software: handle interrupt

□ Save all registers. ☐ Find out which device caused an interrupt □Get additional information (status etc). ☐ Handle the interrupt (e.g. I/O error, start next program etc.) □ Signal to the device that the interrupt was serviced □ Restore all registers □ Execute ReTurn from Interrupt instruction, or ERET.



Interrupts: remarks

- ☐ Interrupt handlers must be transparent.
- □ If execution of an interrupt handler takes too long (longer than the period of the interrupt) the main program will not proceed.
- □ High priority interrupts are used for time critical applications (e.g. while burning a DVD, this was once modern... ⊖)
- ☐ Interrupts are needed for multitasking.

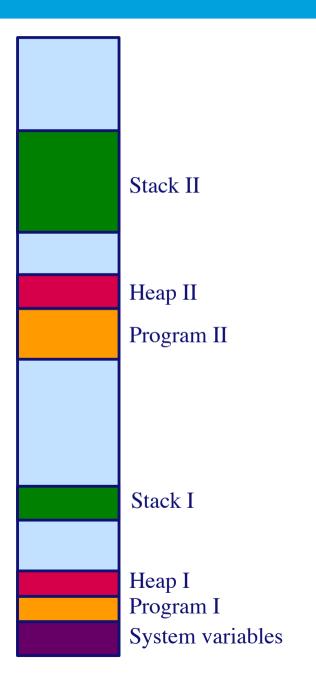


Questions?





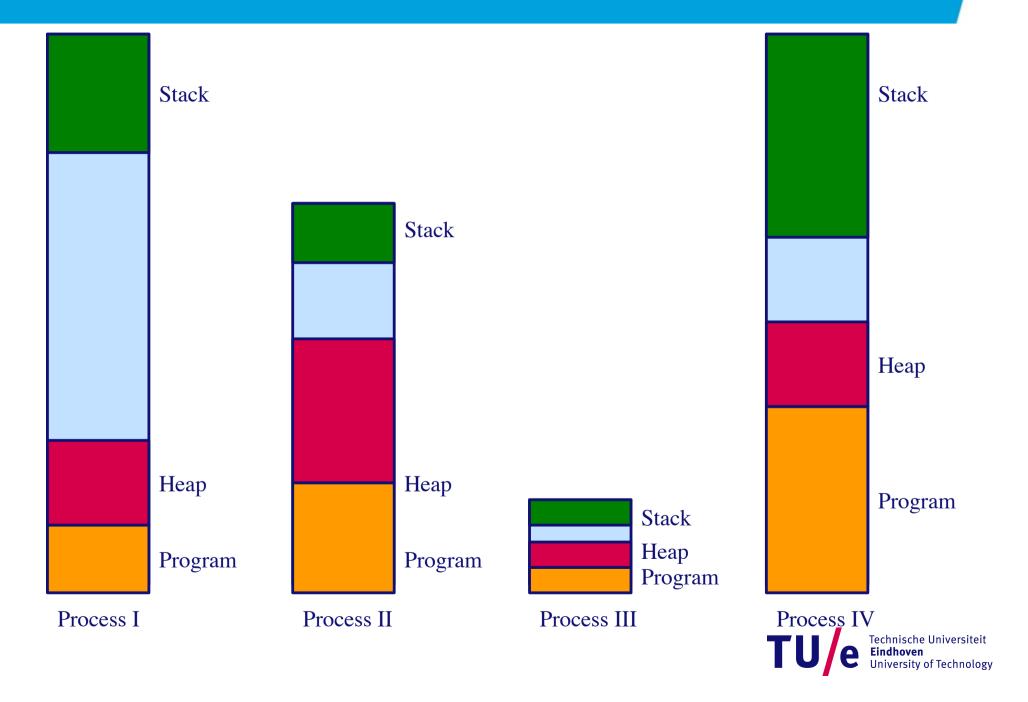
Elementary multitasking.



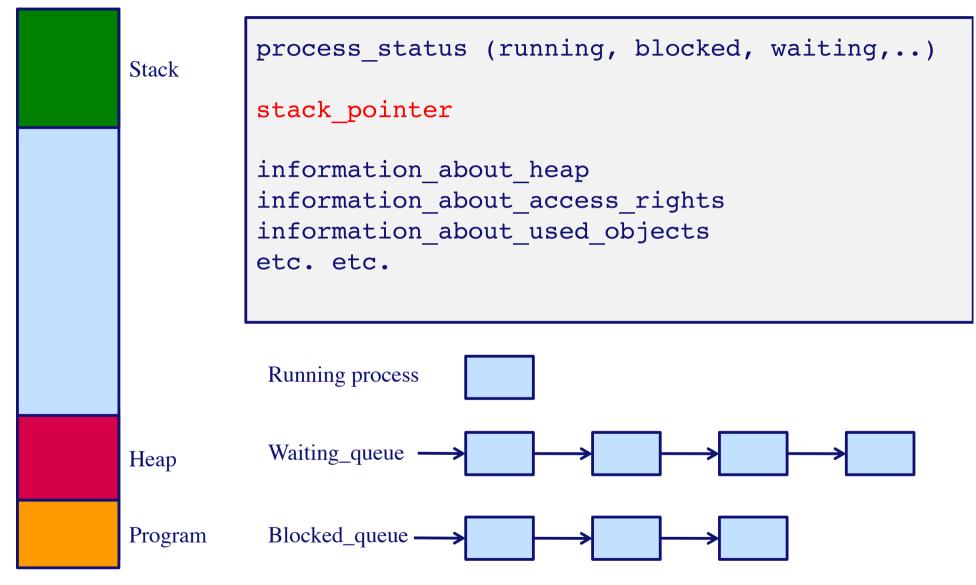
We can store more programs in memory, with their own stack and heap.



Another view on processes.



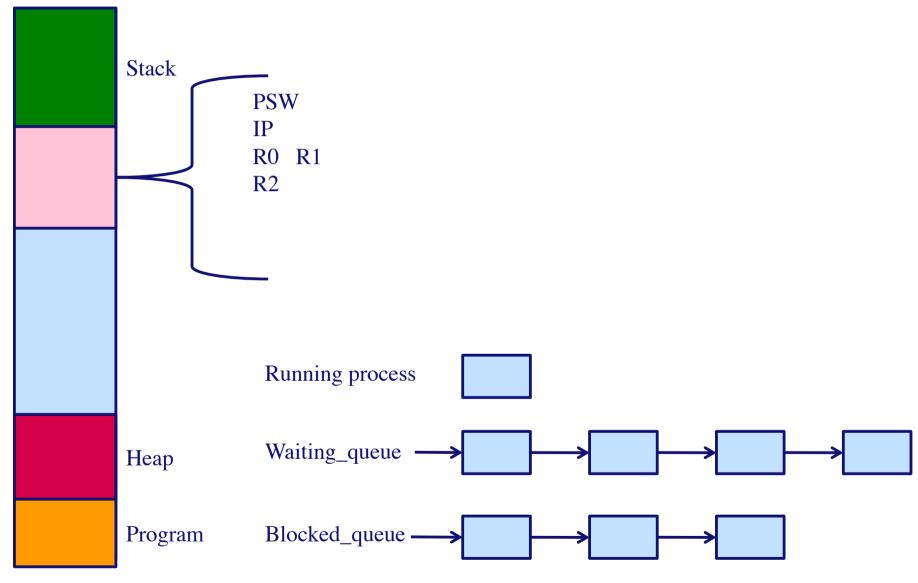
Each process has a process record







Blocked/waiting processes have info on the stack



Waiting/idling process

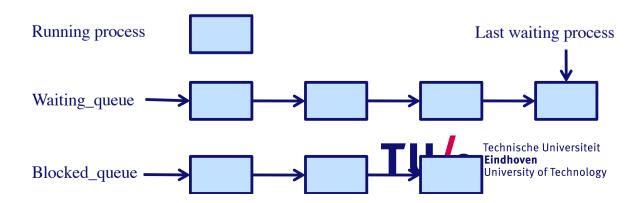


```
// When a process timer interrupt comes.
Save registers on stack of the current process;
If the waiting queue is not empty
 Save the stackpointer in the current process record;
 Set the status of the current process to wait;
 Put the process record at the end of the waiting list;
 Take a new process record from the waiting queue;
 Set its status to running;
 Put the stackpointer of the new process in the stack register.
Restore the registers of the new process;
RTI; Return from interrupt.
                              Running process
                              Waiting_queue
                                                                   Technische Universiteit
```

University of Technology

```
// When a process timer interrupt comes.
; Save registers on stack of the current process
PUSH R0;
; If the waiting queue is not empty
LOAD RO [waiting queue];
     finish context switch;
BEO
PUSH R1;
PUSH R2;
; Save the stackpointer in the current process record;
LOAD R1 [running process]
STOR SP [R1 + stackpointer];
; Set the status of current process to waiting;
LOAD R2 waiting status;
                                 Running process
STOR R2 [R1 + status];
                                 Waiting_queue
                                                                       Technische Universiteit
                                 Blocked_queue
                                                                       University of Technology
```

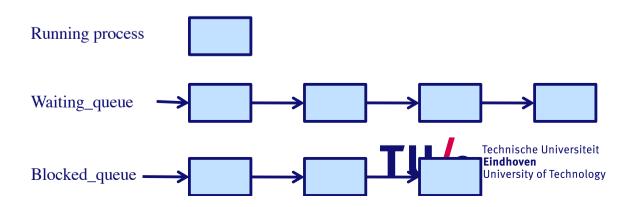
```
; Put the process record at the end of the waiting list;
    LOAD R2 [last waiting process];
    STOR R1 [R2+next];
    STOR R1 [last waiting process];
; Remove the first process record from the waiting queue;
; Set its status to running;
    LOAD R1 running status;
    STOR R1 [R0 + status]; R0 points to the new process.
; Put the stackpointer of the new process in the stack register.
    LOAD SP [R0 + stackpointer]
```



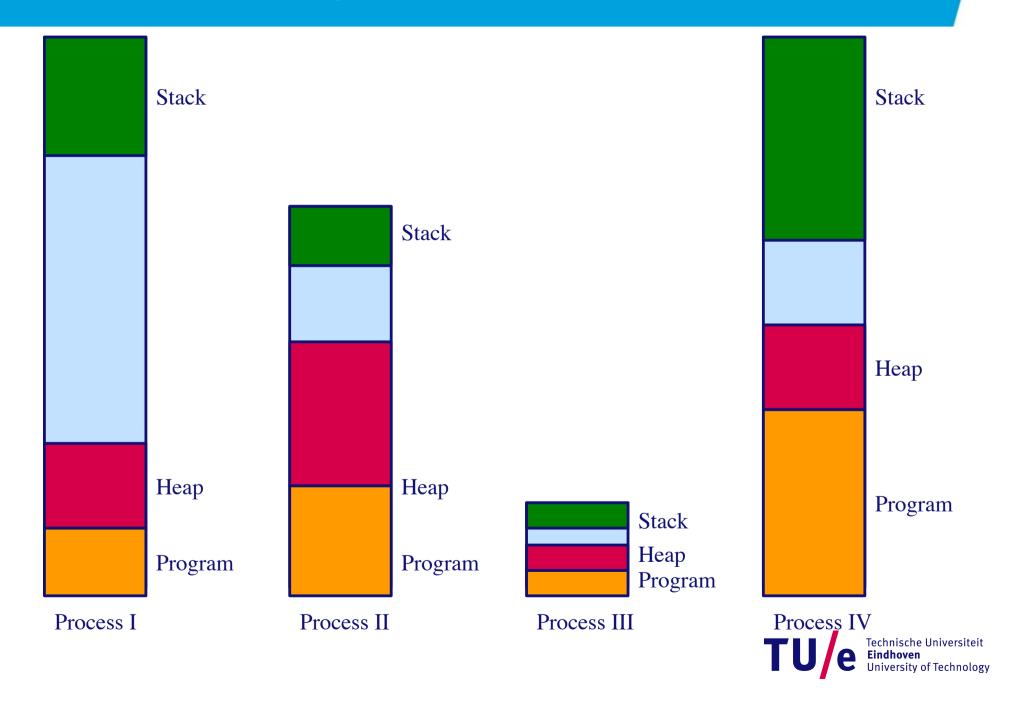
```
PULL R2;
PULL R1;

finish_context_switch:
; Restore the registers of the new process;
PULL R0;

RTI; Return from interrupt.
```



The processor jumps from process to process.



Questions?





Non atomicity.

Translation to assembly code:

Process 1: Process 2:
$$x:=x+1$$
 $x:=x+1$

LOAD R0 [BP+
$$x$$
]
ADD R0 1
STOR R0 [BP+ x]

Executing both processes can result in

when both processes are executed simultaneously. Probability is low. This is hard to test.



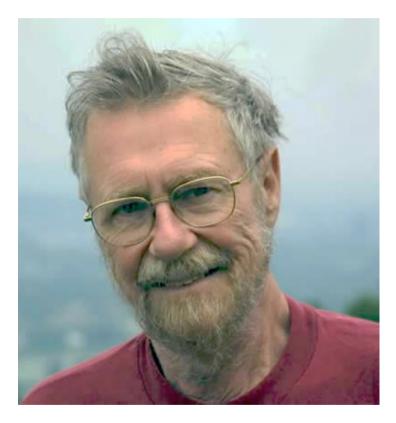
Semaphores/mutexes: tool for process synchronisation





Passeer (Pass):
 if passable
 then indicate non passability
 else
 move to blocking queue
 queue for this semaphore

Vrijgeven (Release):
 if queue is empty
 then indicate passable
 otherwise move a process in the queue
 to the processor waiting queue.



Edsger Wiebe Dijkstra (1930-2002)

These days called mutex variables.



Questions?





Summary

What did we learn:

- We can write programs in assemby (PP2, ARM).
- The stack is used for parameters of functions, local variables, the computation of expressions and returning results.
- We can systematically translate higher level programs into assembly code. This is what compilers do.
- We saw how to transform a single processor machine into a multithreading machine.
- We saw the need for mutual exclusion, and know that semaphores/mutexes can be used for this purpose.

