Embedded Systems Programming - PA8001

http://bit.ly/15mmqf7

Lecture 5

Mohammad Mousavi

m.r.mousavi@hh.se

Center for Research on Embedded Systems
School of Information Science, Computer and Electrical Engineering

Concurrency

```
struct Params params;
```

```
void controller_main() {
  int dist, signal;
  while(1){
    dist = sonar_read();
    control(dist,
            &signal,
            &params);
    servo_write(signal);
  }
}
```

```
void decoder_main() {
   struct Packet packet;
   while(1){
      radio_read(&packet);
      decode(&packet,&params);
   }
}
```

Providing means for two mains to execute concurrently! As if we had 2 CPUs!

Concurrent programming

```
main(){
  spawn(decoder_main);
  controller_main();
}
```

Notice that spawn takes a *function* as an argument.

spawn: provides an extra **Program Counter** and **Stack Pointer**

We also need to interleave the threads.

yield: seizing control to another thread

Calling `yield()`

Explicitly

```
ld a, r1
ld b, r2
add r, r2
st r2, c
jsr yield
ld c, r0
cmp #37, r0
ble label34
...
```

```
yield:
   sub #2, sp
   ...
   mov #0, r0
   rts
```

Calling `yield()`

Implicitly

```
ld a, r1
ld b, r2
add r, r2
st r2, c
```

⟵ Interrupt on pin 3!

```
ld c, r0
cmp #37, r0
ble label34
...
```

```
vector_3:
    push r0-r2
    jsr yield
    pop r0-r2
    rti
```

```
yield:
    sub #2, sp
    ...
    mov #0, r0
    rts
```

Installing interrupt handlers

```
#include<avr/interrupt.h>

...
ISR(interrupt_name){
...
 // code as in a function body!
...
}
```

Preventing interrupts in avr-gcc

```
cli();
// ... code that must not be interrupted ...
sei();
```

Preventing interrupts

Why should we consider disabling interrupts? What parts of the program should be protected?

The critical section problem

What if `params` is read (by the controller) at the same time as it is written (by the decoder)?

I.e., what if the scheduler interleaves read and write instructions from the controller and the decoder?

Mutual exclusion: a central issues in concurrency.

Our embedded system

```
struct Params p;
```

```
while(1){
   ...
   p.minDistance = e1;
   p.maxSpeed = e2;
}
```

```
while(1){
   local_minD = p.minDistance;
   local_maxS = p.maxSpeed;
   ...
}
```

Possible interleaving

```
p.minDistance = 1;
p.maxSpeed = 1;
                              local_minD = 1;
p.minDistance = 200;
p.maxSpeed = 150;
                              local_maxS = 150
```

The classical solution

Apply an access protocol to the critical sections that ensures mutual exclusion

Require that all parties follow the protocol

Access protocols are realized by means of a shared datastructure known as a mutex or a lock.

Mutual exclusion

```
struct Params p;
mutex m;
```

```
while(1){
    ...
    lock(&m);
    p.minDistance = e1;
    p.maxSpeed = e2;
    unlock(&m);
}
```

```
while(1){
    lock(&m);
    local_minD = p.minDistance;
    local_maxS = p.maxSpeed;
    unlock(&m);
    ...
}
```

The datatype `mutex` and the operations `lock` and `unlock` are defined in the kernel: each mutex has a queue of threads that are not in the ready queue. The operations move threads to and from the ready queue!

What we have learned ...

- We know how to read and write to I/O device registers
- We know how to run several computations in parallel by time-slicing the CPU
- We know how to protect critical sections by means of a mutex

But ...

Still not satisfied!

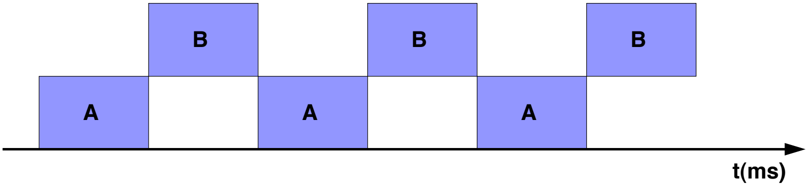
```
void controller_main() {
  int dist, signal;
  while(1){
    dist = sonar_read();
    control(dist,
            &signal,
            &params);
    servo_write(signal);
  }
}
```

```
void decoder_main() {
   struct Packet packet;
   while(1){
      radio_read(&packet);
      decode(&packet,&params);
   }
}
```

⟵ Time slicing ⟶

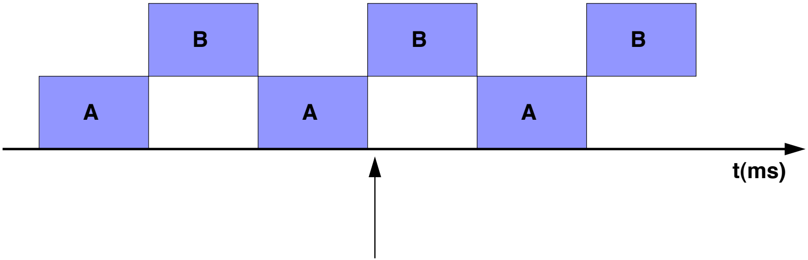
Each thread gets half of the CPU cycles, irrespective of whether it is waiting or computing!

Consequence 1



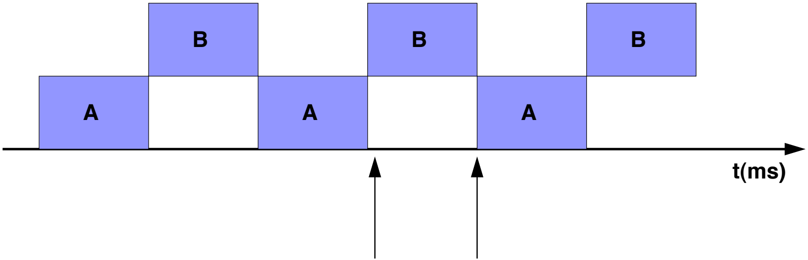
Say each thread gets **T**ms for execution, both waiting and computing!

Consequence 1



Say that an event that **A** is waiting for occurs now . . .

Consequence 1

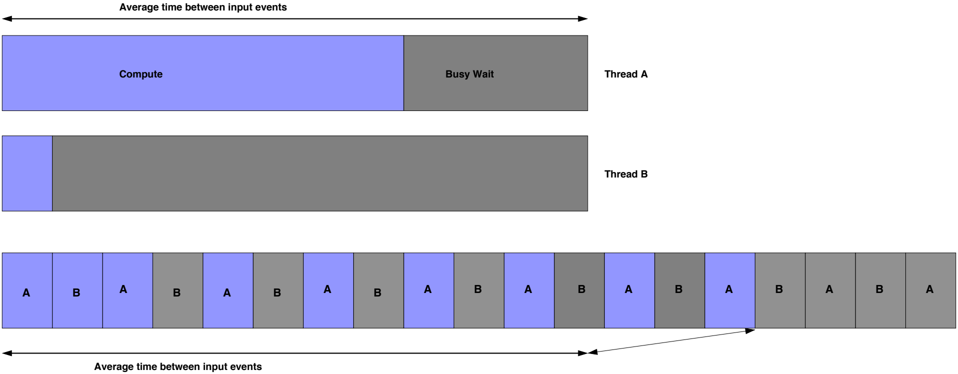


. . . it will not be noticed until now!

Consequence 1

With **N** threads in the system, each getting **T**ms for execution, a status change might have to wait up to **T*(N-1)**ms to be noticed!

Consequence 2



Busy waiting makes waiting indistinguishable from computing.
Thread A cannot keep up with event rate!

Busy waiting and Time slicing

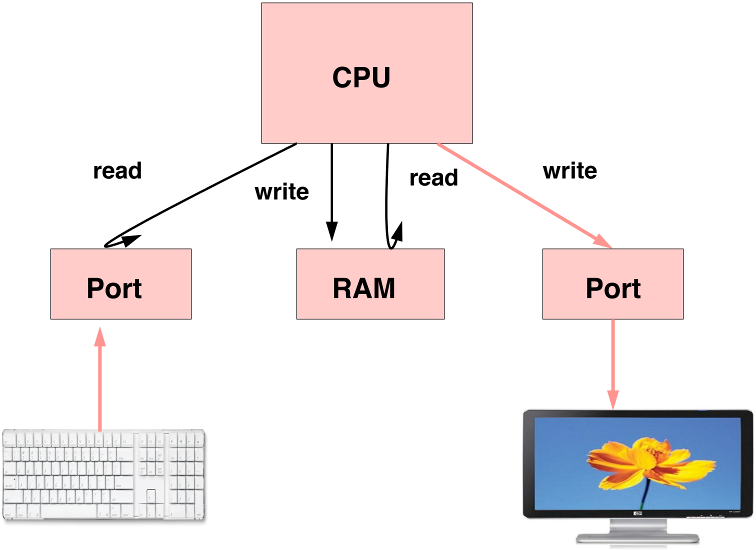
Minus . . .

1. Not a satisfactory technique for input synchronization if the system must meet real-time constraints!
2. Not a satisfactory technique for a system that is battery driven: 100% CPU cycle usage (100% power usage!).

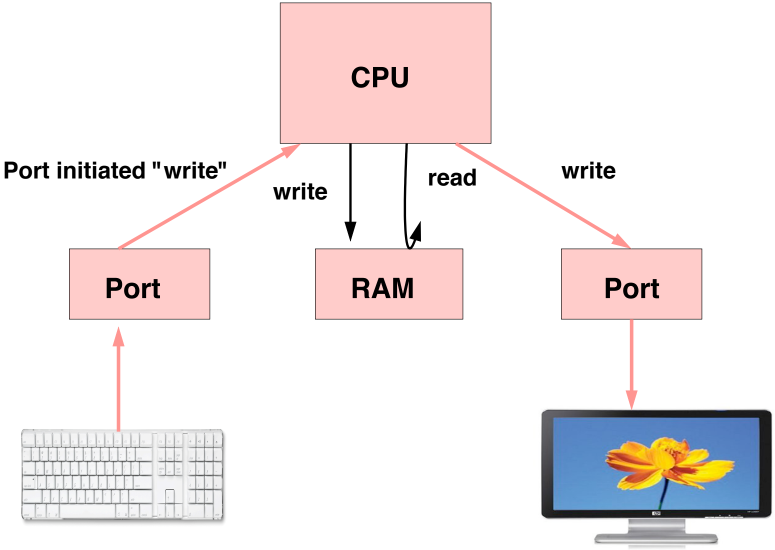
Could we do otherwise?

An input synchronization technique that does not require the receiver of data to actively ask whether data has arrived.

The naked computer – a mismatch



The naked computer – alternative



An analogy

You are expecting delivery of your latest web-shop purchase

Busy waiting

Go to the post-office again and again to check if the delivery has arrived.

Reacting to an interrupt

Receive a note in your mailbox that the goods can be picked up.

The CPU reacts to an interrupt signal by executing a designated ISR (interrupt service routine)

This has consequences for the way we structure programs. They become inside-out!

ISRs vs functions

Busy waiting

We defined functions like `sonar_read` that can be called in the program. The CPU decides when to call the function:

```
while(1){
   sonar_read();
   control();
}
```

Input detection = the exit from the busy waiting fragment (a function return)

Reacting

We define ISRs. These are not called from the program, but the code is executed when an interrupt occurs:

```
ISR(SIG_SONAR){
  control();
}
```

Input detection = invocation of the ISR (as if the hardware did a function call)

Two ways of organizing programs

CPU centric

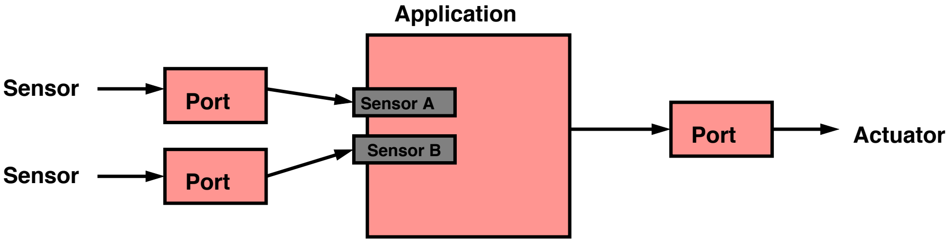
One thread of control that runs from start to stop (or forever) reading and writing data as it goes.

Reacting CPU

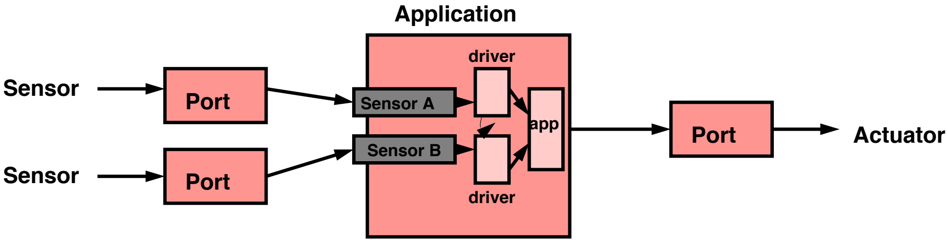
A set of code fragments that constitute the reactions to recognized events.

The main part of the course from now on will focus on the reactive view.

The reactive embedded system



The reactive embedded system



Reactive Objects

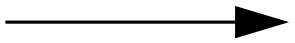
Boxes

Represent software or hardware reactive objects that:

- Maintain an internal state (variables, registers, etc)
- Provide a set of methods as reactions to external events (ISRs, etc)
- Simply rest between reactions!

Arrows

Represent event or signal or message flow between objects that can be either

- asynchronous 
- synchronous 

Hardware objects

Hardware devices are reactive objects

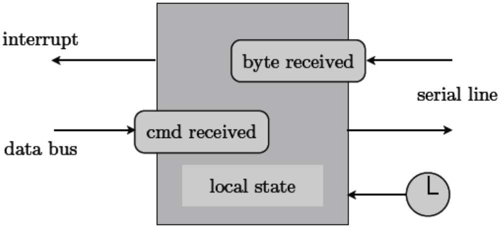
A black box that does nothing unless stimulated by external events.

Serial port - state

Internal registers

Serial port - stimuli

- Signal change
- Bit pattern received
- Clock pulse



Serial port - emissions

- Signal change
- Interrupt signal

Software objects

We would like to regard software objects as reactive objects . . .

The Counter example

```
class Counter{
  int x;
  Counter(){x=0;}
  void inc(){x++;}
  int read(){return x;}
  void reset(){x=0;}
  void show(){
     System.out.print(x);}
}
```

Counter state

`x`

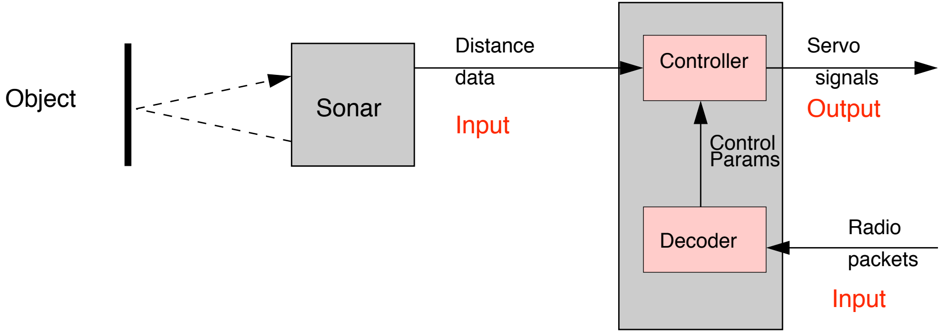
Counter - stimuli

`inc()`, `read()`, `reset()`, `show()`

Counter - emissions

`print()` to the object `System.out`

Back to our running example



All messages/events are asynchronous! Either generated by the CPU or by the sonar hw or by the communication hardware.

Reactive Objects

Object Oriented Programming?

- Objects have local state
- Objects export methods
- Objects communicate by sending messages
- Objects rest between method invocation

Examples of intuitive objects

People, cars, molecules, ...

Bonus

Principles and methodologies from OOP become applicable to embedded, event-driven and concurrent systems!

Java? C++?

The Counter example again

```
class Counter{
    int x;
    Counter(){x=0;}
    void inc(){x++;}
    int read(){return x;}
    void reset(){x=0;}
}
```

One thread

```
public static void main(){
  Counter c = new Counter();
  c.inc();
  System.out.println(c.read());
}
```

Creating a new object just creates a passive piece of storage! Not a thread of control!

Other threads that use the same counter are sharing the state!

OO and Concurrency

OO Languages:

- An object is a passive piece of global state
- A method is a function
- Sending a message is calling a function

Our model says

- An object is an independent process with local state
- A method is a process fragment
- Sending a message is interprocess communication

This is one of the reasons why we choose to build our own kernel supporting reactive objects and programming in C.

Reactive objects in C

We will need to provide ways for

- Create reactive objects
- Declare protected local state
- Receive messages
  - synchronously
  - asynchronously
- Bridge the hardware/software divide (run ISRs)
- Schedule a system of reactive software objects.

This will be the contents of a kernel called TinyTimber that we will learn how to design and use!

Hardware objects

Black boxes that do nothing unless stimulated by external events.

Class
The kind or type or model of a circuit.

Instance
A particular circuit on a particular board.

State
Internal register status or logic status of an object instance.

Provided interface
The set of pins on a circuit that recognize signals.

Required interface
The set of pins on a circuit that generate signals.

Method call
To raise an input signal and wait for a response (synchronous) or just continue (asynchronous).

Software Objects

Black boxes that do nothing unless stimulated by external events.

Class
Program behaviour expressed as state variable layout and method code.

Instance
A record of state variables at at a particular address (the object's identity).

State
Current state variable contents of a particular object.

Provided interface
The set of methods a class exports.

Required interface
Method calls issued to other objects.

Method call
Call to a function with the designated object address as the first argument.

Encoding state layout

We will use a little kernel called `TinyTimber`. We will use files as modules in C.

In `MyClass.h`

```
#include "TinyTimber.h"

typedef struct{
    Object super;
    int x;
    char y;
} MyClass;

#define initMyClass(z) \
  { initObject ,0,z}
```

- Mandatory! Specified and used by the kernel!
- Unconstrained!
- `initMyClass` corresponds to a constructor, it includes programmer defined intialization.

Using it

```
#include "MyClass.h"
MyClass a = initMyClass(13);
```

Comparing with Java

```
class MyClass{
  int x;
  char y;
  MyClass(int z){
     x=0;
     y=z;
  }
}
```

In our programs we do not allocate objects in the heap (as Java does!).

Our constructors are just preprocessor macros!

```
MyClass a = new MyClass(13);
```

Encoding methods declarations

In `MyClass.h`

```
typedef struct{
   Object super;
   int x;
   char y;
} MyClass;
...
int myMethod( MyClass *self , int q);
```

In `MyClass.c`

```
int myMethod( MyClass *self , int q){
  self -> x  =  self -> y  + q;
}
```

In Java

```
class MyClass{
  int x;
  char y;
  ...
  int myMethod(int q){
    x=y+q;
  }
}
```

Encoding function calls

In Java

```
...
MyClass a = new MyClass(13);
a.myMethod(44);
```

In our C programs

```
...
MyClass a = initMyClass(13);
myMethod( &a ,44);
```

But, we are doing all this to do something different than just function calls! We want to have the possibility of introducing the distinction between synchronous and asynchronous messages!