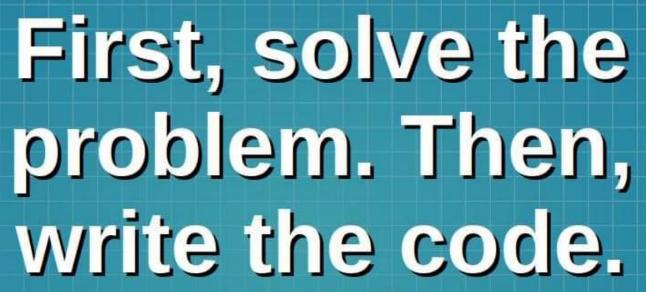
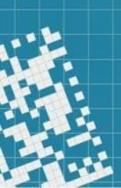
Automata and machines

Paolo Burgio paolo.burgio@unimore.it





- John Johnson





Industrial embedded systems

What they do

- > Monitor physical properties of the system/plant (via sensors)
- > Might perform some control, or part of, control algos
- > Via actuators

Control can be

- > Continuous in time
- > Discrete in time
- → Control theory



Industrial controls in a nutshell



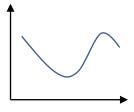
In their generic form,

$$F: \{S, I\} \rightarrow \{O\}$$

computed ...when?

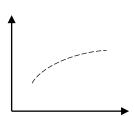
If continuous

- > Physical properties and actuators are continuous in time
- > *F(t)* continuous
- Combinatorial logic/analogic systems



If discrete

- > Computed at pre-determined instance in time
- > Event-driven (e.g., timeout, interrupt)
- > Sequential logics/digital systems





Finite state automations for discrete controls

E.g., an elevator, reacts to multiple events

- > Typically in idle state
- > If you are <u>press</u> the button, the door opens
- > You select the floor, doors close
- > Then, it <u>reaches</u> the floor (feat. velocity control)
- > Then, it opens the door, which subsequently closes <u>after X seconds</u>

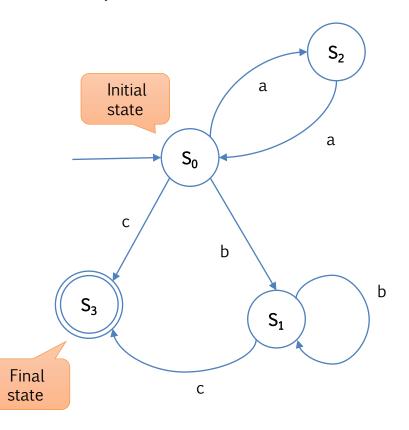
This behavior is controlled by a **finite state automations/machine**



Finite State Automations/Machines

Problem

> Identify even sequences of *a* (even empty), followed by one, or more, or no, *b*, ended by *c*



Given an <u>alphabet</u> V,

...that identifies a language (we'll see)..

define FSA as

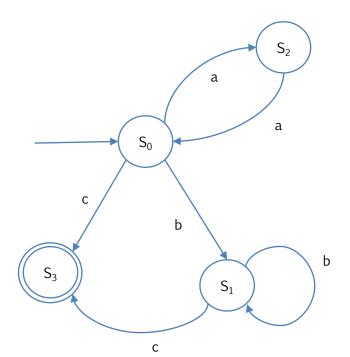
- > S: a non-empty states set
- $\Rightarrow s_0 \in S$: initial state
- $\rightarrow S_f \subseteq S$: final states set
- \rightarrow t: $S \times V \rightarrow S$: states transaction func



FSMs and languages

Let $V^* = \{v, w, ...\}$ contain all the combinations of words using V symbols

- > Including the empty word arepsilon
- > For instance, ac, aabbc, abbabbbc belong to V*
- > (note that, we can associate words in be to inputs, or combination of them)



A language L is a subset of V*

(abbabbbc does not belong to L, as previously defined)

"Identify even sequences of a (even empty), followed by one, or more, or no, b, ended by c"



State transaction function

$$\Rightarrow t(s_0, b) = s_1 \mid s_0 \xrightarrow{b} s_1$$

 $t: S \times V \xrightarrow{b} S$

- > s_v is <u>reachable</u> by s_x if there exists a path from s_x to s_v
 - a combination of alpabet symbols I (letters in our case)

$$\rightarrow$$
: $S \times V^ \times S$: $S_X \xrightarrow{w^*} S_Y$

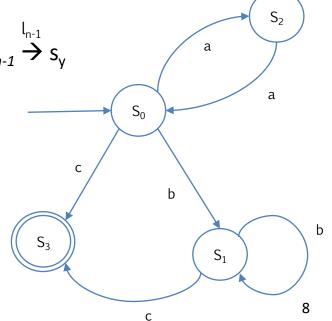
iff

$$w = I_1 I_2 ... I_n$$

$$W = I_1 I_2 ... I_n$$
 $\exists s_1, s_2, ..., s_n : s_x \xrightarrow{l_1} s_1 \xrightarrow{l_2} s_2 ... s_{n-1} \xrightarrow{l_{n-1}} s_y$

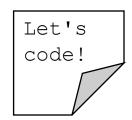


w = aaab
$$\exists s_1, s_2, ..., s_n : s_2 \xrightarrow{a} s_0 \xrightarrow{a} s_2 \xrightarrow{a} s_0 \xrightarrow{b} s_1$$





Exercise



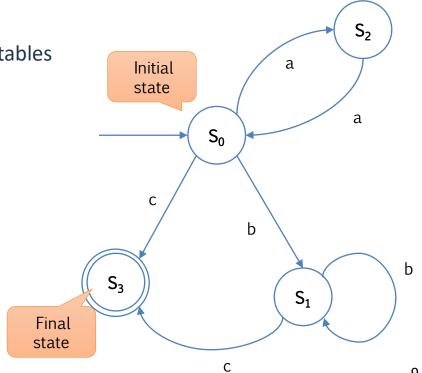
Implement the automata that understands whether a words is from L "Identify even sequences of a (even empty),

followed by one, or more, or no, b, ended by c"

Use the language that you want

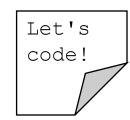
- You just need IFs, CASE-SWITCH, recursion, tables
- Receive the target word from stdin
- Hint: start simple...

What's missing?





Exercise



Implement the automata that understands whether a words is from L

"Identify even sequences of a (even empty),

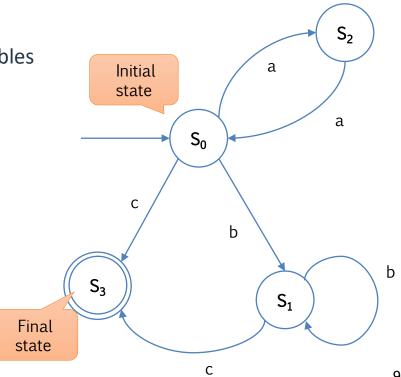
followed by one, or more, or no, b, ended by c"

Use the language that you want

- You just need IFs, CASE-SWITCH, recursion, tables
- > Receive the target word from stdin
- > Hint: start simple...

What's missing?

- > In case of error => default error state
- > Typically implicit in state diagrams





Grammars

> A standard way of representing languages (Noam Chomsky, 1950)

$$G = \langle VT, VN, P, S \rangle$$

> VT : terminal symbols ⊆ V

> VN : non-terminal symbols ⊆ V (aka: syntax categories)

> P : production rules $P \subseteq VN \times (VN \cup VT)$

> S ∈ VN: initial symbol

VT and VN disjoint $VT \cap VN = \emptyset$

VT and VN are VVT U VN = V

A language L_G generated by grammar G is the set of V* elements derived by start symbol S through productions in P



Backus-Naur Form

> Productions rules have form

$$\alpha := \beta$$
, $\alpha \in VN \beta \in V$

- $x \in VN$ have the form <name>
- > | specifies an option



Another example

> Natural numbers

> **Challenge**: extend it with sign (+, -)!



Another example: solution

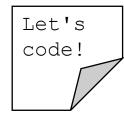
> Natural numbers

> Challenge: extend it with sign (+, -)!



In reality...

> We only give production rules: VN, VT, S are implicitly defined...



Want to try?

- Implement a machine that recognizes whether a sentence (aka: a word of the Language L) is legal for that language
- > ("our" words are symbols of L)



Chomsky classification

> 4 types of grammars, with increasing constraints on production rules structures

Type 0

- > No restriction on productions
- > Phrases can even become shorter!



Type 1 grammars/languages

- > Context-sensitive
- > Production must be in the form

$$x \land y \rightarrow x \alpha y$$

where

 $x, y, \alpha \in (VT \cup VN)^*, \land \in VN, \alpha != \varepsilon$

- \rightarrow A can be replaced with α only if in the context of (surrounded by) x and y
- > Phrases never get shortened
- $\rightarrow \alpha \rightarrow \beta \text{ con } |\beta| \ge |\alpha|$



Type 2 grammars/languages

- > Context-free
- > Production must be in the form

$$A \rightarrow \alpha$$

where

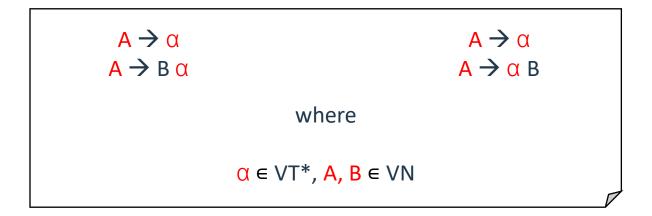
 $\alpha \in (VT \cup VN)^*, A \in VN$

- > α can be ∈
- \rightarrow A can always be replaced with α



Type 3 grammars/languages

- > Regular
- > Production must be in the linear form



- > α can be ∈
- > Either left, or right linear: not both in the same grammar



...and..?

We can build specific machines to recognize/process specific grammar Types

- > Type 0 => Turing machine (if L(G) is recognizable)
- > Type 1 => **Turing machine** with constrained tape length
- > Type 2 => Finite state automations with stack (**Push down automations**)
- > Type 3 => Finite state automations



Hierarchy of machine types

- > Base (combinatorial) machine
- > Finite state machines FSM
- > FSM with stack (PDA)
- > Turing machine

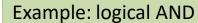


Base combinatorial machine

- > E.g., Logical ports, gates
- > Suitable for continuous control
- Non suitable if you need state/memory
 - Need to model all possible cases!

I: (finite) set of Input symbolsO: (finite) set of output symbols

mfn: I → ○ machine function



$$I = \{ \{0,1\} \times \{0,1\} \}$$

$$0 = \{0,1\}$$

mfn defined by a table

	0	1
0	0	0
1	0	1



Finite state machine

< I, O, S, mfn, sfn >

- > Partly already seen
- > Has memory
- Memory is a limitation

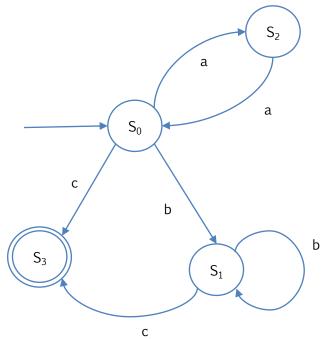
I: (finite) set of Input symbols

: (finite) set of output symbols

S: (finite) set of states

 $mfn: I \times S \rightarrow O \text{ machine function}$

 $sfn: I \times S \rightarrow S$ state function





Finite state machine with stack

< I, O, A, S, mfn, sfn >

I: (finite) set of Input symbols

A: (finite) set of stack alphabet symbols

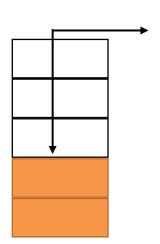
○ : (finite) set of output symbols

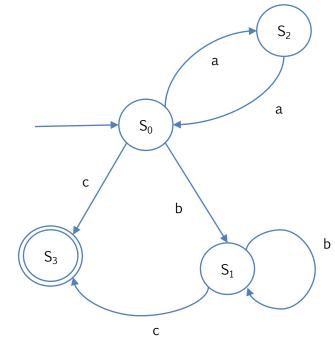
S: (finite) set of states

 $mfn: I \times S \times A \rightarrow O$ machine function

 $sfn: I \times S \times A \rightarrow S$ state function

- > Also known as Push-Down Automata (PDA)
- > Uses a stack
- > We'll see them...







Turing machine

< A, S, mfn, sfn, dfn >

A: (finite) set of in/out symbols

S: (finite) set of states

mfn: A x S → A machine function

sfn: A x S → S state function (inc. HALT)

dfn: A x S → { left, right, none }

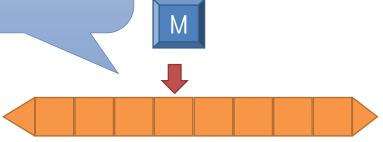
> Unlimited memory

Possible operations:

- · Read from tape
- Write (mfn) to tape
- Change internal status (sfn)
- Move tape in any (dfn) direction

Church-Turing thesis

A function on the natural numbers can be calculated by an effective method if and only if it is computable by a Turing machine



direction function



A Universal Turing Machine

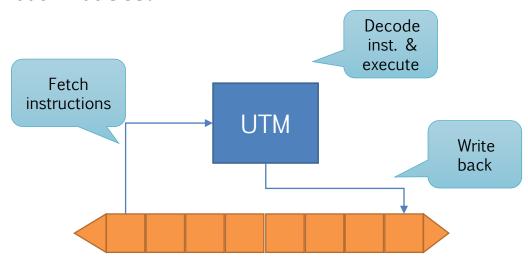
> In TM, the algorithm is inside the machine M, we write results in the tape

What if istruction as well is **in** the tape?

- > We have a programmable machine, with a memory
- >does this remind something?

Which are the catch? What do we miss?

- > Ok, the infinite tape makes it infeasible
- > ..but what else?

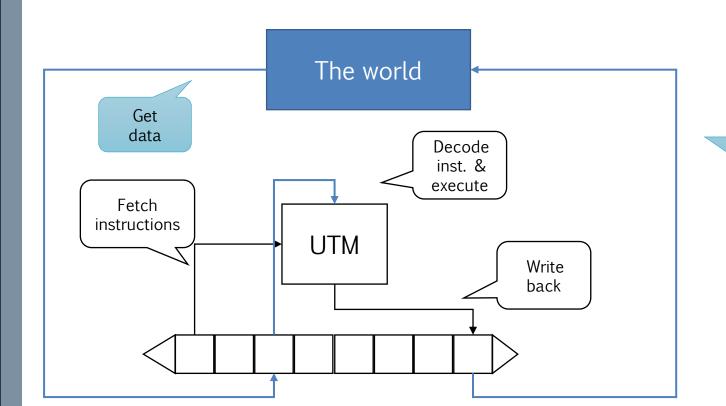




The Van Neumann Machine

We also need to model the interaction with the environment!

- Aka: I/O (HD/SSD is also I/O)
- > Where data comes from!
- > It is a real machine: we can **build** it



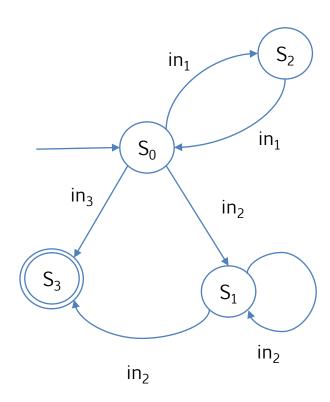
Store data

How to implement a FSM



A generic FSM

- > Till now, we only saw machines that can recognize a word from a language
 - I say "word", you might want to understand "sentence"
- > Let's now see how a machine can actually **produce** an output





The Machine of Mealy

> When crossing an edge, produce an output

< I, O, S, mfn, sfn >

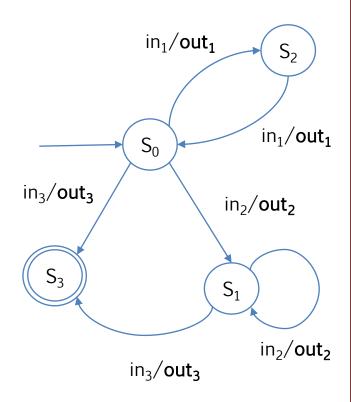
I: (finite) set of Input symbols

○ : (finite) set of output symbols

S: (finite) set of states (s₀ initial state)

 $mfn: I \times S \rightarrow O$ machine/output function

 $sfn: I \times S \rightarrow S$ state transition function





The Machine of Moore

> When in a state an edge, produce an output

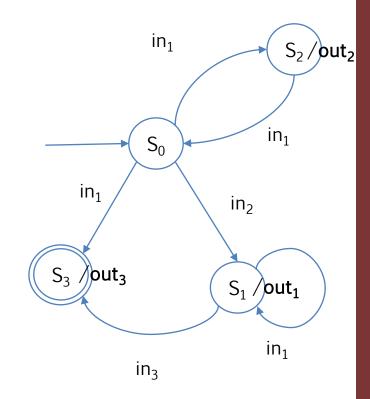
I: (finite) set of Input symbols

○ : (finite) set of output symbols

S: (finite) set of states (s₀ initial state)

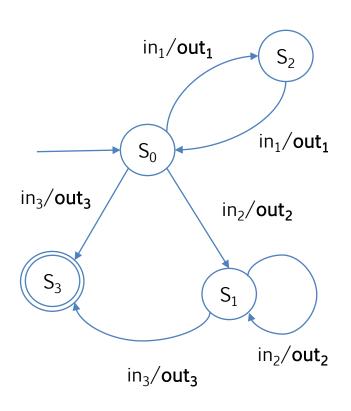
mfn: S → O machine/output function

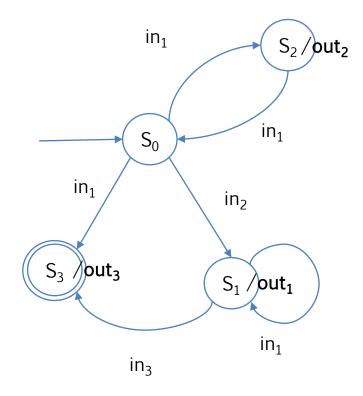
 $sfn: I \times S \rightarrow S$ state transition function





What's the difference?







What's the difference?

Mathematically equivalent

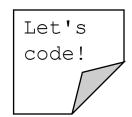
> One can be transformed in another

..but..

- Mealy can potentially have different outs, to different inputs/transitions
 - Less states, if output depends on inputs one can add an edge to the machine
- > Moore potentially keeps the output stable for all the state
 - Moore requires more states, in case out depends on input and not only on state



Exercise

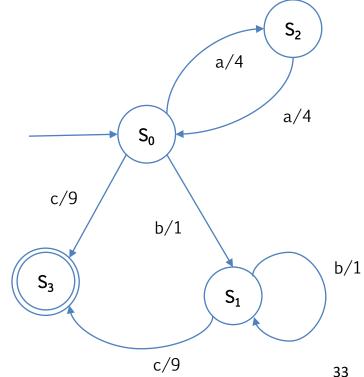


Implement the automata that understands whether a words is from L

"Identify even sequences of a (even empty),

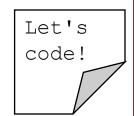
followed by one, or more, or no, b, ended by c"

- > ..and writes the corresponding number (I choose them randomly)
- Mealy? Moore? You choose
 - Here, I show Mealy





Exercise



> Implement the automata that understands whether a words is from L

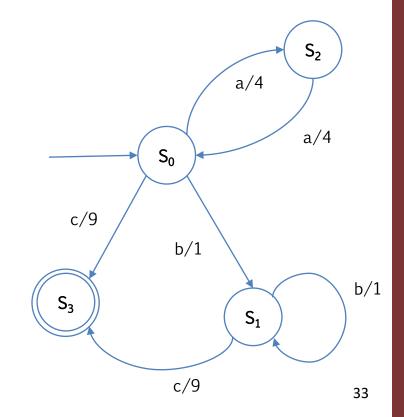
"Identify even sequences of a (even empty),

followed by one, or more, or no, b, ended by c"

- ..and writes the corresponding number (I choose them <u>randomly</u>)
- > Mealy? Moore? You choose
 - Here, I show Mealy

Hint

If not already done, use tables for state/output transactions





What else?

Several tools to support the design

> Matlab Stateflow, UML

Several grammar interpreters to rely the burden of writing FSM code

- > FSF's GNU Bison Included in GCC
- YACC Yet Another Compiler-Compiler



GNU Bison



Converts a context-free grammar into a deterministic LR parser (but not only) in C

- > Recognizes correct sentences from a grammar
- https://www.gnu.org/software/bison/
- > Not part of exam ⁽²⁾

Input format: Bison grammar files

```
% {
    Prologue
% }
Bison declarations
% %
Grammar rules
% %
Epilogue
```



Bison prologue

C-style code that will be appended at the beginning of the generated file

- > Useful for defining macros, includes, headers...
- > ptypes.h contains Bison internal data structures: trees, tokens...

```
응 {
 #define GNU SOURCE
 #include <stdio.h>
 #include "ptypes.h"
응 }
%union {
 long n;
 tree t; /* tree is defined in ptypes.h. */
응 {
  static void print token (yytoken_kind_t token, YYSTYPE val);
응 }
```



Grammar rules

- > Like-BNF syntax
- > Can also include (C) language-specific rules

```
// results => non-terminal;
// components => any
result: components...;
// C statement
{C statements}
// Multiple rules
result:
  rule1-components...
| rule2-components...
// recursive rule
expseq1:
  exp
| expseq1 ',' exp
```



Example - Reverse-polish notation calculator

```
rpcalc.y
      /* empty */
input:
       | input line
     '\n'
line:
       | \exp ' \mid  { printf ("\t%.10g\n", $1); }
                      \{ $$ = $1;
exp:
       NUM
       | \exp \exp '+'  { $$ = $1 + $2;
       | \exp \exp '-'  { $$ = $1 - $2; }
       | \exp \exp '*'  { $$ = $1 * $2;
       | \exp \exp '/'  { $$ = $1 / $2;
     /* Exponentiation */
       | \exp \exp '^{\prime} | \{ \$\$ = pow (\$1, \$2); \}
     /* Unary minus */
       | exp 'n'  { $$ = -$1;
응응
```



Example - Reverse-polish notation calculator

```
"A complete input is either an
                                                            rpcalc.y
                                empty string, or a complete input
         /* empty */
input:
                                   followed by an input line"
         | input line
          '\n'
line:
         | \exp ' \mid  { printf ("\t%.10g\n", $1); }
                             \{ $$ = $1;
           NUM
exp:
                             \{ \$\$ = \$1 + \$2;
         | exp exp '+'
                        \{ \$\$ = \$1 - \$2;
           exp exp '-'
           \exp \exp '*' { $$ = $1 * $2;
         | exp exp '/'
                        \{ \$\$ = \$1 / \$2;
       /* Exponentiation */
         | exp exp '^'
                         \{ \$\$ = pow (\$1, \$2); \}
       /* Unary minus */
                           \{ \$\$ = -\$1;
         | exp 'n'
응응
```



응응

Example - Reverse-polish notation calculator

```
rpcalc.y
input:
             /* empty */
             input line
             '\n'
line:
            exp '\n' { printf ("\t%.10g\n", $1); }
             NUM
exp:
                                       "Can be either a newline, or an expression
             exp exp '+'
                                       followed by a newline"
             exp exp '-'
             exp exp
                                       Also, speficies an action that prints this value
             exp exp '/'
                                       (exp, indicated by $1)
        /* Exponentiation */
                                       Note: we use language-specific features and
           | exp exp '^'
                                       libraries, such as printf (in prologue, I
           Unary minus
                                       included stdio.h)
            exp 'n'
```



Example - Reverse-polish notation calculator

rpcalc.y

Multi-rules expression ("pure" numbers + six arithm operators)

Actions specify how to translate it in C

- \$\$ => result
- \$1, \$2 => operators
- (remember to #include math.h ©)

;

```
exp: NUM { $$ = $1; } } | exp exp '+' { $$ = $1 + $2; } | exp exp '-' { $$ = $1 - $2; } | exp exp '*' { $$ = $1 * $2; } | exp exp '*' { $$ = $1 / $2; } | exp exp '/' { $$ = $1 / $2; } | exp exp '/' { $$ = $1 / $2; } | exp exp '/' { $$ = $1 / $2; } | exp exp '/' { $$ = $0w ($1, $2); } | exp exp '/' { $$ = $0w ($1, $2); } | exp 'n' { $$ = -$1; } | exp 'n' { $$ = -$1
```



Exercise (optional)

Let's code!

Write a parser for the following grammar using Bison

Non-deterministic automata



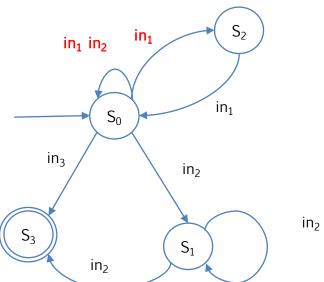
Deterministic vs. non-deterministic

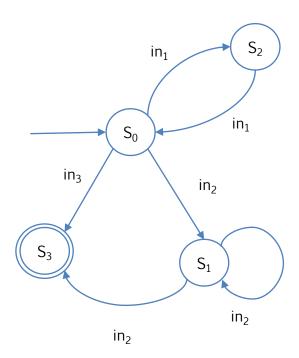
Till now, deterministic automations - **DFA**

- > uniquely identifies a transaction with the couple stateinput
- Only one edge/input connects nodes/states

Non-deterministic finite automations (NFA) can have multiple edges connecting nodes/states

> Also, same inputs can lead to more than one nodes







NFA-to-DFA conversion

- > Non-deterministic automata can be translated into equivalent deterministic one
- > For each NFA, there is a DFA such that it recognizes the same formal language
- > If a DFA cannot recognize a formal language, neither a NFA can

Rabin-Scott powerset construction

- > Catch: if a NFA has n states, the corresponding DFA can have up to 2^n states!
- > We won't see this...

M. O. Rabin and D. Scott, "Finite Automata and Their Decision Problems," in IBM Journal of Research and Development, vol. 3, no. 2, pp. 114-125, April 1959, doi: 10.1147/rd.32.0114.

Event driven Systems



Event driven systems

A system that reacts from external stimula

- > Istantly?
- > Aka: Cyber-Physical Systems (CPS)

Can be

- > Synchronous
- > Asynchronous



Synchronous (Active polling)

> Infinite loop

```
char c;
while (c != SOME_VALUE)
  c = readC();

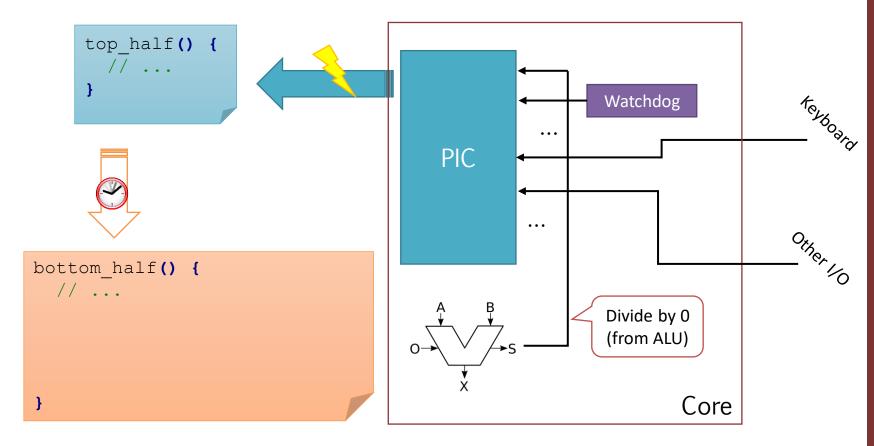
// We can go, now
```

- > **Pros**: extremely fast and reactive
- Cons: waste of resources as one core is busy
 - Possible workaround: insert a sleep



Asynchronous (Interrupt Service Routine)

> Programmable interrupt controller (hierarchy)



- > Pros: "pay-as-you-go"
- Cons: takes more time to issue a ISP



...a mix of the two

Keybaroard management in a General-Purpose system

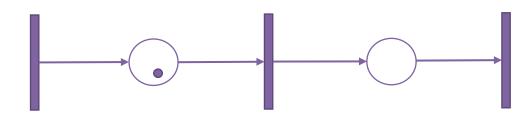
- > GNU/Linux
- > ISP with bottom-half and top-half @ kernel space
- Synchronous, language-specific library API @ user space kernel space user space Core top half() { char c; cin >> c; // Blocking bottom half() { Unlock // Blocked // ... istream &operator>>(istream &, char &); 47 iostream.h

Petri nets



A directed bipartite graph

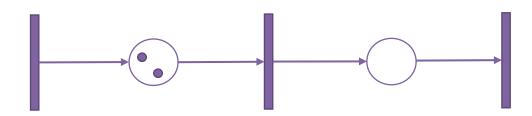
- > Transitions triggered by events (bars)
- > Places, i.e., conditions (circles)
- Arcs connect only places to transitions (or vice-versa), and specify which places are
 pre- or post-conditions for events
- > Every place collects **tokens** (*dots*) which might trigger an event (if multiple events are triggered in the same net, which fires first is non-deterministic





A directed bipartite graph

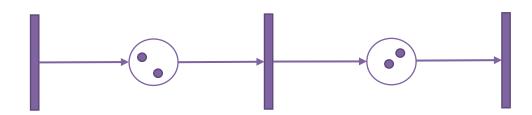
- > Transitions triggered by events (bars)
- > Places, i.e., conditions (circles)
- Arcs connect only places to transitions (or vice-versa), and specify which places are
 pre- or post-conditions for events
- > Every place collects **tokens** (*dots*) which might trigger an event (if multiple events are triggered in the same net, which fires first is non-deterministic





A directed bipartite graph

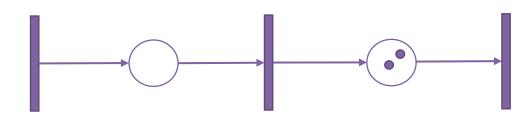
- > Transitions triggered by events (bars)
- > Places, i.e., conditions (circles)
- Arcs connect only places to transitions (or vice-versa), and specify which places are
 pre- or post-conditions for events
- > Every place collects **tokens** (*dots*) which might trigger an event (if multiple events are triggered in the same net, which fires first is non-deterministic





A directed bipartite graph

- > **Transitions** triggered by **events** (*bars*)
- > Places, i.e., conditions (circles)
- Arcs connect only places to transitions (or vice-versa), and specify which places are
 pre- or post-conditions for events
- > Every place collects **tokens** (*dots*) which might trigger an event (if multiple events are triggered in the same net, which fires first is non-deterministic





(Marking) Petri net - formalism

A tuple (S, T, W, M_0)

> S: finite set of states

- > T: finite set of transitions
- $\rightarrow W:(SxT)U(TxS) \rightarrow \mathbb{N}$ multiset of arcs
- > M: (marking) a mapping $S \rightarrow \mathbb{N}$ that assigns to each place a number of tokens
- $\rightarrow M_0$: initial marking

Subject to:

- > S and T are disjoint
- > No arc can connect two states or two transitions among them

How they execute

- > firing a transition t in a marking M consumes W(s, t) tokens from each of its input places, and produces W(t, s) tokens in each of its output places
- > a transition is *enabled* (it may fire) in M if there are enough tokens in its input places for the consumptions to be possible, i.e. if and only if $\forall s : M(s) \ge W(s, t)$



How to run the examples



> Find them in Code/ folder from the course website

For C++: compile

> \$ gcc code.cpp -o code -Wall -lstdc++

Run (Unix/Linux)

\$./code

Run (Win/Cygwin)

\$./code.exe



References



Course website

http://hipert.unimore.it/people/paolob/pub/Industrial Informatics/index.html

My contacts

- > paolo.burgio@unimore.it
- http://hipert.mat.unimore.it/people/paolob/

Resources

- > Alessandro Fantechi, «Informatica Industriale», Città Studi Edizioni
- > For interrupts
 - Robert Love, «Linux kernel development», Pearson
- > For GNU Bison
 - http://dinosaur.compilertools.net/bison/bison_5.html
- A "small blog"
 - http://www.google.com