Initiation à la Vérification
Basics of Verification

Stefan Schwoon
credits for slides: Paul Gastin, Javier Esparza, Keijo Heljanko

M1 MPRI, 2021/2022
## Organisation

### Timetable

<table>
<thead>
<tr>
<th>Event</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Course</strong></td>
<td>Friday 10:45 – 12:45 (Stefan Schwoon)</td>
</tr>
<tr>
<td><strong>Exercises</strong></td>
<td>Friday 8:30 – 10:30 (Nathan Thomasset)</td>
</tr>
<tr>
<td>Midterm exam</td>
<td>(1/3)</td>
</tr>
<tr>
<td>Homework 1</td>
<td>(1/6)</td>
</tr>
<tr>
<td>Final exam</td>
<td>(1/3)</td>
</tr>
<tr>
<td>Second session</td>
<td>Replaces midterm + final exams</td>
</tr>
</tbody>
</table>

**Controls (to be confirmed)**
Organisation

Timetable

- Course: Friday 10:45 – 12:45 (Stefan Schwoon)
- Exercises: Friday 8:30 – 10:30 (Nathan Thomasset)

Controls (to be confirmed)

- Homework 1 (1/6)
- Midterm exam (1/3)
- Homework 2 (1/6)
- Final exam (1/3)

Second session: replaces midterm + final exams
Need for formal verifications methods

Critical systems
- Transport
- Energy
- Medicine
- Communication
- Finance
- Embedded systems
- ...

...
Disastrous software bugs

Mariner 1 probe, 1962

See http://en.wikipedia.org/wiki/Mariner_1

- Destroyed 293 seconds after launch
- Missing hyphen in the data or program? No!
- Overbar missing in the mathematical specification:

\[ \hat{R}_n: \text{nth smoothed value of the time derivative of a radius.} \]

Without the smoothing function indicated by the bar, the program treated normal minor variations of velocity as if they were serious, causing spurious corrections that sent the rocket off course.
Disastrous software bugs

Ariane 5 flight 501, 1996

See http://en.wikipedia.org/wiki/Ariane_5_Flight_501

- Destroyed 37 seconds after launch (cost: 370 millions dollars).
- Data conversion from a 64-bit floating point to 16-bit signed integer value caused a hardware exception (arithmetic overflow).
- Efficiency considerations had led to the disabling of the software handler (in Ada code) for this error trap.
- The fault occurred in the inertial reference system of Ariane 5. The software from Ariane 4 was re-used for Ariane 5 without re-testing.
- On the basis of those calculations the main computer commanded the booster nozzles, and somewhat later the main engine nozzle also, to make a large correction for an attitude deviation that had not occurred.
- The error occurred in a realignment function which was not useful for Ariane 5.
Disastrous software bugs

Spirit Rover (Mars Exploration), 2004


- Ceased communicating on January 21.
- Flash memory management anomaly: too many files on the file system
- Resumed to working condition on February 6.
Other well-known bugs


Formal verifications methods

Complementary approaches

- Theorem prover
- Model checking
- Static analysis
- Test
What does “Model-Checking” mean?
What does “Model-Checking” mean?
Model Checking

- Purpose 1: automatically finding software or hardware bugs.
- Purpose 2: prove correctness of abstract models.
- Should be applied during design.
- Real systems can be analysed with abstractions.

E.M. Clarke
E.A. Emerson
J. Sifakis

Prix Turing 2007.
Model Checking

3 steps

- Constructing the model $M$ (transition systems)
- Formalizing the specification $\varphi$ (temporal logics)
- Checking whether $M \models \varphi$ (algorithmics)

Main difficulties

- Size of models (combinatorial explosion)
- Expressivity of models or logics
- Decidability and complexity of the model-checking problem
- Efficiency of tools

Challenges

- Extend models and algorithms to cope with more systems. Infinite systems, parameterized systems, probabilistic systems, concurrent systems, timed systems, hybrid systems, . . .
- Scale current tools to cope with real-size systems. Needs for modularity, abstractions, symmetries, . . .
References

Bibliography

*Principles of Model Checking*.  

*Systems and Software Verification. Model-Checking Techniques and Tools*.  

*Model Checking*.  


*Temporal Verification of Reactive Systems: Safety*.  
Outline

Introduction

2 Models
- Transition systems
- ... with variables
- Concurrent systems
- Synchronization and communication

Specifications

Linear Time Specifications

Branching Time Specifications
Constructing the model

Example: Men, Wolf, Goat, Cabbage

Model = Transition system

- **State** = who is on which side of the river
- **Transition** = crossing the river
- **Specification**
  - Safety: Never leave WG or GC alone
  - Liveness: Take everyone to the other side of the river.
Definition: TS

\[ M = (S, \Sigma, T, I, \text{AP}, \ell) \]

- \( S \): set of states (finite or infinite)
- \( \Sigma \): set of actions
- \( T \subseteq S \times \Sigma \times S \): set of transitions
- \( I \subseteq S \): set of initial states
- \( \text{AP} \): set of atomic propositions
- \( \ell : S \rightarrow 2^{\text{AP}} \): labelling function.

Example: Digicode ABA

Every discrete system may be described with a TS.
**Peterson’s algorithm (1981)**

Process $i$:
- Loop forever
  - $req[i] := true; turn := 1-i$
  - wait until ($turn = i$ or $req[1-i] = false$)
- Critical section
  - $req[i] := false$

---

Exercise:
- Draw the concrete TS assuming the first two assignments are atomic.
- Is the algorithm still correct if we swap the first two assignments?
Description Languages

Pb: How can we easily describe big systems?

Description Languages (high level)

- Programming languages
- Boolean circuits
- Modular description, e.g., parallel compositions
  problems: concurrency, synchronization, communication, atomicity, fairness, ...
- Petri nets (intermediate level)
- Transition systems (intermediate level)
  with variables, stacks, channels, ...
  synchronized products
- Logical formulae (low level)

Operational semantics

High level descriptions are translated (compiled) to low level (infinite) TS.
## Transition systems with variables

**Definition:** TSV $M = (S, \Sigma, V, (D_v)_{v \in V}, T, I, AP, \ell)$

- $V$: set of (typed) variables, e.g., boolean, [0..4], ...
- Each variable $v \in V$ has a domain $D_v$ (finite or infinite)
- Guard or Condition: unary predicate over $D = \prod_{v \in V} D_v$
  - Symbolic descriptions: $x < 5$, $x + y = 10$, ...
- Instruction or Update: map $f : D \rightarrow D$
  - Symbolic descriptions: $x := 0$, $x := (y + 1)^2$, ...
- $T \subseteq S \times (2^D \times \Sigma \times D^D) \times S$
  - Symbolic descriptions: $s \xrightarrow{x<50,?coin,x:=x+coin} s'$
- $I \subseteq S \times 2^D$
  - Symbolic descriptions: $(s_0, x = 0)$

### Example: Vending machine

- coffee: 50 cents, orange juice: 1 euro, ...
- possible coins: 10, 20, 50 cents
- we may shuffle coin insertions and drink selection
Transition systems with variables

Semantics: low level TS

- $S' = S \times D$
- $I' = \{(s, \nu) \mid \exists (s, g) \in I \text{ with } \nu \models g\}$
- Transitions: $T' \subseteq (S \times D) \times \Sigma \times (S \times D)$

\[
\frac{s \xrightarrow{g, a, f} s' \wedge \nu \models g}{(s, \nu) \xrightarrow{a} (s', f(\nu))}
\]

SOS: Structural Operational Semantics

- $AP'$: we may use atomic propositions in $AP$ or guards in $2^D$ such as $x > 0$.

Programs = Kripke structures with variables

- Program counter = states
- Instructions = transitions
- Variables = variables

Example: GCD
Example: Digicode

\[\begin{align*}
&\text{OPEN} \\
&\text{A B A} \\
&\text{cpt} < n \\
&\text{B, C} \\
&\text{cpt}++ \\
&\text{B, C} \\
&\text{cpt}++ \\
&\text{A} \\
&\text{cpt}++ \\
&\text{C} \\
&\text{cpt}++ \\
&\text{B, C} \\
&\text{cpt}++ \\
&\text{A} \\
&\text{cpt} = n \\
&\text{B, C} \\
&\text{cpt}++ \\
&\text{C} \\
&\text{cpt}++ \\
&\text{B, C} \\
&\text{cpt}++ \\
&\text{A, C} \\
&\text{cpt}++ \\
&\text{C} \\
&\text{cpt} ++ \\
&\text{B, C} \\
&\text{cpt}++ \\
&\text{A} \\
&\text{cpt} = n \\
&\text{B, C} \\
&\text{cpt}++ \\
&\text{A} \\
&\text{cpt} = n \\
&\text{B, C} \\
&\text{cpt}++ \\
\end{align*}\]
... and its semantics \((n = 2)\)

Example: Digicode

\[
\begin{align*}
1,0 & \xrightarrow{A} 2,0 & \xrightarrow{B} 3,0 & \xrightarrow{A} 4,0 \\
1,1 & \xrightarrow{A} 2,1 & \xrightarrow{B} 3,1 & \xrightarrow{A} 4,1 & \text{OPEN} \\
1,2 & \xrightarrow{A} 2,2 & \xrightarrow{B} 3,2 & \xrightarrow{A} 4,2 & \text{OPEN} \\
5,3 & \xrightarrow{B, C} 2,0 & \xrightarrow{B, C} 3,0 & \xrightarrow{B, C} 4,0 & \text{ERROR}
\end{align*}
\]
Modular description of concurrent systems

\[ M = M_1 \parallel M_2 \parallel \cdots \parallel M_n \]

Semantics

- Various semantics for the parallel composition \( \parallel \)
- Various communication mechanisms between components:
  - Shared variables, FIFO channels, Rendez-vous, ...
- Various synchronization mechanisms

Example: Elevator with 1 cabin, 3 doors, 3 calling devices
Example: Elevator

- Cabin:
  - 0 -> 1 -> 2

- Door for level $i$:
  - Closed -> Opened

- Call for level $i$:
  - False -> True

The actual system is a synchronized product of all these automata. It consists of (at most) $3 \times 2^3 \times 2^3 = 192$ states.
Synchronized products

**Definition: General product**

- **Components:** \( M_i = (S_i, \Sigma_i, T_i, I_i, AP_i, \ell_i) \)
- **Product:** \( M = (S, \Sigma, T, I, AP, \ell) \) with
  \[
  S = \prod_i S_i, \quad \Sigma = \prod_i (\Sigma_i \cup \{\varepsilon\}), \quad \text{and} \quad I = \prod_i I_i
  \]
  \[
  T = \{(p_1, \ldots, p_n) \xrightarrow{(a_1, \ldots, a_n)} (q_1, \ldots, q_n) \mid \text{for all } i, (p_i, a_i, q_i) \in T_i \text{ or } p_i = q_i \text{ and } a_i = \varepsilon\}
  \]
  \[
  AP = \biguplus_i AP_i \quad \text{and} \quad \ell(p_1, \ldots, p_n) = \bigcup_i \ell(p_i)
  \]

**Synchronized products: restrictions of the general product.**

**Parallel compositions**

- **Synchronous:** \( \Sigma_{sync} = \prod_i \Sigma_i \)
- **Asynchronous:** \( \Sigma_{sync} = \biguplus_i \Sigma_i' \) with \( \Sigma_i' = \{\varepsilon\}^{i-1} \times \Sigma_i \times \{\varepsilon\}^{n-i} \)

**Synchronizations**

- **By states:** \( S_{sync} \subseteq S \)
- **By labels:** \( \Sigma_{sync} \subseteq \Sigma \)
- **By transitions:** \( T_{sync} \subseteq T \)
Example: Asynchronous product
Synchronization by states: \((P, P)\) is forbidden
Example: **Synchronous product**

Synchronization by transitions

---

**Example: digicode**
Synchronization by Rendez-vous

Synchronization by transitions is universal but too low-level.

**Definition: Rendez-vous**

- \(!m\) sending message \(m\)
- \(?m\) receiving message \(m\)
- **SOS: Structural Operational Semantics**

Local actions:

\[
\begin{align*}
& s_1 \xrightarrow{a_1} s'_1 \\
& (s_1, s_2) \xrightarrow{a_1} (s'_1, s_2)
\end{align*}
\]

Rendez-vous:

\[
\begin{align*}
& s_1 \xrightarrow{!m} s'_1 \land s_2 \xrightarrow{?m} s'_2 \\
& (s_1, s_2) \xrightarrow{m} (s'_1, s'_2)
\end{align*}
\]

- It is a kind of synchronization by actions.
- Essential feature of process algebra.

**Example: Elevator with 1 cabin, 3 doors, 3 calling devices**

- \(?\text{up}\) is uncontrollable for the cabin
- \(?\text{leave}_i\) is uncontrollable for door \(i\)
- \(?\text{call}_0\) is uncontrollable for the system
We should design the controller
Shared variables

Definition: Asynchronous product + shared variables

\( \bar{s} = (s_1, \ldots, s_n) \) denotes a tuple of states
\( \nu \in D = \prod_{v \in \mathcal{V}} D_v \) is a valuation of variables.

Semantics (SOS)

\[
\nu \models g \wedge s_i \xrightarrow{g, a, f} s'_i \wedge s'_j = s_j \text{ for } j \neq i
\]

\[
(\bar{s}, \nu) \xrightarrow{a} (\bar{s}', f(\nu))
\]

Example: Mutual exclusion for 2 processes satisfying

- **Safety**: never simultaneously in critical section (CS).
- **Liveness**: if a process wants to enter its CS, it eventually does.
- **Fairness**: if process 1 wants to enter its CS, then process 2 will enter its CS at most once before process 1 does.

using shared variables but no synchronization mechanisms: the atomicity is

- testing or reading or writing a single variable at a time
- no test-and-set: \( \{x = 0; x := 1\} \)