Formal Specification and Verification

Formal specification (2) 6.12.2016

Viorica Sofronie-Stokkermans

e-mail: sofronie@uni-koblenz.de

Until now

- Logic
- Formal specification (generalities)

Algebraic specification

Transition systems

Transition systems

Transition systems

- Executions
- Modeling data-dependent systems

Transition systems

- Model to describe the behaviour of systems
- Digraphs where nodes represent states, and edges model transitions
- State: Examples
 - the current colour of a traffic light
 - the current values of all program variables + the program counter
 - the current value of the registers together with the values of the input bits
- **Transition** ("state change"): Examples
 - a switch from one colour to another
 - the execution of a program statement
 - the change of the registers and output bits for a new input

Transition systems

Definition.

- A transition system TS is a tuple $(S, Act, \rightarrow, I, AP, L)$ where:
 - *S* is a set of states
 - Act is a set of actions
 - $\rightarrow \subseteq S \times Act \times S$ is a transition relation
 - $I \subseteq S$ is a set of initial states
 - AP is a set of atomic propositions
 - $L: S \rightarrow 2^{AP}$ is a labeling function

S and Act are either finite or countably infinite Notation: $s \xrightarrow{\alpha} s'$ instead of $(s, \alpha, s') \in \rightarrow$.

Direct successors and predecessors

$$Post(s, \alpha) = \{s' \in S \mid s \xrightarrow{\alpha} s'\}, \qquad Post(s) = \bigcup_{\alpha \in Act} Post(s, \alpha)$$

$$Pre(s, \alpha) = \{s' \in S \mid s' \xrightarrow{\alpha} s\}, \qquad Pre(s) = \bigcup_{\alpha \in Act} Pre(s, \alpha)$$

$$Post(C, \alpha) = \bigcup_{s \in C} Post(s, \alpha),$$

$$Post(C) = \bigcup_{\alpha \in Act} Post(C, \alpha) \quad \text{for } C \subseteq S$$

$$Pre(C, \alpha) = \bigcup_{s \in C} Pre(s, \alpha),$$

$$Pre(C) = \bigcup_{\alpha \in Act} Pre(C, \alpha) \quad \text{for } C \subseteq S$$

State s is called terminal if and only if $Post(s) = \emptyset$

Non-determinism

Nondeterminism is a feature!

- to model concurrency by interleaving
 - no assumption about the relative speed of processes
- to model implementation freedom
 - only describes what a system should do, not how
- to model under-specified systems, or abstractions of real systems
 - use incomplete information

Non-determinism

Nondeterminism is a feature!

- to model concurrency by interleaving
 - no assumption about the relative speed of processes
- to model implementation freedom
 - only describes what a system should do, not how
- to model under-specified systems, or abstractions of real systems
 - use incomplete information

In automata theory, nondeterminism may be exponentially more succinct but that's not the issue here! **Definition.** State $s \in S$ is called reachable in *TS* if there exists an initial, finite execution fragment

$$s_0 \stackrel{\alpha_1}{\rightarrow} s_1 \stackrel{\alpha_2}{\rightarrow} \cdots \stackrel{\alpha_n}{\rightarrow} s_n = s$$

Reach(TS) denotes the set of all reachable states in TS.

Detailed description of states

Variables; Predicates

Beverage vending machine revisited



Program graph representation

Program graph representation

Some preliminaries

- typed variables with a valuation that assigns values in a fixed structure to variables
 - e.g., $\beta(x) = 17$ and $\beta(y) = -2$
- Boolean conditions: set of formulae over Var
 - propositional logic formulas whose propositions are of the form " $x \in D$ "

-
$$(-3 < x \le 5) \land (y = green) \land (x \le 2 * x')$$

• effect of the actions is formalized by means of a mapping:

$$Effect : Act \times Eval(Var) \rightarrow Eval(Var)$$

- e.g., $\alpha \equiv x := y + 5$ and evaluation $\beta(x) = 17$ and $\beta(y) = -2$

- Effect $(\alpha, \beta)(x) = \beta(y) + 5 = 3$,
- Effect $(\alpha, \beta)(y) = \beta(y) = -2$

Program graph representation

Program graphs

A program graph PG over set Var of typed variables is a tuple

(Loc, Act, Effect,
$$\rightarrow$$
, Loc₀, g_0)

where

- Loc is a set of locations with initial locations $Loc_0 \subseteq Loc$
- Act is a set of actions
- Effect : $Act \times Eval(Var) \rightarrow Eval(Var)$ is the effect function
- $\rightarrow \subseteq Loc \times ($ Cond(Var) $\times Act) \times Loc$, transition relation

Boolean conditions on Var

• $g_0 \in Cond(Var)$ is the initial condition.

Notation: $I \xrightarrow{g:\alpha} I'$ denotes $(I, g, \alpha, I') \in \rightarrow$.

Beverage Vending Machine

- $Loc = {start, select}$ with $Loc_0 = {start}$
- Act = {bget, sget, coin, ret-coin, refill}
- Var = {nsprite, nbeer} with domain {0, 1, ..., max}
- Effect : Act × Eval(Var) → Eval(Var) defined as follows:

• $g_0 = (nsprite = max \land nbeer = max)$

From program graphs to transition systems

- Basic strategy: unfolding
 - state = location (current control) I + data valuation β
- (I, β)
- initial state = initial location + data valuation satisfying

the initial condition g_0

- Propositions and labeling
 - propositions: "at I" and " $x \in D$ " for $D \subseteq dom(x)$
 - < I, β > is labeled with "at I" and all conditions that hold in β .
- $I \xrightarrow{g:\alpha} I'$ and g holds in β then $< I, \beta > \xrightarrow{\alpha} < I'$, $Effect(< I, \beta >) >$

Transition systems for program graphs

The transition system TS(PG) of program graph

$$PG = (Loc, Act, Effect, \rightarrow, Loc_0, g_0)$$

over set *Var* of variables is the tuple $(S, Act, \rightarrow, I, AP, L)$ where:

- $S = Loc \times Eval(Var)$
- $\rightarrow S \times Act \times S$ is defined by the rule: If $I \stackrel{g:\alpha}{\rightarrow} I'$ and $\beta \models g$ then $\langle I, \beta \rangle \stackrel{\alpha}{\rightarrow} \langle I', Effect(\langle I, \beta \rangle) \rangle$
- $I = \{ < I, \beta > | I \in Loc_0, \beta \models g_0 \}$
- $AP = Loc \cup Cond(Var)$ and
- $L(\langle I, \beta \rangle) = \{I\} \cup \{g \in Cond(Var) \mid \beta \models g\}.$

Transition systems for program graphs



Generalizations of transition systems

- More detailed description of states: Abstract state machines
- Emphasis on processes and their interdependency: CSP
- Durations: Timed automata
- Continuous evolution + discrete control: Hybrid automata

Abstract state machines (ASM)

Purpose

Formalism for modelling/formalising (sequential) algorithms Not: Computability / complexity analysis

Invented/developed by

Yuri Gurevich, 1988

Old name

Evolving algebras

ASMs

Three Postulates

Sequential Time Postulate:

An algorithm can be described by defining a set of states, a subset of initial states, and a state transformation function

Abstract State Postulate:

States can be described as first-order structures

Bounded Exploration Postulate:

An algorithm explores only finitely many elements in a state to decide what the next state is. There is a finite number of names (terms) for all these "interesting" elements in all states.

Example: Computing Squares

Initial State

square = 0

count = 0

ASM for computing the square of input

```
if input < 0 then
    input := - input
else if input > 0∧ count < input then
    par
        square := square + input
        count := count +1
    endpar</pre>
```

The Sequential Time Postulate

Sequential algorithm

An algorithm is associated with

- a set S of states
- a set $I \subseteq S$ of initial states
- A function *τ* : *S* → *S* (the one-step transformation of the algorithm)

Run (computation)

A run (computation) is a sequence $X_0, X_1, X_2 \dots$ of states such that

- $X_0 \in I$
- $au(X_i) = X_{i+1}$ for all $i \ge 0$

Remark

Remark: In this formalism, algorithms are deterministic

au:S o S can be also viewed as a relation $R\subseteq S imes\{ au\} imes S$ with $(s, au,s')\in R$ iff au(s)=s'.

The Abstract State Postulate

States are first-order structures where

- all states have the same vocabulary (signature)
- the transformation τ does not change the base set (universe)
- *S* and *I* are closed under isomorphism
- if f is an isomorphism from a state X onto a state Y, then f is also an isomorphism from $\tau(X)$ onto $\tau(Y)$.

Example: Trees

Vocabulary

nodes:	unary, boolean:	the class of nodes
		(type/universe)
strings:	unary, boolean:	the class of strings
parent:	unary:	the parent node
firstChild:	unary:	the first child node
nextSibling:	unary:	the first sibling
label:	unary:	node label
С:	constant:	the current node

Signatures: A signature is a finite set of function symbols, where

- each symbol is assigned an arity $n \ge 0$
- symbols can be marked relational (predicates)
- symbols can be marked static (default: dynamic)

Signatures: A signature is a finite set of function symbols, where

- each symbol is assigned an arity $n \ge 0$
- symbols can be marked relational (predicates)
- symbols can be marked static (default: dynamic)

Remark: This is not a restriction

• predicates with arity n can be regarded as functions with arity $s \dots s \rightarrow bool$

where s is the usual sort (for terms) and bool is a different sort

- The sort bool is described using a unary predicate Bool
- The sort Bool contains all formulae, in particular also \top , \perp ("relational constants")

Signatures: A signature is a finite set of function symbols, where

- each symbol is assigned an arity $n \ge 0$
- symbols can be marked relational (predicates)
- symbols can be marked static (default: dynamic)

Each signature contains

- the constant *undef* ("undefined")
- the relational constants op (true), op (false)
- the unary relational symbols Boole, \neg
- the binary relational symbols $\land,\lor,\rightarrow,\leftrightarrow,\approx$

These special symbols are all static

Signatures: A signature is a finite set of function/predicate symbols, where

- each symbol is assigned an arity $n \ge 0$
- symbols can be marked static (default: dynamic)

Each signature contains

- the constant *undef* ("undefined")
- the relational constants true, false
- the unary relational symbols Boole, \neg
- the binary relational symbols $\land, \lor, \rightarrow, \leftrightarrow, \approx$

These special symbols are all static

There is an infinite set of variables Terms are built as usual from variables and function symbols Formulae are built as usual

First-order Structures (States)

First-order structures (states) consist of

- a non-empty universe (called BaseSet)
- an interpretation of the symbols in the signature

Restrictions on states

- 0, 1, $undef \in BaseSet$ (different)
- $\perp_{\mathcal{A}} = 0$, $\top_{\mathcal{A}} = 1$
- $undef_{\mathcal{A}} = undef$
- If f relational then $f_{\mathcal{A}}$: BaseSet $\rightarrow \{0, 1\}$
- $Boole_{\mathcal{A}} = \{0, 1\}$
- \neg , \lor , \land , \rightarrow , \leftrightarrow are interpreted as usual

The reserve of a state

Reserve: Consists of the elements that are "unknown" in a state

The reserve of a state must be infinite

Extended States

Variable assignment

A function $\beta: Var \rightarrow BaseSet$

(boolean variables are assigned 0 or 1)

Extended state

A pair (\mathcal{A}, β) consisting of a state \mathcal{A} and a variable assignment β .

Extended States

Variable assignment

A function $\beta: Var \rightarrow BaseSet$

(boolean variables are assigned 0 or 1)

Extended state

A pair (\mathcal{A}, β) consisting of a state \mathcal{A} and a variable assignment β .

Evaluation of terms and formulae: as usual

Example: Trees

Vocabulary

nodes:	unary, boolean:	the class of nodes
		(type/universe)
strings:	unary, boolean:	the class of strings
parent:	unary:	the parent node
firstChild:	unary:	the first child node
nextSibling:	unary:	the first sibling
label:	unary:	node label
С:	constant:	the current node

Example: Trees

Terms

parent(parent(c)) label(firstChild(c)) parent(firstChild(c)) = c $nodes(x) \rightarrow parent(x) = parent(nextSibling(x))$ (x is a variable)

(Boolean, formula)

Isomorphism

Lemma (Isomorphism)

Isomorphic states (structures) are indistinguishable by ground terms:

Justification for postulate

Algorithm must have the same behaviour for indistinguishable states

Isomorphic states are different representations of the same abstract state!

Locations. A location is a pair (f, \overline{a}) with

- f an *n*-ary function symbol
- $\overline{a} \in \mathsf{BaseSet}^n$ an *n*-tuple

Examples

(parent, a), (firstChild, a), (nextSibling, a), (c,)

Locations. A location is a pair (f, \overline{a}) with

- f an *n*-ary function symbol
- $\overline{a} \in \mathsf{BaseSet}^n$ an *n*-tuple

Examples

(parent, a), (firstChild, a), (nextSibling, a), (c,)

An update is a triple (f, \overline{a}, b) with

- (f, \overline{a}) a location
- f not static
- $b \in \mathsf{BaseSet}$
- if f is relational, then $b \in \{0, 1\}$

Locations. A location is a pair (f, \overline{a}) with

- f an *n*-ary function symbol
- $\overline{a} \in \mathsf{BaseSet}^n$ an *n*-tuple

Examples

(parent, a), (firstChild, a), (nextSibling, a), (c,)

An update is a triple (f, \overline{a}, b) with

- (f, \overline{a}) a location
- f not static

• $b \in \mathsf{BaseSet}$

• if f is relational, then $b \in \{0, 1\}$

Intended meaning:

f is changed by changing $f(\overline{a})$ to b.

Locations. A location is a pair (f, \overline{a}) with

- f an *n*-ary function symbol
- $\overline{a} \in \mathsf{BaseSet}^n$ an *n*-tuple

Examples

(parent, a), (firstChild, a), (nextSibling, a), (c,)

An update is a triple (f, \overline{a}, b) with

- (f, \overline{a}) a location
- f not static
- $b \in \mathsf{BaseSet}$
- if f is relational, then $b \in \{tt, ff\}$ An update is trivial if $f_{\mathcal{A}}(\overline{a}) = b$

Intended meaning:

f is changed by changing $f(\overline{a})$ to b.

Generalizations of transition systems

- More detailed description of states: Abstract state machines
- Emphasis on processes and their interdependency: CSP
- Durations: Timed automata
- Continuous evolution + discrete control: Hybrid automata

• transition systems + timing constraints

A timed automaton is a finite automaton extended with a finite set of real-valued clocks. During a run of a timed automaton, clock values increase all with the same speed. Along the transitions of the automaton, clock values can be compared to integers. These comparisons form guards that may enable or disable transitions and by doing so constrain the possible behaviors of the automaton. Further, clocks can be reset.

A timed automaton is a finite automaton extended with a finite set of real-valued clocks. During a run of a timed automaton, clock values increase all with the same speed. Along the transitions of the automaton, clock values can be compared to integers. These comparisons form guards that may enable or disable transitions and by doing so constrain the possible behaviors of the automaton. Further, clocks can be reset.

Timed automata can be used to model and analyse the timing behavior of computer systems, e.g., real-time systems or networks.

Example: Simple Light Control



WANT: if press is issued twice quickly then the light will get brighter; otherwise the light is turned off.

Example: Simple Light Control



Solution: Add a real-valued clock ×

Adding continuous variables to transition systems

Timed automata: Syntax

- A finite set *Loc* of locations
- A subset $Loc_0 \subseteq Loc$ of initial locations
- A finite set Act of labels (alphabet, actions)
- A finite set X of clocks
- Invariant Inv(I) for each location $I \in Loc$: (clock constraint over X)
- A finite set E of edges. Each edge has:
 - source location I, target location I'
 - label $a \in Act$ (empty labels also allowed)
 - guard g (a clock constraint over X)
 - a subset X' of clocks to be reset

For a timed automaton

$$A = (Loc, Loc_0, Act, X, \{Inv_l\}_{l \in Loc}, E)$$

define an infinite state transition system S(A):

- States S: a state s is a pair (I, v), where
 I is a location, and
 v is a clock vector, mapping clocks in X to R, satisfying Inv(I)
- Initial States: (I, v) is initial state if I is in Loc_0 and v(x) = 0
- Elapse of time transitions: for each nonnegative real number d, $(l, v) \xrightarrow{d} (l, v + d)$ if both v and v + d satisfy Inv(l)
- Location switch transitions: $(I, v) \xrightarrow{a} (I', v')$ if there is an edge (I, a, g, X', I') such that v satisfies g and $v' = v[\{x \mapsto 0 \mid x \in X'\}].$

Example: Simple Light Control



Timed automaton:

 $Loc = \{Off, Light, Bright\}, Loc_0 = \{Off\}, Act = \{Press\}$ $X = \{x\}; Inv(Off) = Inv(Light) = Inv(Bright) = (x \ge 0)$

Edges: (Off, Press, \top , {x}, Light), (Light, Press, x > 3, \emptyset , Off) (Light, Press, $x \le 3$, \emptyset , Bright), (Bright, Press, \top , \emptyset , Off)

Example: Simple Light Control



States: (Off, v), (Light, v), (Bright, v) (v value of clock x). **Initial state:** (Off, 0).

Transitions (Examples) **Elapse of time:** $(Off, 10) \xrightarrow{5} (Off, 15)$ **Location switch:** $(Off, 10) \xrightarrow{Press} (Light, 0)$



Hybrid Automata



f : R -> R evolution of external temperature

h : R -> R evolution of heater temperature

Hybrid Automata

Hybrid automaton (HA) S = (X, Q, flow, Inv, Init, E, jump) where:

- (1) $X = \{x_1, ..., x_n\}$ finite set of real valued variables Q finite set of control modes
- (2) {flow_q | q ∈ Q} specify the continuous dynamics in each control mode (flow_q predicate over {x₁,..., x_n} ∪ {x₁,..., x_n}).
- (3) {Inv_q | $q \in Q$ } mode invariants (predicates over X).
- (4) {Init_q | $q \in Q$ } initial states for control modes (predicates over X).
- (5) E: control switches (finite multiset with elements in $Q \times Q$).
- (6) {guard_e | $e \in E$ } guards for control switches (predicates over X).
- (7) Jump conditions {jump_e | $e \in E$ }, (predicates over $X \cup X'$), where $X' = \{x'_1, \ldots, x'_n\}$ is a copy of X consisting of "primed" variables.

Linear Hybrid Automata

Atomic linear predicate: linear inequality (e.g. $3x_1 - x_2 + 7x_5 \le 4$).

Convex linear predicate: finite conjunction of linear inequalities.

A state assertion s for S: family $\{s(q) \mid q \in Q\}$, where s(q) is a predicate over X (expressing constraints which hold in state s for mode q).

Definition [Henzinger 1997] A linear hybrid automaton (LHA) is a hybrid automaton which satisfies the following requirements: (1) Linearity:

- For every $q \in Q$, flow_q, Inv_q , and $Init_q$ are convex linear predicates.

- For every $e = (q, q') \in E$, jump_e and guard_e are convex linear predicates. We assume that flow_q are conjunctions of *non-strict* inequalities.

(2) Flow independence:

For every $q \in Q$, flow_q is a predicate over X only.



Chemical plant

Two substances are mixed; they react; the resulting product is filtered out; then the procedure is repeated.





Chemical plant

Two substances are mixed; they react; the resulting product is filtered out; then the procedure is repeated.

Check:



- No overflow
- Substances in the right proportion
- If substances in wrong proportion, tank can be drained in \leq 200s.



Mode 1: Fill Temperature is low, 1 and 2 do not react. Substances 1 and 2 (possibly mixed with a small quantity of 3) are filled in the tank in equal quantities up to a margin of error.



If proportion not kept: system jumps into mode 4 (**Dump**); If the total quantity of substances exceeds level L_f (tank filled) the system jumps into mode 2 (**React**).



Mode 2: React Temparature is high. Substances 1 and 2 react. The reaction consumes equal quantities of substances 1 and 2 and produces substance 3.



If the proportion between substances 1 and 2 is not kept the system jumps into mode 4 (**Dump**); If the total quantity of substances 1 and 2 is below some minimal level min the system jumps into mode 3 (**Filter**).



Mode 3: Filter Temperature is low. Substance 3 is filtered out.



If proportion not kept: system jumps into mode 4 (**Dump**); Otherwise, if the concentration of substance 3 is below some minimal level min the system jumps into mode 1 (**Fill**).



Mode 4: Dump The content of the tank is emptied. For simplicity we assume that this happens instantaneously:

$$Inv_4 : \bigwedge_{i=1}^3 x_i = 0$$
 and $Iow_4 : \bigwedge_{i=1}^3 x_i = 0$.



The material on ASMs is not required for the exam (only the general idea) The definitions of timed automata and hybrid automata are required for the exam.