Formal Specification and Verification

Formal specification (2) 3.06.2014

Viorica Sofronie-Stokkermans

e-mail: sofronie@uni-koblenz.de

Until now

- Logic
- Formal specification (generalities)

Algebraic specification

Transition systems

Formal specification

• Specification languages for describing programs/processes/systems

- Model based specification
 - transition systems, abstract state machines, specifications based on set theory

Axiom-based specification

- algebraic specification
- Declarative specifications
 - logic based languages (Prolog)
 - functional languages, λ -calculus (Scheme, Haskell, OCaml, ...)
 - rewriting systems (very close to algebraic specification): ELAN, SPIKE, ...
- Specification languages for properties of programs/processes/systems Temporal logic

Formal specification

• Specification languages for describing programs/processes/systems

- Model based specification
 - transition systems, abstract state machines, specifications based on set theory
- Axiom-based specification
 - algebraic specification
- Declarative specifications
 - logic based languages (Prolog)
 - functional languages, λ -calculus (Scheme, Haskell, OCaml)
 - rewriting systems (very close to algebraic specification): ELAN, SPIKE
- Specification languages for properties of programs/processes/systems Temporal logic

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

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

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 $(I, g, \alpha, I') \in \rightarrow$.

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
- $L(\langle I, \beta \rangle) = \{I\} \cup \{g \in Cond(Var) \mid \beta \models g\}.$

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.

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 ≤ 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 minima 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 $flow_4 : \bigwedge_{i=1}^3 \dot{x_i} = 0$



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.

More complex specifications and specification languages

- Languages for describing various processes
- Languages based on Set theory (OZ, B)
- Languages for describing durations
- Complex languages

More complex specifications and specification languages

- Languages for describing various processes
- Languages based on Set theory (OZ, B)
- Languages for describing durations
- Complex languages

Communicating Sequential Processes, or CSP, is a language for describing processes and patterns of interaction between them.

It is supported by an elegant, mathematical theory, a set of proof tools, and an extensive literature. Communicating Sequential Processes, or CSP, is a language for describing processes and patterns of interaction between them.

It is supported by an elegant, mathematical theory, a set of proof tools, and an extensive literature.

- Each process: transition system
- Operations on processes: sequential, parallel composition

efects on states

General idea:

Given:

- Set of event names
- Process: behaviour pattern of an object (insofar as it can be described in terms of the limited set of events selected as its alphabet)

CSP

Example:



Events: insert-coin, get-sprite, get-beer

Prefix:

 $P = a \rightarrow Q$

where a is an event and Q a process

After event a, process P behaves like process Q

(a then Q)

CSP: Example



A simple vending machine which consumes one coin before breaking

(insert-coin \rightarrow STOP)

CSP: Example



A simple vending machine that successfully serves two customers before breaking

 $(insert-coint \rightarrow (get-sprite \rightarrow (insert-coin \rightarrow (get-beer \rightarrow STOP))))$

Example: (recursive definitions)

Consider the simplest possible everlasting object, a clock which never does anything but tick (the act of winding is deliberately ignored)

 $Events(CLOCK) = \{tick\}$

Consider next an object that behaves exactly like the clock, except that it first emits a single tick

$$(tick \rightarrow CLOCK)$$

The behaviour of this object is indistinguishable from that of the original clock. This reasoning leads to formulation of the equation

$$CLOCK = (tick \rightarrow CLOCK)$$

This can be regarded as an implicit definition of the behaviour of the clock.

Modular Specifications: CSP-OZ-DC (COD)

COD [Hoenicke,Olderog'02] allows us to specify in a modular way:

- the control flow of a system using Communicating Sequential Processes (CSP)
- the state space and its change using Object-Z (OZ)
- (dense) real-time constraints over durations of events using the Duration Calculus (DC)



RBC method enter : [s1? : Segment; t0? : Train; t1? : Train; t2? : method leave : [ls? : Segment; lt? : Train] local_chan alloc. reg. updPos. updSpd	Train]	
$\begin{array}{rcl} \text{main} & \stackrel{c}{=} & ((enter \rightarrow \text{main}) \\ & & & (leave \rightarrow \text{main}) \\ & & & (updSpd \rightarrow State1)) \\ \text{State1} & \stackrel{c}{=} & ((enter \rightarrow State1) \\ & & & (leave \rightarrow State1) \\ & & & (leave \rightarrow State2)) \\ \hline & & & \text{SegmentData} \\ \hline & & & \text{train : Segment} \rightarrow \text{Train} & & [Train \text{ on segment}] \\ & & & \text{req : Segment} \rightarrow \mathbb{Z} & & [\text{Requested by train}] \\ & & & alloc : \text{Segment} \rightarrow \mathbb{Z} & & [\text{Allocated by train}] \\ \end{array}$	$\begin{array}{llllllllllllllllllllllllllllllllllll$	CSP
$\begin{array}{l} sd: SegmentData\\ td: TrainData\\\\\hline \forall t: Train\Gamma tid(t) > 0\\ \forall t1, t2: Train \mid t1 \neq t2\Gamma tid(t1) \neq tid(t2)\\ \forall s: Segment\Gamma prevs(nexts(s)) = s\\ \forall s: Segment\Gamma nexts(prevs(s)) = s\\ \forall s: Segment\Gamma sid(s) > 0\\ \forall s: Segment\Gamma sid(nexts(s)) > sid(s)\\ \forall s1, s2: Segment \mid s1 \neq s2\Gamma sid(s1) \neq sid(s2)\\ \forall s: Segment \mid s \neq snil\Gamma length(s) > d + gmax \cdot \Delta t\\ \end{array}$	$ \begin{array}{l} \text{Init} \\ \forall t : \text{Train} \lceil \text{train}(\text{segm}(t)) = t \\ \forall t : \text{Train} \lceil \text{next}(\text{prev}(t)) = t \\ \forall t : \text{Train} \lceil \text{prev}(\text{next}(t)) = t \\ \forall t : \text{Train} \lceil 0 \leq \text{pos}(t) \leq \text{length}(\text{segm}(t)) \\ \forall t : \text{Train} \lceil 0 \leq \text{spd}(t) \leq \text{lmax}(\text{segm}(t)) \\ \forall t : \text{Train} \lceil \text{alloc}(\text{segm}(t)) = tid(t) \\ \forall t : \text{Train} \lceil \text{alloc}(\text{nexts}(\text{segm}(t))) = tid(t) \\ \lor \text{ length}(\text{segm}(t)) - bd(\text{spd}(t)) > \text{pos}(t) \\ \forall s : \text{Segment} \lceil \text{segm}(\text{train}(s)) = s \end{array} $	ΟΖ
$ \begin{array}{l} \forall s : Segment \mid s \neq snill(0 < lmax(s) \land lmax(s) \leq gmax \\ \forall s : Segment \mid s \neq snill(0 < lmax(s) \land lmax(s) \leq gmax \\ \forall s : Segment \lceil lmax(s) \geq lmax(prevs(s)) - decmax \cdot \Delta t \\ \forall s1, s2 : Segment \lceil tid(incoming(s1)) \neq tid(train(s2)) \\ \hline \\ \hline \\ \Delta(spd) \\ \hline \\ \hline \\ \forall t : Train \mid pos(t) < length(segm(t)) - d \land spd(t) - decmax \\ (0, spd(t) - decmax \cdot \Delta t) \leq spd'(t) \leq lmax(segm(t)) - d \land alloc(nexts(segm(t)) - all$	$\begin{array}{l} \max \cdot \Delta t > 0 \\ n(t)) \\ gm(t))) = tid(t) \\ ((segm(t)), max(nexts(segm(t)))\} \\ (segm(t))) = tid(t) \end{array}$	

CSP part: specifies the processes and their interdependency.

The RBC system passes repeatedly through four phases, modeled by events:

- updSpd (speed update)
- req (request update)
- alloc (allocation update)
- updPos (position update)



Between these events, trains may leave or enter the track (at specific segments), modeled by the events leave and enter.

CSP part: specifies the processes and their interdependency.

The RBC system passes repeatedly through four phases, modeled by events with corresponding COD schemata:

CSP:

method enter : [s1? : Segment; t0? : Train; t1? : Train; t2? : Train]
method leave : [ls? : Segment; lt? : Train]
local_chan alloc, req, updPos, updSpd

 $\begin{array}{ll} \text{main} \stackrel{c}{=} ((updSpd \rightarrow State1) & State1 \stackrel{c}{=} ((req \rightarrow State2) & State2 \stackrel{c}{=} ((alloc \rightarrow State3) & State3 \stackrel{c}{=} ((updPos \rightarrow \text{main})) \\ & \Box (leave \rightarrow \text{main}) & \Box (leave \rightarrow State1) & \Box (leave \rightarrow State2) & \Box (leave \rightarrow State3) \\ & \Box (enter \rightarrow \text{main})) & \Box (enter \rightarrow State1)) & \Box (enter \rightarrow State2)) & \Box (enter \rightarrow State3)) \end{array}$

OZ part. Consists of data classes, axioms, the Init schema, update rules.

OZ part. Consists of data classes, axioms, the Init schema, update rules.

• 1. Data classes declare function symbols that can change their values during runs of the system



SegmentData	
train : Segment \rightarrow Train	
req : Segment $\rightarrow \mathbb{Z}$	[Train on segment] [Requested by train]
and . Segment $\rightarrow \mathbb{Z}$	[Allocated by train]

TrainData	
segm : Train \rightarrow Segment	
	[Train segment]
next : $\mathit{Train} \rightarrow \mathit{Train}$	[Next train]
spd : Train $ ightarrow \mathbb{R}$	[Speed]
pos : Train $ ightarrow \mathbb{R}$	[Current position]
prev : Train $ ightarrow$ Train	[Prev. train]

OZ part. Consists of data classes, axioms, the Init schema, update rules.

- 1. Data classes declare function symbols that can change their values during runs of the system, and are used in the OZ part of the specification.
- 2. Axioms: define properties of the data structures and system parameters which do not change
 - $gmax : \mathbb{R}$ (the global maximum speed),
 - $decmax : \mathbb{R}$ (the maximum deceleration of trains),
 - $d : \mathbb{R}$ (a safety distance between trains),
 - Properties of the data structures used to model trains/segments

OZ part. Consists of data classes, axioms, the Init schema, update rules.

- 3. Init schema. describes the initial state of the system.
 - trains doubly-linked list; placed correctly on the track segments
 - all trains respect their speed limits.
- 4. Update rules specify updates of the state space executed when the corresponding event from the CSP part is performed.

Example: Speed update

 $\begin{array}{l} \begin{array}{c} \begin{array}{c} \mbox{effect_updSpd_} \\ \Delta(spd) \end{array} \end{array} \\ \hline \forall t: \mbox{Train} \mid pos(t) < \mbox{length}(segm(t)) - d \land spd(t) - \mbox{decmax} \cdot \Delta t > 0 \\ & \mbox{\sc max}\{0, spd(t) - \mbox{decmax} \cdot \Delta t\} \leq \mbox{spd}^{\prime}(t) \leq \mbox{Imax}(segm(t)) \\ \forall t: \mbox{Train} \mid pos(t) \geq \mbox{length}(segm(t)) - d \land \mbox{alloc}(nexts(segm(t))) = \mbox{tid}(t) \\ & \mbox{\sc max}\{0, spd(t) - \mbox{decmax} \cdot \Delta t\} \leq \mbox{spd}^{\prime}(t) \leq \mbox{min}\{\mbox{Imax}(segm(t)), \mbox{Imax}(nexts(segm(t)))\} \\ \forall t: \mbox{Train} \mid pos(t) \geq \mbox{length}(segm(t)) - \mbox{d} \land \neg \mbox{alloc}(nexts(segm(t))) = \mbox{tid}(t) \\ & \mbox{\sc spd}^{\prime}(t) = \mbox{max}\{0, spd(t) - \mbox{decmax} \cdot \Delta t\} \end{array}$

Formal specification

- Specification for program/system
- Specification for properties of program/system

Verification tasks:

Check that the specification of the program/system has the required properties.