

Semantics of UML State Machines

Alexander Knapp

Institut für Informatik, Ludwig-Maximilians-Universität München
knapp@pst.ifi.lmu.de

Abstract. The abstract syntax and semantics of a simplified subclass of UML state machines is defined.

1 UML State Machines

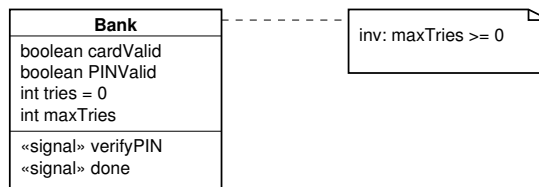
We illustrate the main concepts of UML state machines by a simple UML model of an automatic teller machine (ATM), shown in Fig. 1: The class diagram in Fig. 1(a) specifies an (active) class `Bank`. Classes define *attributes*, i.e., local variables of its instances, and *operations* and *signals* that may be invoked on instances by call and send actions, respectively.

The state machine for class `Bank` is shown in Fig. 1(b), consisting of *states* and *transitions* between states (we number the states for short reference later on). States can be *simple* (such as `Idle` and `DispenseMoney`) or *composite* (such as `Verifying`); a *concurrent* composite state contains several *orthogonal regions*, separated by dashed lines. Moreover, *fork* and *join* (pseudo-)states, shown as bars, synchronize several transitions to and from orthogonal regions; *junction* (pseudo-)states, represented as filled circles, chain together multiple transitions. Transitions between states are triggered by *events*. Transitions may also be guarded by conditions and specify actions to be executed or events to be emitted when the transition is fired. For example, the transition leading from state `Idle` to the fork pseudostate requires signal `verifyPIN` to be present; the transition branch from `VerifyingCard` to `CardValid` requires the guard `cardValid` to be true; the transition branches to `Idle` set the `Bank` attributes `tries` and `cardValid`. Events may also be emitted by *entry* and *exit* actions that are executed when a state is activated or deactivated. Transitions without an explicit trigger (e.g. the transition leaving `DispenseMoney`), are called *completion transitions* and are triggered by *completion events* which are emitted when a state completes all its internal activities.

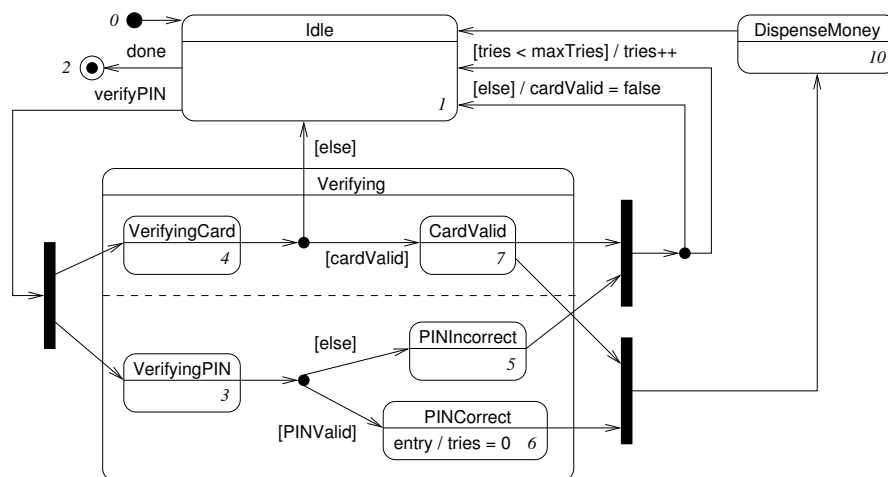
The actual state of a state machine is given by its *active state configuration* and by the contents of its *event queue*. The active state configuration is the tree of active states; in particular, for every concurrent composite state each of its orthogonal regions is active. The event queue holds the events that have not yet been handled by the machine. The *event dispatcher* dequeues the first event from the queue; the event is then processed in a *run-to-completion* (RTC) step. First, a maximally consistent set of enabled transitions is chosen: a transition is *enabled* if all of its source states are contained in the active state configuration, if its trigger is matched by the current event, and if its guard is true; two enabled transitions are *consistent* if they do not share a source state. For each transition in the set, its *least common ancestor* (LCA) is determined,

i.e. the lowest composite state that contains all the transition's source and target states. The transition's main source state, that is the direct substate of the LCA containing the source states, is deactivated, the transition's actions are executed, and its target states are activated.

The example state machine simulates card and PIN validation of a bank computer. After initialization the bank computer is in state *Idle*. The reception of signal *done* leads to finalizing the state machine, whereas on reception of signal *verifyPIN* the verification process is started in state *Verifying*. If the card is invalid, the bank computer immediately returns to state *Idle*. If the PIN is invalid, it is checked whether the maximum number of trials is exceeded. If this is the case, the card is marked invalid; otherwise the number of trials is incremented by one. In both cases, the bank computer returns to state *Idle*. If the PIN is valid, the number of trials is reset to zero. If both the PIN and the card are valid, state *DispenseMoney* is entered from which the bank computer returns to state *Idle*.



(a) Class diagram



(b) State machine diagram for class Bank

Fig. 1. UML model of an ATM

2 Semantics of UML State Machines

The semantics of UML state machines is defined by an execution algorithm. This algorithm forms the basis for embedding UML state machines into temporal logic and, in particular, the symbolic execution technique. Our semantical account of UML state machines follows the semantics definition of the UML 1.5 specification [2] as closely as possible, but fills in some of the gaps of the specification. We mainly follow the ideas presented by Lilius and Porres [1]. However, we refine the definition of compound transitions and maximally consistent sets of compound transitions.

We first define the abstract syntax of the sublanguage of UML state machines from the UML specification [2] for which our semantics is valid. Apart from those language constructs which we do not discuss here (i.e. history, sync, and choice pseudostates, call and deferred events, and internal transitions), this definition introduces the following restriction: A transition from an initial pseudostate must target a non-pseudostate contained in the same composite state as the initial pseudostate. Abandoning this restriction would lead to a more intricate definition of compound transitions. In particular, when an initial pseudostate could target a junction pseudostate, this would amount to treating dynamic choice points as the guards on the transitions from such a junction pseudostate have to be evaluated only after the state containing the initial pseudostate has been entered. However, such a transition could leave the initial pseudostate's container and thus render the choice of transitions to be executed invalid. This case has no relevance in most practical applications.

The semantics of UML state machines is defined in two steps: First, we present a procedure for statically computing the configurations and the compound transitions of a state machine. Due to our syntactical restriction, compound transitions are trees of state machine transitions (describing the chaining of junctions, forks, and entries of states) with a possible fan-in prefix (describing joins) that represent a move from a configuration into another configuration. Second, we define an algorithm for run-to-completion steps which first computes a maximally consistent set of compound transitions for a given event and then executes the set of compound transitions.

2.1 Abstract Syntax of UML State machines

We assume an expression language Exp that at least includes boolean expressions (like true, false, $e_1 \wedge e_2$, etc.) and an action language Act that at least includes a skip statement, and sequential (;) and parallel (||) composition of statements. Furthermore, we assume a set of events $Event$ which includes * denoting a completion event.

A state s has a kind $kind(s) \in \{\text{initial, final, simple, composite, concurrent, junction, join, fork}\}$, an entry action $entry(s) \in Act$, and an exit action $exit(s) \in Act$. A pseudostate is a state s with $kind(s) \in \{\text{initial, junction, join, fork}\}$; we require that $entry(s) = \text{skip}$ and $exit(s) = \text{skip}$ for each pseudostate s . A composite state is a state s with $kind(s) \in \{\text{composite, concurrent}\}$.

A state hierarchy is given by a tree (S, E) where S is a finite set of states and $E \subseteq S \times S$ a non-empty substate relation such that the constraints below are satisfied. We write $substates(s) = \{s' \in S \mid (s, s') \in E\}$ for the substates of state s :

1. If $substates(s) \neq \emptyset$ then $kind(s) \in \{\text{composite}, \text{concurrent}\}$.
2. If $kind(s) = \text{concurrent}$ then $\#substates(s) \geq 2$ and $kind(s') = \text{composite}$ for all $s' \in substates(s)$.
3. If $kind(s) = \text{composite}$ then $\#\{s \in substates(s) \mid kind(s') = \text{initial}\} \leq 1$.

We further write $container(s)$ for the container state of state s if s is not the root state; $substates^+(s) = \{s' \in S \mid (s, s') \in E^+\}$ for the set of transitive substates of state s ; $substates^*(s) = substates^+(s) \cup \{s\}$ for the set of reflexive, transitive substates of s ; and $initial(s)$ for the initial state contained in the composite state s if it exists. The *least common ancestor* of a set of states $M \subseteq S$ not containing the root state, denoted by $lca(M)$, is the least composite state c w.r.t. E such that $M \subseteq substates^+(c)$; the *least common reflexive ancestor* of $M \subseteq S$, written $lca^-(M)$, is the least state s w.r.t. E such that $M \subseteq substates^*(s)$.

Given a state hierarchy $H = (S, E)$, a *transition* t over H has a source state $source(t) \in S$, a target state $target(t) \in S$, a triggering event $trigger(t) \in Event$, a guard expression $guard(t) \in Exp$, and an effect action $effect(t) \in Act$, such that the following constraints are satisfied:

1. $kind(source(t)) \neq \text{final}$.
2. $kind(target(t)) \neq \text{initial}$.
3. If $kind(source(t)) = \text{initial}$ then $target(t)$ is not a pseudostate.
4. If $kind(source(t)) = \text{initial}$ then $container(target(t)) = container(source(t))$.
5. If $kind(source(t)) = \text{fork}$ then $target(t)$ is not a pseudostate.
6. If $kind(target(t)) = \text{join}$ then $source(t)$ is not a pseudostate.
7. If $kind(source(t)) = \text{composite}$ then $kind(container(source(t))) \neq \text{concurrent}$.
8. If $kind(target(t)) = \text{composite}$ then $kind(container(target(t))) \neq \text{concurrent}$.
9. If $kind(source(t)) \in \{\text{initial}, \text{fork}, \text{join}\}$ then $guard(t) = \text{true}$.
10. If $kind(target(t)) = \text{join}$ then $guard(t) = \text{true}$.
11. If $kind(source(t)) = \text{initial}$ then $effect(t) = \text{skip}$.
12. If $source(t)$ is a pseudostate, then $trigger(t) = *$.

A *state machine* (for a class C) is given by a pair (H, T) where $H = (S, E)$ is a state hierarchy and T a finite set of transitions over H such that the constraints below are satisfied for all $t \in T$. We write $outgoings(s)$ for the set $\{t \in T \mid source(t) = s\}$; $incomings(s)$ for the set $\{t \in T \mid target(t) = s\}$; $sources(M)$ for the set $\{source(t) \mid t \in M\}$; and $targets(M)$ for the set $\{target(t) \mid t \in M\}$:

1. If $kind(s) = \text{initial}$ then $\#outgoings(s) = 1$.
2. If $kind(s) = \text{junction}$ then $\#incomings(s) = 1$ and $\#outgoings(s) \geq 1$.
3. If $kind(s) = \text{fork}$ then $\#incomings(s) = 1$ and $\#outgoings(s) \geq 2$.
4. If $kind(s) = \text{join}$ then $\#outgoings(s) = 1$ and $\#incomings(s) \geq 2$.
5. If $kind(s) = \text{fork}$ then there is an $s' \in S$ with $kind(s') = \text{concurrent}$ such that $targets(outgoings(s)) \subseteq substates^+(s') \setminus substates(s')$ and the following holds: if $t, t' \in outgoings(s)$ such that $\{target(t), target(t')\} \subseteq substates^+(s')$ for some $s'' \in substates^+(s')$ then $t = t'$.
6. If $kind(s) = \text{join}$ then there is an $s' \in S$ with $kind(s') = \text{concurrent}$ such that $sources(incomings(s)) \subseteq substates^+(s') \setminus substates(s')$ and the following holds: if $t, t' \in incomings(s)$ such that $\{source(t), source(t')\} \subseteq substates^+(s')$ for some $s'' \in substates^+(s')$ then $t = t'$.

Conditions (5) and (6) require forks and joins to come from and go to different orthogonal regions of a concurrent composite state.

2.2 Configurations and Compound Transitions

The *configurations* of a state machine $((S, E), T)$ are given by the smallest subsets C of S that satisfy the following conditions:

1. The root state of S is in C .
2. No state $s \in C$ is a pseudostate.
3. If $kind(s) = \text{composite}$ then there is a single $s' \in C$ such that $container(s') = s$.
4. If $kind(s) = \text{concurrent}$ then all states $s' \in S$ with $container(s') = s$ are in C .

In particular, composite states are or-states, concurrent states and-states.

The *compound transitions* of a state machine $((S, E), T)$ represent semantically complete transition paths that originate from a set of non-pseudostates and target a set of simple states. More precisely, a compound transition consists of three parts: The optional tail part of a compound transition may have multiple transitions in T originating from a set of mutually orthogonal regions that are joined by a join pseudostate. The middle part of a compound transition is a finite chain of transitions in T joined via junction pseudostates. Finally, the optional head part of a compound transition is a tree of transitions in T : If a transition in the middle part of a compound transition or in its head part itself targets a composite state the head part continues at the initial state of this composite transition; if a transition targets a concurrent composite state the head part continues at all initial states of the orthogonal regions of the concurrent composite state; if a transition targets a fork pseudostate the head part continues with the transitions outgoing from the fork pseudostate which target mutually orthogonal regions and simultaneously continues at the initial states of all those orthogonal regions that are not targeted by transitions outgoing from the fork pseudostate.

In the ATM example, the compound transitions outgoing from `VerifyingCard` just consist of middle parts:

$$\langle \{\}, \langle \text{VerifyingCard} \rightarrow \text{junction}, \text{junction} \rightarrow \text{Idle} \rangle, \{\} \rangle,$$

$$\langle \{\}, \langle \text{VerifyingCard} \rightarrow \text{junction}, \text{junction} \rightarrow \text{CardValid} \rangle, \{\} \rangle.$$

The fork transition from `Idle` consists of a middle part and a tail part:

$$\langle \{\}, \langle \text{Idle} \rightarrow \text{fork} \rangle, \{ \langle \text{fork} \rightarrow \text{VerifyingCard} \rangle, \langle \text{fork} \rightarrow \text{VerifyingPIN} \rangle \} \rangle.$$

The join transition to `DispenseMoney` consists of a head part and a middle part:

$$\langle \{ \langle \text{CardValid} \rightarrow \text{join} \rangle, \langle \text{PINCorrect} \rightarrow \text{join} \rangle \}, \langle \text{join} \rightarrow \text{DispenseMoney} \rangle, \{\} \rangle.$$

We present an algorithm for computing the compound transitions of a state machine $((S, E), T)$ in two steps: The procedure `compounds` in Fig. 2 for computing the compound transitions outgoing from a non-pseudostate in S relies on the procedure `forwardTrees` in Fig. 3 that computes the middle and head parts of compound transitions,

```

compounds(state)  $\equiv$ 
  compounds  $\leftarrow \emptyset$ 
  joinTargets  $\leftarrow \emptyset$ 
  for forwardTree  $\in$  forwardTrees(state) do
    if join  $\in \{kind(s) \mid s \in targets(forwardTree)\}$ 
      then joinTarget  $\leftarrow choose(\{s \mid s \in targets(forwardTree) \wedge kind(s) = join\})$ 
      if joinTarget  $\notin joinTargets$ 
        then joinTargets  $\leftarrow joinTargets \cup \{joinTarget\}$ 
        for joinFT  $\in$  forwardTrees(joinTarget) do
          compounds  $\leftarrow compounds \cup \{(incomings(joinTarget), joinFT)\}$ 
        od
      fi
    else compounds  $\leftarrow compounds \cup \{(\emptyset, forwardTree)\}$ 
    fi
  od
  compounds ]

```

Fig. 2. Compound transitions algorithms (1).

also called *forward trees*, outgoing from an arbitrary state in S . Herein, the (simple) target states of the leaf transitions of a forward tree f are denoted by $targets(f)$.

Note that our definition of compound transitions deviates from the explanations in the UML specification [2]. There, compound transitions are not required to target simple states only, but may as well stop at composite states. The proper initialization of composite and concurrent composite states is left to the entry procedure for composite states. In fact, as mentioned above, this is necessary for handling choice pseudostates and initial in conjunction with junction pseudostates.

The notions of source states, target states, trigger, guard, and effect are transferred from transitions to compound transitions in the following way: The *source states* of a compound transition τ , written $sources(\tau)$, are the source states of the transitions in the tail part of τ , if τ shows a tail part, and the source state of the middle part, otherwise. Analogously, the target states of τ , written $targets(\tau)$, are the target states of the transitions in the head part of τ , if τ shows a head part, and the target state of the middle part, otherwise. The *trigger* of τ is the set of triggers of the transitions in the tail part of τ , if τ shows a tail part, and the trigger of the first transition in the middle part otherwise. The *guard* of τ is the conjunction of all guards of transitions in τ . Finally, the *effect* of τ is the sequential composition of the effects of the tail, the middle, and the head part of τ , where the effects in the tail and the head are conjoined in parallel whereas the effects in the middle part are composed sequentially. These definitions are naturally extended to sets of compound transitions which show the same trigger.

We recall some notions on compound transitions τ from the UML specification that will be used for the definition of the execution semantics of state machines, in particular, when computing maximally conflict free sets of compound transitions in a given configuration C : The *main source state* of τ , $mainSource(\tau)$, is given by the state $s = lca(lca^=(sources(\tau)), lca^=(targets(\tau)))$ if $kind(s) = \text{concurrent}$, and it is given by the state $s' \in substates(s)$ with $lca^=(sources(\tau)) \in substates^*(s')$, oth-

```

forwardTrees(state) ≡
  forwardTrees ← ∅
  for outgoing ∈ outgoings(state) do
    target ← target(outgoing)
    if kind(target) = junction
      then for targetFT ∈ forwardTrees(target) do
        forwardTrees ← forwardTrees ∪ {⟨outgoing, targetFT⟩}
      od
    elseif kind(target) = fork
      then container ← lca(targets(outgoings(target)))
        regionsFTs ← ⟨⟩
        for region ∈ substates(container) do
          regionTs ← {t ∈ outgoings(target) | target(t) ∈ substates(region)}
          if regionTs ≠ ∅
            then transition ← choose(regionTs)
              if kind(target(transition)) ∈ {composite, concurrent}
                then regionFTs ← ⟨⟩
                  for regionFT ∈ forwardTrees(target(transition)) do
                    regionFTs ← regionFTs ∪ {⟨transition, regionFT⟩}
                  od
                regionsFTs ← regionsFTs :: ⟨regionFTs⟩
              else regionsFTs ← regionsFTs :: {⟨transition⟩}
            fi
          else regionsFTs ← regionsFTs :: ⟨forwardTrees(initial(region))⟩
        fi
      od
      n ← #substates(container)
      for ⟨regionFTsi⟩1 ≤ i ≤ n ∈ regionsFTs do
        forwardTrees ← forwardTrees ∪ {⟨outgoing, ⟨regionFTsi⟩1 ≤ i ≤ n⟩}
      od
    elseif kind(target) = composite
      then for initialFT ∈ forwardTrees(initial(target)) do
        forwardTrees ← forwardTrees ∪ {⟨outgoing, initialFT⟩}
      od
    elseif kind(target) = concurrent
      then regionsFTs ← ⟨⟩
        for region ∈ substates(target) do
          regionsFTs ← regionsFTs :: ⟨forwardTrees(initial(region))⟩
        od
        n ← #substates(target)
        for ⟨regionFTsi⟩1 ≤ i ≤ n ∈ regionsFTs do
          forwardTrees ← forwardTrees ∪ {⟨outgoing, ⟨regionFTsi⟩1 ≤ i ≤ n⟩}
        od
      else forwardTrees ← forwardTrees ∪ {⟨outgoing⟩}
    fi
  od
forwardTrees ]

```

Fig. 3. Compound transitions algorithms (2).

erwise. The *main target state* of τ , $\text{mainTarget}(\tau)$ is defined analogously, but exchanging $\text{sources}(\tau)$ and $\text{targets}(\tau)$. The set of states *exited* by τ in configuration C , $\text{exited}(C, \tau)$, consists of $\text{substates}^*(\text{mainSource}(\tau)) \cap C$. The set of states *entered* by τ in configuration C , $\text{entered}(\tau)$, is $\text{substates}^*(\text{mainTarget}(\tau)) \cap C$. Again, the definitions for *entered* and *exited* are naturally extended to sets of compound transitions.

Two compound transitions τ_1 and τ_2 are *in conflict* in configuration C , written $\tau_1 \#_C \tau_2$, if $\text{exited}(C, \tau_1) \cap \text{exited}(C, \tau_2) \neq \emptyset$; more generally, a compound transition τ is in conflict with a set of compound transitions T in configuration C , written $\tau \#_C T$, if $\tau \#_C \tau'$ for some $\tau' \in T$. If $\tau_1 \#_C \tau_2$ let S_1 and S_2 be the sets of states in $\text{sources}(\tau_1)$ and $\text{sources}(\tau_2)$, resp., that show the maximal numerical distance from the root state of (S, E) ; τ_1 is *prioritized* over τ_2 in configuration C , written $\tau_1 \prec_C \tau_2$, if $S_1 \subseteq \text{substates}^+(S_2)$. Again, $\tau \prec_C T$ for a compound transition τ and a set of compound transitions T in configuration C if $\tau \prec_C \tau'$ for some $\tau' \in T$.

2.3 Run-To-Completion Semantics

The execution semantics of a UML state machine is described in the UML specification as a sequence of *run-to-completion steps*. Each such step is a move from a configuration of the state machine to another configuration. The sequence of steps starts in the *initial configuration* of the state machine, i.e., the configuration that is targeted by the forward tree outgoing from the initial state of the root state of the state machine's state hierarchy. In a run-to-completion step from some configuration, first, an event is fetched from the event queue. Second, a maximally consistent set of enabled compound transitions outgoing from the states of the current configuration and whose guards are satisfied is chosen. If such a set, also called a *step*, exists, all its compound transitions are fired simultaneously: First, all states that are exited by the step are left in an inside-out manner, executing the exit actions of these states; each such that is marked to be not completed, as it is not part of the configuration any more. Second, the gathered effect of the step is executed. Third, all states that are entered by the step are entered in an outside-in manner, executing the entry actions of these states. Furthermore, after executing the entry action of a state this state is marked as complete, i.e. a completion event for this state is generated.

More formally, let $((S, E), T)$ be a state machine. We assume a structure of *environments* η for state machines that provides the following primitive operations: An event can be fetched by $\text{fetch}(\eta)$; the completion of a state s can be recorded by $\text{complete}(\eta, s)$; the revocation of a state s from being completed can be recorded by $\text{uncomplete}(\eta, s)$; a statement a can be executed by $\text{exec}(\eta, a)$; given a configuration C and an event e all compound transitions of $((S, E), T)$ that are triggered by e can be computed by $\text{enabled}(\eta, C, e)$; and, finally, the validity of an expression g can be checked by $\eta \models g$.

The enabledness of compound transitions in a configuration C by an event e is indeed solely defined on the basis of the triggers of compound transitions and thus only involves the completed states that have been previously recorded with the environment. The fireable sets of compound transitions, which are maximally consistent sets of enabled compound transitions are computed by the *steps* algorithm in Fig. 4(a). The execution of a state machine in some configuration and some environment is defined by the


```

steps(env, conf, event) ≡
  [ transitions ← enabled(env, conf, event)
    { step | ⟨guard, step⟩ ∈ steps(conf, transitions) ∧ env ⊨ guard } ]

steps(conf, transitions) ≡
  [ steps ← {⟨false, ∅⟩}
    for transition ∈ transitions do
      for ⟨guard, step⟩ ∈ steps(transitions \ {transition}) do
        if transition †conf step
          then if transition <conf step
            then guard ← guard ∧ ¬ guard(transition)
            fi
          else step ← step ∪ {transition}
            guard ← guard ∧ guard(transition)
          fi
        steps ← steps ∪ {⟨guard, step⟩}
      od
    od
  steps ]

```

(a) Transition selection algorithm.

<pre> RTC(env, conf) ≡ [⟨event, env⟩ ← fetch(env) steps ← steps(env, conf, event) if steps ≠ ∅ then choose step ∈ steps ⟨env, conf⟩ ← fire(env, conf, step) fi ⟨env, conf⟩] </pre>	<pre> fire(env, conf, step) ≡ [for state ∈ insideOut(exited(conf, step)) do env ← exec(env, exit(state)) conf ← conf \ {state} env ← uncomplete(env, state) od env ← exec(env, effect(step)) for state ∈ outsideIn(entered(conf, step)) do env ← exec(env, entry(state)) conf ← conf ∪ {state} env ← complete(env, state) od ⟨env, conf⟩] </pre>
--	--

(b) Run-to-completion step algorithm.

(c) Transition firing algorithm.

Fig. 4. State machine execution algorithms.

RTC algorithm in Fig. 4(b) which uses the algorithm for firing a compound transitions step in Fig. 4(c).

References

1. Johan Lilius and Iván Porres Paltor. Formalising UML State Machines for Model Checking. In Robert France and Bernhard Rumpe, editors, *Proc. 2nd Int. Conf. Unified Modeling Language (UML '99)*, volume 1723 of *Lect. Notes Comp. Sci.*, pages 430–445. Springer, Berlin, 1999.

2. Object Management Group. Unified Modeling Language Specification, Version 1.5. Specification, OMG, 2003. <http://www.omg.org/cgi-bin/doc?formal/03-03-01>.