

A New Coordination Language MediatE

Yi Li and Meng Sun

LMAM and Department of Informatics, School of Mathematical Sciences, Peking
University, Beijing, China
liyi_math@pku.edu.cn, sunmeng@math.pku.edu.cn

Abstract. tbd

1 Introduction

2 Overview

The goal of this language MediatE mainly focus on:

1. Compositional Verification.

3 Syntax of MediatE

$$\langle program \rangle ::= (\langle import \rangle \mid \langle typedef \rangle \mid \langle function \rangle \mid \langle automaton \rangle \mid \langle system \rangle)^*$$

In the following subsections, we are going to take a simple *queue* as an example, to illustrate how certain language elements are used to compose a model.

3.1 Type System

MediatE provides a rich-featured type system that supports various commonly-used data types in both formal modeling languages and programming languages.

Primitive Type. Table. 1 shows the primitive types supported in MediatE.

Table 1. Primitive Data Types

Name	Declaration	Term Example
Bounded Integer	<code>int lowerBound .. upperBound</code>	<code>-1, 0, 1</code>
Integer	<code>int</code>	<code>-1, 0, 1</code>
Real	<code>real</code>	<code>0.1, 1E-3</code>
Boolean	<code>bool</code>	<code>true, false</code>
Character	<code>char</code>	<code>'a', 'b'</code>
Enumeration	<code>enum item₁, ..., item_n</code>	<code>enumname.item</code>

Composite Type. Composite type offers an approach to construct complex data types with simpler ones. Several composite patterns are introduced as follows,

Table 2. Composite Data Types (T denotes an arbitrary data type)

Name	Declaration
Tuple	T_1, \dots, T_n
Union	$T_1 \mid \dots \mid T_n$
Array	$T \text{ [length]}$
Slice	$T \text{ []}$
Map	$\text{map } [T_{key}] T_{value}$
Struct	$\text{struct } \{ \text{field}_1:T_1, \dots, \text{field}_n:T_n \}$
Initialized	$T_{base} \text{ init term}$

- *Tuple*. The *tuple* operator ‘,’ can be used to construct a finite tuple type with several base types.
- *Union*. The *union* operator ‘|’ is designed to combine *disjoint* types as a more complicated one. This is similar to the union type in C language but much easier to use.
- *Array* and *Slice*. An *array* $T[n]$ is a finite ordered collection containing exactly n elements of type T . Moreover, a *slice* is an array of which the capacity is not specified, i.e. slice is a dynamic array.
- *Map*. A *map* $[T_{key}] T_{val}$ is a dictionary that maps a key of type T_{key} to a value of type T_{val} .
- *Struct*. A *struct* $\{ \text{field}_1 : T_1, \dots, \text{field}_n : T_n \}$ contain n fields, each has a particular type T_i and a unique identifier id_i .
- *Initialized*. A initialized type make it able to specify default values to types.

For simplicity in formalizing data types, we introduce the concept *domain* of a type.

Formalization 1 (Domain) We use $Dom(T)$ to denote the value domain of type T , i.e. the set of all possible value of T .

Example 1 (Types Used in A Queue). Now let us introduce some type declarations and local variables used in an automaton **Queue**. As shown in the following code fragment, we declares a singleton enumeration **NULL**, which contains only one element **null**. The buffer of a queue is in turn formalized as an array of **T** or **NULL**, indicating that a queue element can be either an assigned item or empty. The head and tail pointer are defined as two bounded integers.

```

1  typedef enum {null} init null as NULL;
2  automaton <T:type,size:int> Queue(A:in T, B:out T) {
3      variables {
4          buf : ((T | NULL) init null) [size];
5          phead : int 0 .. (size - 1) init 0;
6          ptail : int 0 .. (size - 1) init 0;
7      }
8      ...
9  }
```

3.2 Functions

The abstract syntax tree of functions is shown as follows.

```

<funcDecl> ::= function <template>? <identifier> ( <arguments> ) {
    ( variables { <varDecl>* } )?
    statements { <assignStmt>* <returnStmt> }
<assignStmt> ::= <term> := <term>
<returnStmt> ::= return <term>
<varDecl> ::= <identifier> : <type> ( init <term> )?

```

Basically, definition of a function includes:

- An optional template including a set of parameters. A parameter can be either a type parameter (decorated by **type**) or a value parameter (decorated by its type). All possible parameter values of a function should be located statically. Parameters in the template can be used in all the following language elements, e.g. type of input variables and return value, local variables and function statements.
- An identifier that indicates the name of this function.
- A set of read-only input variables.
- A optional set of local variables.
- A list of ordered statements that describes how the return value is calculated. Such a list must be ended by a **return** statement.

Functions in MediatE are side-effect free. In other words, only local variables are writable in its assignment statements.

Example 2 (Incline Operation on a Queue Pointer). The simple function describes how pointers are inclined. When a pointer is going to exceed its upper bound (determined by the parameter *size*), we will reset it to zero.

```

1  function <size:int> next(pcurr:int 0..(size-1)) : int 0..(size-1) {
2      statements { return (pcurr + 1) % size; }
3  }

```

3.3 Automata : The Basic Behavioral Unit

```

<automaton> ::= automaton <template>? <identifier> ( <port>* ) {
    ( variables { <varDecl>* } )?
    transitions { <transition>* } }
<port> ::= <identifier> : ( in | out ) <type>
<transition> ::= <guardedStmt> | group { <guardedStmt>* }
<guardedStmt> ::= <term> -> ( <stmt> | { <stmt>* } )
<stmt> ::= <term> := <term> | perform <identifier>+

```

Template. Very similar to functions, a automaton can also be decorated with a set of template parameters, either value parameters or type parameters.

Ports. Each automaton contains a set of ports, either **in**-coming or **out**-going, to communicate with the environment. To ensure the well-defineness of automata, ports are required to have an *initialized* type, e.g. `int 0..1 init 0` instead of ~~`int 0..1`~~.

Variables. Two types of variables are used in a automaton definition, they are:

1. *Local variables* that are declared in the *variables* section. A local variable can only be referenced in its scope, i.e. the automaton definition. And similar to the ports, only initialized types are permitted when declaring local variables.
2. *Adjoint variables* that are used to describe the status of ports. For a port **A**, we assume that it has two boolean fields **A.reqRead** and **A.reqWrite** indicating if there is a pending *read* or *write* request on this port, and a data field **A.value** indicating the current value of this port (if a write operation is performed, **A.value** will be reassigned).

A reasonable rule comes up that, both the **reqWrite** field of a input port and the **reqRead** field of a output port are *read-only*. Similarly, we cannot rewrite the **value** field of a input port.

Transitions. Similar to the PRISM[6] language, behavior of a channel in MediatE is described by a series of guarded transitions (groups). As shown in Example 3, a *transition* comprises two parts: a boolean term *guard* that shows on what condition the transition could be fired, and a (set of) statement(s) that describe what will happen if the transition is fired. Two types of statements are supported in automata,

- *Assignment Statements* (`var1, ..., varn := term1, ..., termn`). Local variables and writable adjoint variables are permitted to be assigned here. We can also assign several variables at the same time (similar to the tuple assignment in Python).
- *Perform Statements* (`perform port1, ..., portn`). Informally speaking, perform statements tell the environment to fire data operations on the output ports, or wait until being noticed that data operation on the input ports are fired by the environment (other automata, actually). Consequently, it's reasonable to require that the value of an input port should never be referred until the port is performed. Similarly, the value of an output port should never be assigned after the port is performed. Perform statements are mainly used when combining multiple automata, where they determine how transitions are synchronized. (See in Section 4.3)

When guard of a transition is satisfied by the context, we say the transition is *activated*. However, being activated is only necessary condition of being fired. When choosing a transition to fire, we have to consider other criteria, which will be introduced later.

Though not mentioned explicitly, transitions can be divided into two classes: *external* and *internal*. A transition is external iff. perform assignments are involved to model the interaction with environment. Hereinafter we use $g \rightarrow S$ to denote a transition, where g is the guard formula and S is the set of statements.

Example 3 (Transitions in Queue). In a **Queue**, we use internal transitions to formalize the changes of its state. For example, becoming writable when buffer is not full, and readable when buffer is not empty. External transitions, on the other hand, mainly show how the read and write operations are performed.

```

1  // Internal Transitions
2  true -> B.reqWrite := (buf[ptail] != null) ;
3  true -> A.reqRead := (buf[phead] == null) ;
4
5  // External Transitions
6  (A.reqRead && A.reqWrite) -> {
7      perform A; buf[phead] := A.value; phead := next(phead);
8  }
9  (B.reqRead && B.reqWrite) -> {
10     B.value := buf[ptail]; ptail := next(ptail); perform B;
11 }

```

All the transitions are supposed to have the following features. They are declared on the syntax level, i.e. we will resolve this feature when discussing the formal aspect of MediatE and use a simple and standard automata model to capture all these features (see in Section. 4).

- *Alternative.* A transition won't be fired if it changes nothing in its context. For example, the first internal transition in a **Queue** will not be activated if **B.reqWrite** is already equal to **buf[ptail] != null**. This assumption is mainly used to avoid useless executions.
- *Urgent.* In some formal models, e.g. CSP[5] and Timed Automata[2], transitions may not be triggered even the guard is satisfied. On contrast, such behavior is strictly prohibited in our model. Once a transition is activated (i.e. its guard is satisfied), it have to be fired unless another guard with higher priority is also activated.
- *Ordered.* A channel may includes a list of transitions (single or group). They are ordered by their appearance. In other words, first written, first executed.

Formalization 2 (Transition Groups) *A transition group t_G can be formalized as a list of guarded transitions*

$$t_G = \{t_1, \dots, t_n\}, t_i = g_i \rightarrow S_i$$

where t_i is a single transition with guard g_i and a set of statements S_i .

Formalization 3 (Automata) *We use a tuple $A = \langle Ports, Vars, Trans_G \rangle$ to represent an automaton instance, where $Ports$ is a set of ports, $Vars$ is a set of local variables and $Trans_G = \{t_{G_1}, \dots, t_{G_n}\}$ is a set of transition groups.*

3.4 System : The Composition Approach

Theoretically, an automaton in MediatE is powerful to represent any classical software system (without consideration of time and probability, of course). However, modeling complex systems in transitions and tons of local variables may become a real disaster. That's why we are going to introduce a new block, called *system*, to help reuse existing automata (systems as well), and construct clear and comprehensible high-level models.

To solve this problem, hierarchical diagrams are widely used in various modeling tools (e.g. SCADE[1, 4], Simulink and LabVIEW) and formal languages (e.g. Reo[3], AADL). In such diagrams, blocks can be declared as *components* and organized by a set of connections to capture more powerful behavior, where these connections are called *channels*.

Both *components* and *connectors* (or *channel*) are well-known concepts in component-based software engineering. However, in semantics level, they all turn out to the same nature – *automata*. Following this idea, we present a compositional block named *system*, where automata can be instantiated as either components or channels. The abstract syntax tree of systems is shown as follows.

$$\begin{aligned}
 \langle system \rangle &::= \mathbf{system} \langle template \rangle^? \langle identifier \rangle (\langle port \rangle^*) \{ \\
 &\quad (\mathbf{internals} \langle identifier \rangle^+)^? \\
 &\quad (\mathbf{components} \{ \langle componentDecl \rangle^* \})^? \\
 &\quad \mathbf{connections} \{ \langle connectionDecl \rangle^* \} \} \\
 \langle componentDecl \rangle &::= \langle identifier \rangle : \langle systemType \rangle \\
 \langle connectionDecl \rangle &::= \langle systemType \rangle \langle params \rangle (\langle portName \rangle^+)
 \end{aligned}$$

The interface of a system (i.e. its template, name, and ports) shares exactly the same form with interfaces of automata, which also implies that system is not a special semantics unit, but simply an approach to construct more complex automata. A system is composed of internal nodes (optional), components(optional) and a set of connections.

Internals.

Components.

Connections.

Example 4 (Alternator).

4 Semantics

4.1 Configurations of Automata

Configurations are used to represent the state of an automaton. Since we don't have locations here, it only depends on the values of its locally accessible variables, which includes both *adjoint variables* and *local variables*.

```

1  system <T:type> alternator (A:in T, B:in T, C: out T) {
2    internals A1,A2,B1,B2,C1,C2;
3    connections {
4      Replicator<T>(A,A1,A2);
5      Replicator<T>(B,B1,B2);
6      Merger<T>(C1,C2,C);
7
8      Sync<T>(A1,C1);
9      SyncDrain<T>(A2,B1);
10     Fifo1<T>(B2,C2);
11   }
12 }

1  system testBench () {
2    components {
3      writer1 : Writer<int>; writer2 : Writer<int>;
4      reader : Reader<int>;
5    }
6    connections {
7      Alternator<int>(writer1.out, writer2.out, reader.in);
8    }
9  }

```

Definition 1 (Valuation). An evaluation of a set of variables $Vars$ is defined as a function v that satisfies $\forall x \in Vars, v(x) \in Dom(type(x))$. We denote all the possible evaluations of a variable set $Vars$ as $Val(Vars)$.

Definition 2 (Configuration). A configuration of an automaton $A = \langle Ports, Vars, Trans_G \rangle$ is defined as a tuple (v_{loc}, v_{adj}) where $v_{loc} \in Val(Vars)$ is a valuation on local variables, and $v_{adj} \in Val(Adj(P))$ is a valuation on adjoint variables.

4.2 Canonical Form of Transitions

Definition 3 (Canonical Transitions). A transition $t = g \rightarrow \{s_1, \dots, s_n\}$ is canonical iff. its statements $\{s_i\}$ is an interleaving sequence of assignments and performs which start from an assignment, e.g. $a := exp_1$; **perform** A; $b := exp_2$; \dots .

We only need to simple steps to canonicalize a transition, they are:

1. Merging the contiguous assignments. As mentioned before, an assignment statement is represented as a function $f : EV \rightarrow EV$. Thus a list of multiple assignments f_1, \dots, f_n can be simplified by $f = f_1 \circ \dots \circ f_n$.
2. Any two adjacent performs should be separated by a identical assignment id_{EV} .

Observable. A transition is always *observable*, i.e. it will makes some difference to the context. For example, without this assumption, a transition $\mathbf{true} \rightarrow \mathbf{x} := \mathbf{x}$ will block the whole model by endless meaningless executions.

Definition 4 (Canonical Transition Groups). *A transition group is canonical iff. it only contains one canonical transition.*

Priority. Given a set of ordered transitions.

$$\{g_1 \rightarrow S_1, g_2 \rightarrow S_2, \dots, g_n \rightarrow S_n\}$$

As required by the *priority* assumption, a transition can be fired only if all the previous ones are not enabled (i.e. their guards are not satisfied) yet. In MediatE, this feature is resolved simply by adding $\neg g_i$ to all $g_j (j > i)$. E.g.

$$\{g_1 \rightarrow S_1, g_2 \wedge (\neg g_1) \rightarrow S_2, \dots, g_n \wedge (\neg g_1 \wedge \neg g_2 \wedge \dots \wedge \neg g_{n-1}) \rightarrow S_n\}$$

Now let's consider a set of ordered groups t_{G_i} , where t_{G_i} contains l_i transitions,

$$T_G = \{t_{G_1} = \{g_{11} \rightarrow S_{11}, \dots, g_{1l_1} \rightarrow S_{1l_1}\}, \dots, t_{G_n} = \{g_{n1} \rightarrow S_{n1}, \dots, g_{nl_n} \rightarrow S_{nl_n}\}\}$$

Informally speaking, once a transition in t_{G_1} is enabled, all the other transitions in $t_{G_i} (i > 1)$ should be strictly prohibited from being fired. We use $enab(t_G)$ to denote the condition where at least one transition in t_G is enabled, formalized as

$$enab(t_G = \{g_1 \rightarrow S_1, \dots, g_n \rightarrow S_n\}) = g_1 \vee \dots \vee g_n$$

Then we can generate the new set of transitions with no dependency on priority:

$$\begin{aligned} &g_{11} \rightarrow S_{11}, \dots, g_{1l_1} \rightarrow S_{1l_1}, \\ &g_{21} \wedge \neg enab(t_{G_1}) \rightarrow S_{21}, \dots, g_{2l_2} \wedge \neg enab(t_{G_1}) \rightarrow S_{2l_2}, \dots \\ &g_{n1} \wedge \neg enab(t_{G_1}, \dots, t_{G_{n-1}}) \rightarrow S_{n1}, \dots, g_{nl_n} \wedge \neg enab(t_{G_1}, \dots, t_{G_{n-1}}) \rightarrow S_{nl_n} \end{aligned}$$

where $enab(t_{G_1}, \dots, t_{G_{n-1}})$ is an abbreviation of $enab(t_{G_1}) \vee \dots \vee enab(t_{G_{n-1}})$. It indicates that at least one group in t_{G_i} is enabled.

4.3 From System to Automaton

Systems, as shown previously, are simply introduced to construct automata in a more natural way. Now we show how such a system can be flattened as a standard automaton.

Algorithm 1 Scheduling in a Synchronous Set of External Transitions

Require: t_1, t_2, \dots, t_n are transitions (canonical form)

Ensure: $t = \text{Schedule}(t_1, \dots, t_n)$

```
1: if  $\{t_i\}$  don't belong to different automata or  $\exists t_i$  is internal then
2:    $t \leftarrow \text{null}$ 
3:   return
4: end if
5:  $t.g, t.S \leftarrow \bigwedge_i t_i.g, \{\}$ 
6: for  $i \leftarrow 1, \dots, n$  do
7:   if  $t_i.s_1$  is an assignment then
8:     add  $t_i.s_1$  to the head of  $t.S$ 
9:   end if
10:   $p \leftarrow$  the first perform statement
11:  while  $p \neq \text{null}$  do
12:     $a \leftarrow$  the assignment statement after  $p$ 
13:     $p' \leftarrow$  the next perform statement after  $p$ 
14:    if  $p \in t.S$  then
15:      insert  $a$  to  $t.S$  exactly after  $p$ 
16:      remove  $p$  from  $t.S$ 
17:    end if
18:  end while
19: end for
20:  $t \leftarrow \text{Canonicalize}(t)$ 
```

Algorithm 2 Compose Several Automata

Require: A_1, A_2, \dots, A_n are automata

Ensure: $A = \text{Compose}(A_1, \dots, A_n)$

```
1: rename local variables in  $A_1, \dots, A_n$  to avoid duplicated names
2:  $A \leftarrow$  empty automaton
3:  $\text{ext\_trans} \leftarrow \{\}$ 
4: for  $i \leftarrow 1, 2, \dots, n$  do
5:   add all local variables of  $A_i$  to  $A$ 
6:   add all internal transitions of  $A_i$  to  $A$ 
7:   add all external transitions of  $A_i$  to  $\text{ext\_trans}$ 
8: end for
9: for  $\text{set\_trans} \leftarrow$  subset of  $\text{ext\_trans}$  do
10:   $\text{newedge} \leftarrow \text{Schedule}(\text{set\_trans})$ 
11:  if  $\text{newedge} \neq \text{null}$  then
12:    add  $\text{newedge}$  to  $A$ 
13:  end if
14: end for
```

4.4 Automaton as Labelled Transition System

Definition 5 (Transition System, TS). A transition system is a tuple (S, \rightarrow) where S is a set of states and $\rightarrow \subseteq S \times \Sigma \times S$ is a set of transitions. For simplicity reason, we use $s \rightarrow s'$ to denote (s, s') in \rightarrow .

Suppose $A = \langle Ports, Vars, Trans_G \rangle$ is an automaton, its semantics can be captured by a labelled transition system $\langle S_A, \rightarrow_A \rangle$ where

- S_A is the set of all configurations of A .
- $\rightarrow_A \subseteq S_A \times \Sigma_A \times S_A$ is a set of transitions constructed by the following rules.

$$\begin{array}{c}
\frac{p \in P_{in}}{(v_{loc}, v_{adj}) \rightarrow_A (v_{loc}, v_{adj}[p.reqWrite \mapsto \neg p.reqWrite])} \text{R-INPUTSTATUS} \\
\\
\frac{p \in P_{in}, val \in type(p.value)}{(v_{loc}, v_{adj}) \rightarrow_A (v_{loc}, v_{adj}[p.value \mapsto val])} \text{R-INPUTVALUE} \\
\\
\frac{p \in P_{out}}{(v_{loc}, v_{adj}) \rightarrow_A (v_{loc}, v_{adj}[p.reqRead \mapsto \neg p.reqRead])} \text{R-OUTPUTSTATUS} \\
\\
\frac{\{g \rightarrow \{s\}\} \in Trans_G \text{ is internal}}{(v_{loc}, v_{adj}) \rightarrow_A s(v_{loc}, v_{adj})} \text{R-INTERNAL} \\
\\
\frac{\{g \rightarrow S\} \in Trans_G \text{ is external, } \{s_1, \dots, s_n\} \text{ are the assignments in } S}{(v_{loc}, v_{adj}) \rightarrow_A s_n \circ \dots \circ s_1(v_{loc}, v_{adj})} \text{R-EXTERNAL}
\end{array}$$

The first three rules describe the potential change of context, i.e. the adjoint variables. R-InputStatus and R-OutputStatus shows that the reading status of an output port and status of an input port may changed randomly. And R-InputValue shows that the value of an input port may be updated by the context.

The rule R-Internal models the internal transitions in $Trans_G$. As illustrated previously, an internal transition doesn't contains any perform statement. So its canonical form comprises only one assignment s . Firing such a transition will simply apply s to the current configuration.

Meanwhile, R-External models the external transitions, where the automaton need to interact with its context. Fortunately, since all the context change are captured by the first three rules, we can simply regard the context as a set of local variables. Consequently, the only difference between an internal transition and an external transitions is that the later may contains multiple assignments.

5 Discussion

6 Case Study

7 Conclusion and Future Work

References

1. Abdulla, P., Deneux, J., Stålmarch, G., Ågren, H., Åkerlund, O.: Designing safe, reliable systems using SCADE. In: Tiziana, M., Bernhard, S. (eds.) Proceedings of ISoLA 2004. LNCS, vol. 4313, pp. 115–129. Springer (2006)
2. Alur, R., Dill, D.L.: A theory of timed automata. *Theoretical Computer Science* 126(2), 183–235 (1994)
3. Arbab, F.: Reo: a channel-based coordination model for component composition. *Mathematical Structures in Computer Science* 14(3), 329–366 (2004)
4. Berry, G., Gonthier, G.: The Esterel synchronous programming language: design, semantics, implementation. *Science of Computer Programming* 19(2), 87–152 (1992)
5. Hoare, C.A.R.: *Communicating Sequential Processes*. Prentice-Hall (1985)
6. Kwiatkowska, M., Norman, G., Parker, D.: PRISM 4.0: Verification of Probabilistic Real-Time Systems. In: Gopalakrishnan, G., Qadeer, S. (eds.) Proceedings of CAV 2011. LNCS, vol. 6806, pp. 1–6. Springer (2011)