Transformation of PROMELA to Channel Systems

Sheila Nurul Huda
Jurusan Teknik Informatika, Fakultas Teknologi Industri
Universitas Islam Indonesia
Jl. Kaliurang Km. 14 Yogyakarta, Indonesia, 55501
Tel +62274895287 ext.122, Faks. +62274895007 ext.148
sheilla@staff.uii.ac.id

Reza Pulungan
Jurusan Ilmu Komputer dan Elektronika, Fakultas Matematika
dan Ilmu Pengetahuan Alam, Universitas Gadjah Mada
FMIPA Gedung Selatan, Sekip Unit III
Kotak Pos BLS, 21 Yogyakarta, Indonesia, 55281
Tel/Fax +62274546194
pulungan@ugm.ac.id

Abstract—This paper reports on an implementation of transformation of PROMELA models into Channel Systems, which will be further transformed into Labeled Transition Systems (LTSs). The objective of this transformation is to obtain a formal semantics for further model checking purposes. A Channel System is a way to describe communicating processes that run in parallel, where each process is represented by a Program Graph (PG). The main part of a Program Graph is a location transition which consists of the initial location, a guard which determines whether the transition is executable or not, an action that will be executed in the location transition, and the next location. This paper defined the location transition for PROMELA constructs such as assignments, communication actions, if-fi, do-od, and atomic steps.

Keywords—PROMELA, Channel Systems, Program Graphs, Model Checking.

I. INTRODUCTION

PROMELA is a high-level language to describe a model of a system for model checking purpose. Systems that have been modeled in PROMELA can be verified formally using SPIN model checker. Building an abstraction/model of a system is the first key to a successful model checking. With an easy to understand language and ability to model concurrent systems, PROMELA and SPIN have reached an important position in the model-checking world.

System modeled in PROMELA can be verified using SPIN so that the correctness is guaranteed. After it is verified, the modeled system has to be implemented to become a real application. The transformation from PROMELA model to application, all this time, is performed by human by manually programming it in some high-level programming language. This process is prone to errors and mistakes. It is much more convenient and preferable if a verified PROMELA model can be automatically transformed into an application, e.g. application in Java language.

Transformation from modeling language to implementation language is a big endeavour, and has to be done carefully; this process can be viewed in Fig. 1. The transformation itself has to be verified too; which means that it has to maintain the overall correctness. To achieve the correctness of the transformation, the semantics of the PROMELA model and Java program are needed. We have to make sure that both semantics are equivalent. This can be done by transforming the PROMELA model to a Labeled Transition System (LTS) and, similarly, the Java program to another LTS, then proving that there is an equivalence notion relating both LTSs. LTSs are a universal structure and have a clear semantics. The problem is that a PROMELA model cannot be translated into an LTS as is, because the PROMELA model may contain communication actions among its concurrent processes. One way to transform PROMELA to LTS is via Channel System. Channel System is a formalism that supports communications among concurrent processes via message passing and shared memory. From the resulting Channel System, we can then expand their constituent Program Graphs into an LTS.

For that purpose, it is necessary to transform PROMELA into Channel Systems, as a preliminary step in a work to achieve correct applications from verified models. This paper deals with the transformation of PROMELA to Channel Systems, which is defined in the style of operational semantic rules. There are several paper describing works on operational semantics of PROMELA, e.g. Natarajan and Holzmann into symbolic LTSs [1], where basic operational semantics are provided with symbolic notations. Weise [2] uses Structured Operational Semantics (SOS) approach, where each step is divided into three incremental stages, i.e. declaration, single processes, and parallel processes. Bevier works in transforming PROMELA into ACL2 logic [3]; ACL2 logic is a theorem-proving logic for integrating model checking and theorem prover methods. Del Mar Galardo uses generalized operational semantics [4] for abstract model checking. Channel System itself is formalized by Baier and Katoen [5] and it is used as semantic base for nanoPROMELA language (a derivative and simpler version of PROMELA).
We denote Program Graph PG of set Var of typed variables as a tuple (Loc, Act, Effect, \( \mapsto \), Loc\(_0\), g\(_0\)), where:

- Loc is a finite set of location.
- Act is a set of action.
- Effect: Act \( \times \) Eval(Var) \( \rightarrow \) Eval(Var); is the effect function. For example, if \( \alpha \) is \( x=x+y \) then, Effect(\( \alpha \); \( \{ x=1; y=7 \} \)) = \( \{ x=8; y=7 \} \). Effect describes the result in variable value, which changes when an action \( \alpha \) is executed.
- \( \mapsto \subseteq \) Loc \( \times \) Cond(Var) \( \times \) Act \( \times \) Loc, is the conditional transition relation. Cond(Var) is a boolean condition defined on the set of variables. If the condition holds true then action \( \alpha \) will be executed and then there will be transition to next location.
- Loc\(_0\) \( \subseteq \) Loc is a set of initial locations.
- g\(_0\) \( \subseteq \) Cond(Var) is the initial condition.

A. Channel Systems

Baier dan Katoen [5] describes that inside a system, a process will communicate with other processes using channels. Channel System is a way to describe communications among concurrent processes inside a system. A Channel System CS over (Var, Chan) consists of set Var of typed variables and a set Chan of channels, and atomic regions that avoid undesired interleavings. Guarded commands are used as labels for the edges of Program Graphs and Channels System.

B. Formal Definitions

1. Program Graph

According to Baier and Katoen [5] Program Graph PG over set Var of typed variables is a tuple (Loc, Act, Effect, \( \mapsto \) , Loc\(_0\), g\(_0\)), where:

- Loc is a finite set of location.
- Act is a set of action.
- Effect: Act \( \times \) Eval(Var) \( \rightarrow \) Eval(Var); is the effect function. For example, if \( \alpha \) is \( x=x+y \) then, Effect(\( \alpha \); \( \{ x=1; y=7 \} \)) = \( \{ x=8; y=7 \} \). Effect describes the result in variable value, which changes when an action \( \alpha \) is executed.
- \( \mapsto \subseteq \) Loc \( \times \) Cond(Var) \( \times \) Act \( \times \) Loc, is the conditional transition relation. Cond(Var) is a boolean condition defined on the set of variables. If the condition holds true then action \( \alpha \) will be executed and then there will be transition to next location.
- Loc\(_0\) \( \subseteq \) Loc is a set of initial locations.
- g\(_0\) \( \subseteq \) Cond(Var) is the initial condition.

A. Channel Systems

Baier dan Katoen [5] describes that inside a system, a process will communicate with other processes using channels. Channel System is a way to describe communications among concurrent processes inside a system. A Channel System CS over (Var, Chan) consists of set Var of typed variables and a set Chan of channels, and atomic regions that avoid undesired interleavings. Guarded commands are used as labels for the edges of Program Graphs and Channel System.

C. PROMELA

PROMELA is an input language for SPIN model checker. It is used to specify or model a system for model checking purpose. PROMELA is an acronym for Process Meta-Language. The use of the term meta is significant in this context. As we shall see, abstraction is often a key to successful verifications. The specification language is intended to make it easier to find good abstractions for systems designs. PROMELA is not meant to be an implementation language but a systems description language. To make this possible, the emphasis in the language is on the modeling of process synchronization and coordination, and not on computation. The language is also targeted to the description of concurrent software systems, rather than the description of hardware circuits (which is more common for model checking applications) [6].

A PROMELA model P consists of a finite number of processes P\(_1\),..., P\(_n\) to be executed concurrently.

PROMELA supports communications over shared variables and message passing along either synchronous or buffered FIFO-channels. The stepwise behaviour of the processes Pi is specified in PROMELA using a guarded command language [7, 8] with several features of classical imperative programming languages (such as variable assignment, conditional and repetitive commands, sequential composition), communication actions where processes may send and receive messages from the channels, and atomic regions that avoid undesired interleavings. Guarded commands are used as labels for the edges of Program Graphs and Channel Systems. They consists of a condition (guard) and an action. PROMELA does not use action names, but specifies the effect of actions by statements of the guarded command language [5].

The main ingredients of the statements that formalize the stepwise behaviour of the processes Pi are the atomic commands skip, variable assignment \( x = expr \), communication actions c?x (reading a value for variable x from channel c) and c!expr (sending the current value of an expression over channel c), conditional commands (if-then-else) and (while) loops. Instead of the standard if-then-else constructs or while-loops, PROMELA supports nondeterministic choices and allows specifying a finite number of guarded commands in conditional and repetitive commands. Syntax of PROMELA statement can be viewed in Fig. 2.

- stmt = skip \[ x = expr \] \[ c?x \] \[ c!expr \]
- stmt\(_1\); stmt\(_2\) | atomic{assignments}
- if \( g_0 \Rightarrow stmt_1 \) \( ... \) \( g_n \Rightarrow stmt_n \) \( \text{fi} \)
- do \( g_1 \Rightarrow stmt_1 \) \( ... \) \( g_n \Rightarrow stmt_n \) \( \text{od} \)

Fig. 2. Syntax of PROMELA statements.
III. SEMANTICS OF PROMELA

The operational semantics of a PROMELA statement with variables and channels from (Var, Chan) is given by a Program Graph over (Var, Chan). The Program Graphs PG₁, ..., PGₙ for the processes P₁, ..., Pₙ of a PROMELA model P = [P₁ | ... | Pₙ] constitute a Channel System over (Var, Chan) [5].

The Program Graph associated with a PROMELA statement stmt formalizes the control flow when executing stmt. That is, the sub-statements play the role of the locations. For modeling termination, a special location exit is used. The set of sub-statements of PROMELA statement stmt is recursively defined. For statement stmt ∈ {skip, assignment, communication action} the set of sub-statements is sub(stmt) = [stmt, exit]. For sequential composition let:

\[
\text{sub}(\text{stmt}_1; \text{stmt}_2) = \{\text{stmt'} \mid \text{stmt'} \in \text{sub}(\text{stmt}_1) \cup \text{sub}(\text{stmt}_2)\}
\]

For conditional commands, the set of sub-statements is defined as the set consisting of the if-fi statement itself and sub-statements of its guarded commands. That is, for if :: g₁ → stmt₁; ...; gₙ → stmtₙ if we have:

\[
\text{sub}(\text{cond_cmd}) = \{\text{cond_cmd} \} \cup \bigcup_{i=1}^{n} \text{sub}(\text{stmt}_i).
\]

The sub-statements of a loop given by do :: g₁ → stmt₁; ...; gₙ → stmtₙ od is defined similarly, but taking into account that control moves back to loop when guarded commands terminate. That is:

\[
\text{sub}(\text{loop}) = \{\text{loop, exit} \} \cup \bigcup_{i=1}^{n} \text{sub}(\text{stmt}_i) \setminus \{\text{exit}\}.
\]

For atomic regions, we let sub(atomic(stmt)) = {atomic{stmt}, exit}.

Baier and Katoen [5] also provide the inference rules for PROMELA constructs. The inference rules for atomic commands (skip, assignment, communication action) and sequential composition, conditional and repetitive commands give rise to the edges of a large Program Graph where the set of locations agrees with the set of PROMELA statements. Thus, the edges have the form:

\[
\text{stmt} \xrightarrow{g} \text{stmt} \text{ or stmt} \xrightarrow{\text{comm}} \text{stmt}'.
\]

where stmt is a PROMELA statement, stmt' a sub-statement of stmt, and g a guard, u an action, and comm a communication action c!x or c?x. The subgraph consisting of the sub-statements of stmt then yields the Program Graph PGₜ of process Pₜ as a component of the program P.

The semantics of the atomic statement skip is given by a single axiom formalizing that the execution of skip terminates in one step without affecting the variables:

\[
\begin{align*}
\text{skip} & \xrightarrow{\text{true, guard}} \text{exit}
\end{align*}
\]

where id denotes an action that does not change the values of the variables, i.e., Effect(id, η) = η for all variable evaluations η. Similarly, the execution of a statement consisting of an assignment x = expr has the trivial guard and terminates in one step:

\[
\begin{align*}
\text{x: expr} & \xrightarrow{\text{true, assign(x, expr)}} \text{exit}
\end{align*}
\]

where assign(x, expr) denotes the action that changes the value of x according to the assignment x = expr and does not affect the other variables. For the communication actions c!expr and c?x, the following axioms apply:

\[
\begin{align*}
\text{c!expr} & \xrightarrow{\text{true}} \text{exit} \\
\text{c?x} & \xrightarrow{\text{true}} \text{exit}
\end{align*}
\]

The effect of an atomic region atomic{x₁ = expr₁; ...; xₘ = exprₘ} is the cumulative effect of the assignments xᵢ = exprᵢ. It can be defined by the rule:

\[
\begin{align*}
\text{atomic{x₁ = expr₁; ...; xₘ = exprₘ}} & \xrightarrow{\text{true, #=}} \text{exit}
\end{align*}
\]

where a₀ = id, aᵢ = Effect(assign(xᵢ, exprᵢ), Effect(aᵢ₋₁, η)) for 1 ≤ i ≤ m.

Sequential composition stmt₁; stmt₂ is defined by two rules that distinguish whether or not stmt₁ terminates in one step. If the first step of stmt₁ leads to a location (statement) different from exit, then the following rule applies:

\[
\begin{align*}
\text{stmt₁} & \xrightarrow{g} \text{stmt₁'} \neq \text{exit} \\
\text{stmt₁; stmt₂} & \xrightarrow{g} \text{stmt₁'; stmt₂}
\end{align*}
\]

If, however, the computation of stmt₁ terminates in one step by executing action α, then control of stmt₁; stmt₂ moves to stmt₂ after executing α:

\[
\begin{align*}
\text{stmt₁} & \xrightarrow{g} \text{exit} \\
\text{stmt₁; stmt₂} & \xrightarrow{g} \text{stmt₂}
\end{align*}
\]

The effect of a conditional command cond_cmd = if :: g₁ → stmt₁; ...; gₙ → stmtₙ fi is formalized by means of the following rule:

\[
\begin{align*}
\text{if :: g₁ → stmt₁; ...; gₙ → stmtₙ fi} & \xrightarrow{\text{gᵢ, id}} \text{stmtᵢ}
\end{align*}
\]

This rule relies on the two step semantics, where the selection of an enabled guarded command and the execution of its first action are split into two steps.

For loops, say loop = do :: g₁ → stmt₁; ...; gₙ → stmtₙ od we deal with two rules. The rules also rely on the two step semantics. The first rule is similar to the rule for conditional commands, but taking into account that control moves back to loop after the execution of the selected guarded command has been completed. This correspond to the following rule:
If the action is break, then the do-od loop will be aborted. This is formalized by the following axiom:

\[
\text{break} \rightarrow \text{exit}
\]

IV. DESIGN AND IMPLEMENTATION

The transformator tool has a PROMELA file with .pml extension as an input file. The content of the file has to be recognized so that the transformation can be done to Program Graph. To recognize the content, we need to build lexer and parser. Lexer works to recognize the input file into a token. Then a parser will recognize the meaning (semantics) of a token. In programming, recognizing the words is done via a method call, which represents the structure of the words, so that the sequence of method calls shows the sequence of input words structures [9]. The meaningful token then will be transformed into part of Channel Systems. Parts of the code that will be transformed into Channel System include global variable declaration, proctype declaration, local variable declaration, and statement inside proctype. Java codes are added to recognize variable, change variable object attribute, change Program Graph and state object attribute, and build internal data structures that will be used to generate Channel Systems. Fig. 3 shows the system workflow of the transformation tool.

A. Output File Design

The general form of Channel Systems that will be built is as shown in Fig. 4. Channel Systems consist of CS declaration, global variable declaration, active Program Graph declaration, and prototype of all Program Graphs (active and inactive). Inside the prototype of Program Graphs, there may be local variable declaration, and transition locations, which consists of initial location, guard, action, and next location. Atomic structure is wrapped in keyword atomic begin-atomic end. This keyword means to mark that all transition locations inside the keyword is done within a single state transition, which means that no interleaving is allowed.

```
Channel System name_CS
over var, chan
...
CS = [PG0 || PG1]
Begin
PG0 instance of p()
over var, chan
...
Begin
loc_init [guard] [action] loc_next;
...
atomic begin
loc_init [guard] [action] loc_next
atomic end
```

Fig. 4. Output File Design

B. PROMELA Grammar

PROMELA file is parsed using ANTLR (Another Tool for Language Recognition) [10]. First, the grammar of PROMELA has to be specified. The grammar used in this system is based on PROMELA grammar of Holzmann [11] with proper changes, as shown in Fig. 5 and Fig. 6.
grammar promela;
spec : module+ EOF ;
module : proctype | init | mtype | decl_lst;
proctype : ('active' | 'run' | 'init')? 'ifoption' 'fi' | 'proctype' name '(' (arg_lst)? ')' | 'proctype' name '(' (arg_lst)? ')';
init : 'init' '(' sequence ')' | 'proctype' name '(' (arg_lst)? ')';
ifoption : ':' ':' (guard 'ifoption' 'fi'   |  'do' dooption 'od'  | 'else';
dooption : ':' ':'  (guard 'ifoption' 'fi'   |  'do' dooption 'od'  | 'else');
sequence : step ( separator step )* (';'? )? ;
step : one_decl | stmnt ;
assignx :  varref ('=' any_expr  | '+' '+=' | '-=' )? ';';
constant : 'true' | 'false' | 'skip' | NUMBER;
MOD : '%';
NUMBER : '0'..'9'+;
ID : ('a'..'z' |'A'..'Z' | '_' )+ ;
WS  :   (' '|'	'|'\n'|'\r')? ;
name: ID (ID | NUMBER)*;
sequence: step ( separator step )* (';')? ;
sep : ';'  |  ':' | '@' ;
module : proctype | init | mtype | decl_lst;
Keyword run is considered as an action so it follows assignment semantics.

The second example is if-fi construct in PROMELA as shown in Fig. 9.

![Fig. 9. if-fi Construct](image)

There are two branches from the if statement, one with the keyword a == 0, other if a is not 0 (else keyword). According to if semantics, we use two step semantics:

$$
\text{if } a = 0 \rightarrow b++; \\
\text{else } b--; a = a+b; \\
\text{fi}
$$

When making transition to check the guard, it will execute id (action without effect), then it will continue with transition for executing command after the guard (i.e. b++ for a == 0 guard, sequential composition for else guard).

Here we can see that the sequential composition with the computation of first statement terminates in one step, so it must follow the semantics:

$$
\text{if } a = 0 \rightarrow b++; \\
\text{else } b--; a = a+b; \\
\text{fi}
$$

The transformation result is depicted in Fig. 10 and it complies with the semantics.

![Fig. 10. Transformation Output](image)

For nondeterministic if-fi, the guard will be trivial, all true, so it will not check the guard and terminates in one step (not two step). The third PROMELA example shown in Fig. 11 is nondeterministic if-fi.

![Fig. 11. Nondeterministic if-fi](image)

The transformation result will shown two branches from loc1 and there is no guard checking, as shown in Fig. 12.

![Fig. 12. Transformation Output](image)

The fourth example is do-od construct as shown in Fig. 13.

![Fig. 13. do-od Construct](image)

Execution of a++ will be continued as long as a<5 holds. If not, it will execute break. The loop uses two step semantics, so first it has to check the guard, and then executes the statement after the guard. After executing the statement, it will back to loop. This semantics can be written as:

$$
\text{do } a < 5 \rightarrow a++; \\
\text{else } \text{break;}
$$

Break is a keyword for terminating the loop, so when it encounters break, it will go to location exit, as in semantics:

$$
\text{break} \rightarrow \text{exit}
$$

The transformation result complies with the semantics, as shown in Fig. 14.
The validation is also done using other constructs and has shown proper results. More complex examples are also used to make sure all constructs will make a proper composition, such as the Szymanski mutual exclusion protocol [12] is modeled in PROMELA then transformed into Channel Systems. The transformation result shows location transitions comply with the semantics.

VI. CONCLUDING REMARKS

The PROMELA language has been successfully transformed into Channel Systems. Channel Systems as an operational semantics of PROMELA have a clear semantics and can be implemented as above. All results show that every construct used in the example is transformed well and complied with the semantics. The more complex example is not shown here due to page limitation, but has been checked and gives proper result that follows with the semantics. As the Channel System can be obtained from PROMELA language, the work can continue to form Labeled Transition Systems (LTSS) which is the final structure that is intended to achieve.

REFERENCES


[9] P. Sestoft, Grammars and Parsing with JAVA, Department of Mathematics and Physics Royal Veterinary and Agricultural University, Denmark, 1999.

