INTEGRATION OF AUTOMATIC THEOREM PROVERS IN EVENT-B PATTERNS

Eman Elsayed¹, Gaber El-Sharawy² and Enas El-Sharawy³

¹²³Faculty of Science(girls), Al-Azhar University, Cairo, Egypt
emankaram10@azhar.ed.eg, gelsharawy@gmail.com, dr_enas_idm@hotmail.com

ABSTRACT
Event-B is a formal method for the system level modeling and analysis of dependable applications. It is supported by an open and extendable Eclipse-based tool set called Rodin. In this paper we proposed using Automatic theorem provers known as SMT-solvers with event-B pattern. The benefits of that are to reduce the proving effort, to reuse a model and to increase the degree of automation. The proposed approach has been applied successfully on two different case studies.

KEYWORDS
Formal methods, Event-B, Design patterns, SMT-Solver, Refinement.

1. INTRODUCTION
Nowadays, there are many formal methods used in various domains for constructing models of complex system together with a several advanced theories and tools. But more experiments in this area are still needed to be carried out to provide significant evidence for convincing and encouraging other users to benefit from those theories and tools, and make formal methods more accessible to software industries. However, there is no systematic approach for refinement model with formal methods so we design pattern. Since refinement model is monotonic, the final resulting model shall be a refinement of the original model. Practically, but the last refinement model of the pattern’s refinement-chain is incorporated in the development. This feature allows us to reuse formal models more flexibly.

Also when used formal methods as event-b, produces large quantities of proof obligations (Pos). For each such PO can be valid, but cannot be automatically proved because the proof system lacks some axioms, as the specification logic is incomplete. In that case, it is sometimes possible to patch the prover with additional rules so that it finds a proof for the verification condition.

The goal of the work presented in this paper is to study the possibility of enhancing event-b patterns. We want to bring the Satisfiability Modulo Theory (SMT) into formal methods and in particular SMT-solvers tools into event-B patterns. It considers the POs generated in the Rodin Platform and using SMT-solvers and Pattern design tools. This approach should also be readily applied in other development environments.

For our purposes we apply the proposed approach on two case studies of a sequential algorithms first binary search algorithm. The second case study is finding the minimum value of an array of natural numbers. We design pattern and prove the correctness of binary search algorithm automatically discharges a part of the proof obligations and also provides the interactive prover to discharge the remaining proofs. The rest of the paper is organized as follows. The next section presents the main concepts then section three introduces the literature review. The proposal
The approach is described in section four. The approach is applied on two case studies in section five and the result analyzed in section six. Conclusions and future work are discussed in section seven. Finally, list the references.

2. MAIN CONCEPTS:

This section presents the background material for the paper. We start by overviewing the Event-B formalism. Then we describe the RODIN platform tool.

2.1. Event-B and Rodin

The B Method in reference [3] is a formal approach for the specification and rigorous development of highly dependable software. The method has been successfully used in the development of several complex real-life applications [5]. Event-B [10] is a formal framework derived from the B Method to model and reason about parallel, distributed and reactive systems. Event-B has the associated RODIN platform [4,5,6], which provides automated tool support for modeling and verification by theorem proving. Then Event-B is an extension of the B-method for specifying and reasoning about complex systems including concurrent and reactive systems. An Event-B model is described in terms of contexts and machines as shown in Figure 1.

Where, contexts contain the static parts of a model [1]. Each context may consist of carrier sets and constants as well as axioms which are used to describe the properties of those sets and constants. Contexts may contain theorems for which it must be proved that they follow from the preceding axioms and theorems. Moreover, contexts can be extended by other contexts and seen by more than one machine.

Machines contain the dynamic parts of an Event-B model [1]. This part is used to provide behavioral properties of the model. A machine is made of a state, which is defined by means of variables, invariants, events and theorems. A machine contains a number of atomic events which show the way that the machine may evolve. Each event is composed of three elements: an event name, guard(s) and action(s). The guard is the necessary condition for the event. The actions determine the way in which the state variables are going to evolve when performing the event [1]. In addition, machines can be refined by other machines, but each machine can refine at most one machine.

Machine refinement provides a means to introduce more details about the dynamic properties of a model [2]. For more on the well-known theory of refinement, we refer to the Action System formalism that has inspired the development of Event-B [5]. We present some important proof obligations for machine refinement. As mentioned before, the user of Event-B is not presented with a behavioral model but only with proof obligations. The proof obligations describe the semantics of Event-B models.

The Rodin platform is an open and extensible tool for Event-B specification and verification [3,18]. It contains a database of modeling elements used for constructing system models such as
variables, invariants and events. It is accompanied by various useful plug-ins such as a proof-obligation generator, provers, model-checkers, UML transformers, etc.

2.2 Design Pattern

The intention of Design Patterns in Event-B is to have a methodological approach to reuse former developments (referred to as patterns) in a new development. The patterns approach is a promising avenue to let inexperienced designers build conceptual models and as a tool for building domain models [21]. The proofs of the pattern can be reused too. It was shown that for the special case of a model that does not see any context and its events do not have any parameters, the generation of a refinement of the problem at hand is correct by construction and no proof obligation needs to be generated again. The correctness of the construction relies on a correct matching of the pattern with the problem at hand. We summarize the incorporating patterns into Event-B developments:

First of all, in our notion, a pattern is just a development in Event-B including specification p0 and a refinement P1. In p0 we introduce a variable r_pat with event final_pat. P1 will be like the second refinement that in the section 5.1.3 but by another naming for each component in Event-B. As a first step, all the variables of the pattern specification have to be matched with variables of the problem. Furthermore all the events of the pattern have to be matched with a event of the problem. The chosen matching is valid for the construction if the following checks turn out to be true. Note that this is only formally proved for models that do not see any context or having events with parameters. The proof of the correctness of the construction with general models is pending.

- Check that the guards of each event of the pattern specification are syntactically the same as the guards of the corresponding event in the problem.
- Check that the actions of each event of the pattern specification are syntactically the same as the actions of the corresponding event in the problem.
- Check that no event in the problem that is not matched alters a matched variable.

Once all checks are done, the refinement of the problem is generated by merging the pattern refinement with the problem.

2.3 Satisfiability Modulo Theory (SMT)

The Satisfiability Modulo Theory (SMT) problem is a decision problem to determine if a given logic formula is satisfiable with respect to a combination of theories expressed in first-order logic. Theories of interest for the work described in this paper include uninterpreted functions with equality and integer arithmetics. Since the validity of a proof obligation can be decided by checking the unsatisfiability of its negation, SMT-solvers are natural candidates for discharging the verification conditions generated in the application of Event-B methods. SMT-solvers can for example handle a formula like

\[ x \leq y \land y \leq x + f(x) \land P(h(x) - h(y)) \land \neg P(0) \land f(x) = 0 \]

Which contains linear arithmetic on real’s (0,+,−,≤), and un-interpreted symbols (P,h,f) . SMT-solvers use decision procedures for the disjoint languages (for instance, congruence closure for un-interpreted symbols, and simplex for linear arithmetic) and combine them to build a decision procedure for the union of the languages. There are a number of available SMT-solvers as in references [8,9,15,12]. They differ mostly by the theories they handle, their efficiency (which
may vary depending on the theories they handle). The ability to handle quantified formulas as well as to construct certificates (proofs) of their results.

3. LITERATURE REVIEW

Event-B is a language for the formal development of reactive systems. At present the RODIN toolkit [5] for Event-B is used for modeling requirements, specifying refinements and verification. The design pattern is not a new idea and is not restricted to the field of computer science. Design patterns were introduced by Christopher Alexander in the field of architecture. In 1977 he spoke of patterns as, “each pattern describes a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way that you can use this solution a million times over, without ever doing it the same way twice”[AIS+77]. Event-b pattern[17] approach has been applied to formalize communication protocols from SAP. The examples are Buyer/Seller B2B as described in [20] and Ordering/SupplyChain A2A Communications as described in [14, Section 5.3.3].

The proof statistics compare with the developments without patterns and with patterns for the two case studies. More importantly, this approach saves on average of the two case studies 33% of the manual proofs). When applying SMT-solvers on the European project Deploy by the developments made publicly available Web site (http://www.deploy-project.eu/) , all the proof obligations generated by the developments was 2359 proof obligations, and it terminated successfully in 54% of the cases. When applying the SMT-Solver approach, the ratio of discharged proof obligations jumps to 63%. It is important to note that all the runs of the SMT-solver take a negligible amount of time. The remaining proof obligations shall be communicated to the community of researchers on SMT-solving techniques in the hope that use them as a direction for future extensions and improvements [13].

4. THE PROPOSAL APPROACH

The proposed approach is classified into two main phases:

First phase, design the event-B pattern to the final refinement model. That is for guarantee the reusability in any similar models.

Second phase, apply Automatic theorem provers known as SMT-solvers with event-B pattern. That is for guarantee reducing the proving effort and increasing the degree of automation.

The proposed approach has been applied successfully on two case studies. First is a binary search model and the second is Minimum of array model. Each model is created in Rodin platform application once completely and by using design pattern approach.

5. CASE STUDIES

5.1 Overview of a Binary Search:

A binary search or half-interval search algorithm finds the position of a specified value (the input "key") within a sorted array. The algorithm compares the input key value with the key value of the middle element of the array. If the keys match then a matching element has been found so its index, or position, is returned. Otherwise, if the sought key is fewer than the middle element's key, then the algorithm repeats its action on the sub-array to the left of the middle element or, if the input key is greater, on the sub-array to the right. If the remaining array to be searched is reduced to zero, then the key cannot be found in the array and a special "Not found" indication is returned. A binary search halves the number of items to check with iteration so locating an item
(or determining its absence) takes logarithmic time. A binary search is a dichotomist divide and conquer search algorithm.

5.1.1 Initial Model

In this abstraction, we begin with an abstract model of binary search by given (Pre-condition) a natural number \( n \in \mathbb{N} \), \( n > 0 \), a sorted array \( f \) of \( n \) elements built on a set \( \mathbb{N} \): \( f \in \mathbb{N}^n \), a value \( v \) known to be in the array: \( v \in \text{ran}(f) \). We are looking for (Post-condition) an index \( r \) in the domain of the array: \( r \in \text{dom}(f) \) such that \( f(r) = v \). Two events are defined called final and progress.

5.1.2 First Refinement

First refinement consist in introducing two new variables \( p \) and \( q \). Variables \( p \) and \( q \) are supposed to be two indices in the array \( f \) (inv1_1 and inv2). The variable \( r \) is within the interval \( p .. q \) (inv3). Moreover, the value \( v \) is supposed to be a member of the set denoting the image of the interval \( p .. q \) under \( f \): that is, \( f[p .. q] \) (inv4). Here is the state of this refinement:

Now, we introduce two events called inc and dec which split abstract anticipated event progress. These events are convergent (see variant1 above). They are increment or decrement \( p \) or \( q \) when \( f(r) \) is smaller or greater than \( v \). They also move \( r \) non deterministically within the new interval \( p .. q \): The total events of first refinement

5.1.3 Second Refinement

At the previous stage, Increment and decrement were non-deterministic and \( r \) was chosen arbitrarily within the interval \( p .. q \). We now remove the non-determinacy in Increment and decrement and \( r \) is chosen to be the middle of the interval \( p .. q \).
5.1.4 Pattern Design

First of all, in our notion, a pattern is just a development in Event-B including specification p0 and a refinement p1

5.1.4.1 Development specification P0

Pattern design specification like initial model in the section 5.1.1

5.1.4.2 Pattern refinement P1

Pattern refinement like first and second refinement in the section 5.1.2 and 5.1.3

5.1.4.3 Applying Pattern

We have implemented our prototype for supporting our approach as a plug-in for the RODIN Platform [16] which is an open source platform based on Eclipse. The plug-in provides a wizard taking users through different steps of applying patterns, namely, matching, syntax checking,
renaming and incorporating. Fig. 2 is a screen-shot of the wizard page for create new event-b pattern. Fig. 3 is a screen-shot of the wizard page for matching step between the problem and the specification. This step includes a dialog for the developers to choose the matching between variables and events. Fig. 4 is a screen-shot of the wizard page for Syntax checking of the matching provided by the user in the previous steps. Fig. 5 is a screen-shot of the wizard page for renaming pattern of variables and events. Fig 6 is a screen-shot of the wizard page for incorporating the refinement of the pattern into the development.
5.1.5 Apply SMT-Solver

The Rodin platform [11] is a development environment for Event-B. It is based on Eclipse [19] and can be extended by developing new plug-ins in the Java programming language. Automatic verification of proof obligations using SMT-solvers. The plug-in communicates with the SMT-solvers using files and operating system commands. The configuration of the plug-in includes a choice of SMT-solvers: Alt-Ergo [7], cvc3, veriT [9] and Z3 [12] at this time (see Fig. 7). It is now available to the formal methods community as an exploratory package through Rodin’s official source code repository. Currently, the verification with the SMT-solver has to be activated by the user by clicking a button (see Fig. 7, left part). Whenever the verification is successful (see Fig. 7, right part), the status of the proof obligation is updated and the user may move to the next proof obligation.

Fig. 6 Fifth step. Renaming

Fig. 7. Screen shot of the proof. On the right, before the proof, the button is active and the status is not proved. On the left, after a successful proof, the button has been disactivated and the status is “proved”.

5.2 Overview of Minimum Model

Minimum model consists in looking for the minimum of the range of a non-empty array of natural numbers.

5.2.1 Initial Model

Let \( n \) and \( f \) be two constants, and \( m \) a variable. Here is our initial model:

Constants: \( n, f \)

Variables: \( m \)

Inv1: \( m \in \mathbb{N} \)

Init

BEGIN

M \in \mathbb{N}

END

minimum

BEGIN

act1 : m \in \min(\text{ran}(f))

END

We also introduce two new events inc and dec. When \( p \) is smaller than \( q \) and \( f(p) \) is greater than \( f(q) \), we can reduce the interval \( p \ldots q \) to \( p + 1 \ldots q \) since \( f(p) \) is certainly not the minimum we are looking for. We have a similar effect with invariant dec. The minimum is then found when \( p \) is equal to \( q \) according to invariant inv1_4:

Init

\[
\begin{align*}
\ m & := 0 \\
\ p & := 1
\end{align*}
\]

Increment

When

p<q

f(p)>f(q)

then

p := p+1

end

decrement

When

p\leq q

f(p)>f(q)

then

q := q-1

end

5.2.3 Pattern Design

First of all, in our notion, a pattern is just a development in Event-B including specification \( p0 \) and a refinement \( p1 \)

5.2.3.1 Development Specification P0

Pattern design specification like initial model in the section 5.2.1
5.2.3.2 Pattern refinement P1

Pattern refinement like first and first refinement in the section 5.2.2

Pattern design steps and applying SMT-solver of this model will be like that done in section 5.1.4.3 and 5.1.5.

6. RESULT ANALYSIS

In section 5.1 we create the first case study that is a binary search model in Rodin platform including initial model and the first and second refinement. Rodin produces proof obligations for each model. Some of them are done automatically and the others need to discharge manually by the user or by proof solver tools as SMT-solver. The following table 1 shows the Proof statistics of Binary search model.
When we create the binary search model in Rodin (section 5.1.1, 5.1.2, 5.1.3) we found 10 (27%) Pos produced automatically from total 37 Pos. The remaining Pos 27 need to discharge manually. Then we create the binary search model in Rodin by installing pattern plug-in (section 5.1.4). We found the required total proof 17 Pos instead of 37 Pos without pattern using that is mean the pattern approach saves 20 (54%) Pos instead of 37 Pos. But when we created the binary search model in Rodin by installing SMT-Solver plug-in as in section 5.1.5, we found 10 Pos produced automatically and 20 Pos produced by SMT-Solver from total 37 Pos. The remaining Pos 7 need to discharge manually that is mean using SMT-Solver approach the proof raised to (81%) proof instead of (27%) that obtained by applying Rodin only. Finally when we applied our proposal on binary search model in Rodin by installing SMT-Solver and pattern plug-ins together we found that the total proved jumps to 61% Pos compare with 27% Pos that is obtained from applying Rodin only. Also the saved proof jumps to 64% of total Pos. The following chart present different among Pos on Rodin, SMT-Solver, Pattern and SMT with Pattern binary search model.

In section 5.2 we create the second case study that is a Minimum model in Rodin platform including initial model and the first refinement. Rodin produces proof obligation for each model. Some of them are done automatically and the others need to discharge manually by the user or by proof solver tools like SMT-solver. The following table 2 shows the Proof statistics of minimum search model.
Table 2: Proof statistics of Minimum model

<table>
<thead>
<tr>
<th>Pos Methods / POS</th>
<th>Total Pos</th>
<th>Auto</th>
<th>Interactive (SMT)</th>
<th>Proved Pos %</th>
<th>Saved Proof</th>
<th>Undercharged</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rodin</td>
<td>25</td>
<td>5</td>
<td>0</td>
<td>5 (2%)</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>SMT</td>
<td>25</td>
<td>5</td>
<td>17</td>
<td>22 (88%)</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Pattern</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1 (2%)</td>
<td>20 (80%)</td>
<td>4</td>
</tr>
<tr>
<td>Smt + Pattern</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3 (75%)</td>
<td>21 (84%)</td>
<td>1</td>
</tr>
</tbody>
</table>

When we create the minimum search model in Rodin (section 5.2.1, 5.2.2) we found 5 (2%) Pos produced automatically from total 25 Pos. The remaining Pos 20 need to discharge manually. Then we create the minimum search model in Rodin by installing pattern plug-in. We found the required total proof 5 Pos instead of 25 Pos without pattern using that is mean the pattern approach saves 20 (80%) Pos from 25 Pos. But when we create the minimum search model in Rodin by installing SMT-Solver plug-in, we found 5 Pos produced automatically and 17 Pos produced by SMT-Solver from total 25 Pos. The remaining Pos 3 need to discharge manually that is mean using SMT-Solver approach the proof raised to (88%) proof instead of (27%) that obtained by applying Rodin only. Finally when we applied our proposal on minimum search model in Rodin by installing SMT-Solver and pattern plug-ins together we found that the total proved jumps to 75% Pos compare with 2% Pos that is obtained from applying Rodin only. Also the saved proof jumps to 84% of total Pos. The following chart present different among Pos on Rodin, SMT-Solver, Pattern and SMT with Pattern of minimum search model.

![Chart 2: Pos of Minimum model](image)

7. CONCLUSION AND FURTHER WORK

In the previous section we analyzed statistically and compare with among applying SMT-Solver and pattern or do not apply. Then we conclude that the proof obligation reduced than using Rodin only. Also when we applied SMT-Solver on pattern model that help for reusability without proof obligation again (saved Pos). Finally, when we present the results of the two different case studies, we found proposal reduced the proof obligation successfully. The method in this paper tends to reduce the proving effort, to reuse a model and to increase the degree of automation.
As for future work, Proof obligations can be discharged in other development platforms, such as Atelier-B. Also other proof solvers may be used to improve discharging the proof obligations. Furthermore, we intend to implement the missing features from the pattern plug-in for the RODIN Platform, e.g. syntax checking and support for extracting information from the pattern refinement. Also, we are going to investigate more examples in other domains that could benefit from pattern approach. However, we also need to “instantiate” the context of the pattern development. In our examples so far, the contexts of the pattern and the problem are the same. Also, we shall like to use the patterns in a more general context. In the future, if this turns out to be too restrictive, we can choose to generate the corresponding proof obligations, again for more flexibility. Note that if a pattern matching can be syntactically checked successfully, the proof obligations generated should be trivial to be discharged. Finally, we can apply Pattern approach with another proof solver and compare with our proposed method.

REFERENCES


