Automated Debugging of SystemVerilog Assertions

Brian Keng1, Sean Safarpour2, Andreas Veneris1,3

Abstract—In the last decade, functional verification has become a major bottleneck in the design flow. To relieve this growing burden, assertion-based verification has gained popularity as a means to increase the quality and efficiency of verification. Although robust, the adoption of assertion-based verification poses new challenges to debugging due to presence of errors in the assertions. These unique challenges necessitate a departure from past automated circuit debugging techniques which are shown to be ineffective. In this work, we present a methodology, mutation model and additional techniques to debug errors in SystemVerilog assertions. The methodology uses the failing assertion, counter-example and mutation model to produce alternative properties that are verified against the design. These properties serve as a basis for possible corrections. They also provide insight into the design behavior and the failing assertion. Experimental results show that this process is effective in finding high quality alternative assertions for all empirical instances.

I. INTRODUCTION

Functional verification and debugging are the largest bottlenecks in the design cycle taking up to 46% of the total development time [1]. To cope with this bottleneck, new methods such as assertion-based verification (ABV) [2], [3] have been adopted by the industry to ease this growing burden. ABV in particular has shown to improve observability, reduce debug time as well as improve overall verification efficiency [2]. However even with the adoption of ABV, debugging remains an ongoing challenge taking up to 60% of the total verification time [1].

Modern ABV environments are typically composed of three main components: design, testbench and assertions. Due to the human factor inherent in the design process, it is equally likely for errors to be introduced into any one of these components. Commercial solutions [4]–[6] aim to help the debugging process by allowing manual navigation and visualization of these components. Most existing research in automated debugging [7]–[11] have focused primarily on locating errors within the design. The absence of automated debugging tools targeting testbenches and assertions remains a critical roadblock in further reducing the verification bottleneck.

The adoption of assertions introduces new challenges to the debugging process. Modern temporal assertion languages such as SystemVerilog assertions and Property Specification Language [12], [13] are foreign to most design engineers who are more familiar with RTL semantics. Temporal assertions concisely define behaviors across multiple cycles and execution threads, which creates a significant paradigm shift from RTL. For example debugging the failing SystemVerilog assertion $\text{req } |\Rightarrow \text{ gnt } \## [1:4] \text{ ack}$, requires the engineer to analyze four threads over five clock cycles to understand the failure. Moreover, a single temporal operator such as a non-consecutive repetition may map to a multiple line RTL implementation, adding to the debugging complexity. For these reasons, debugging complex temporal assertions remains one of the biggest challenges in their wide spread adoption.

Automated circuit debugging techniques have traditionally relied on localizing an error in a circuit. In a similar manner, it is possible to synthesize assertions [14]–[16] and allow one to apply similar circuit localization techniques to assertions. However, this proves ineffective in debugging assertions due to their compact nature, also shown later in the paper. For example, applying path-tracing [7] to the assertion $\text{valid } \#1 \text{ start } |\Rightarrow \text{ go}$, will return the entire assertion as potentially erroneous. Moreover, this type of localization does not provide help in directing the engineer towards correcting it. This suggests an urgent need for a departure from traditional circuit debugging techniques so that we can debug assertions effectively.

In this work, we propose a novel automated debugging methodology for SystemVerilog assertions (SVA) that takes a different approach. It aids debugging by generating a set of properties closely related to the original failing assertion that have been validated against the RTL. These properties serve as a basis for possible corrections to the failing assertion, providing an intuitive method for debugging and correction. They also provide insight into design behavior by being able to contrast their differences with the failing assertion.

In summary, our major contributions are as follows:

- We introduce a language independent methodology for debugging errors in assertions that produce a set of closely related verified properties to aid the debugging process. These properties are generated by an input set of modifications that mutate existing assertions.
- We propose a particular set of modifications for the SystemVerilog assertion language to mutate the failing assertion in order to generate closely related properties.
- We introduce two techniques dealing with vacuity and multiple errors to enhance the practical viability of this approach.

An extensive set of experimental results are presented on real designs with SystemVerilog assertions written from their specifications. They show that the proposed methodology and modification model are able to return high quality verified properties for all instances. In addition, the multiple error and vacuity techniques are able to filter out inessential properties by an average of 23% and 34% respectively.

The remaining sections of the paper are organized as follows. Section II provides background material. Our contributions are presented in Section III, Section IV and Section V. Section VI presents the experimental results and Section VII concludes this work.

II. PRELIMINARIES

This section gives a brief overview of the SystemVerilog assertions (SVA) language as well as concepts used extensively throughout this paper. For a more detailed treatment please refer to [12], [17]. SVA is a concise language for describing

---

1University of Toronto, ECE Department, Toronto, ON M5S 3G4 ({briank, veneris}@eecg.toronto.edu)
2Vennsa Technologies, Inc., Toronto, ON M5V 3B1 (sean@vennsa.com)
3University of Toronto, CS Department, Toronto, ON M5S 3G4

978-3-9810801-7-9/DATE11/©2011 EDAA
TABLE I
COMMON SVA OPERATORS

| seq_expr    | := | bool_expr |
| seq_expr    | := | time_range seq_expr |
| bool_expr   | := | bool_expr bool_abbrv |
| seq_expr    | := | seq_op seq_expr |
| bool_expr   | := | first_match (seq_expr) |
| seq_expr    | := | through seq_expr |

| time_range  | := | \#\#k | \#\#k1;k2 |
| bool_abbrv  | := | \[*k1\] | \[*k1;k2\] | \[#k\] | \[#k1;k2\] |
| seq_op      | := | and | intersect | or | within |
| prop_expr   | := | seq_expr |
| seq_expr    | := | \rightarrow property_expr |
| seq_expr    | := | property_expr |
| property_expr | := | property_expr or property_expr |
| property_expr and property_expr | := |

Example 1
Consider the specification: “If start is active, data repeats 2 times before stop and one cycle later stop should be low.” We can interpret this as the waveform given in Figure 1 with respect to clk. This can be concisely written as the SVA property:

\[
\text{start} \rightarrow \text{data}[-2] \#\#1 \text{ stop} \#\#1 !\text{stop}
\]

III. ASSERTION DEBUGGING METHODOLOGY

This section presents a methodology that automatically debugs errors in failing assertions. It is assumed that errors only exist in the assertions and the design is implemented correctly. This parallels the assumption in automated RTL debuggers [7]–[11] where the verification environment is assumed to be correct. However even with these assumptions, both assertion and RTL debuggers can be used in conjunction to return the union of the two results. Using these results, the engineer can then determine where to make the appropriate fix. As an added benefit, the results from the assertion debugging methodology can also provide insight into the design behavior to aid both assertion and RTL debugging. Note that this methodology makes no assumptions about the assertion language or the types of errors as these are functions of the input model.

The methodology aids debugging by returning a set of verified properties with respect to the design that closely relate to the failing assertion. We denote this set as \( P \). This set of closely related assertions aids in the debugging process in several ways. First, \( P \) serves as a suggestion for possible corrections to the failing assertion. As such, it provides an intuitive method to aid in the debugging and correction process. Second, since the properties in \( P \) have been verified, they give an in-depth understanding of related design behaviors. This information is essential in understanding the reason for the failed assertion. Finally, \( P \) allows the engineer to contrast the failing assertion with closely related ones, a fact that allows the user to build intuition regarding the possible sources and causes of errors.

The overall methodology is shown in Figure 2 and consists of three main steps. After a failing assertion is detected by verification, the first step of applying mutations is performed. This step takes in the failing property along with the mutation model and generates a set of closely related properties, denoted as \( P' \), to be verified. Each property in \( P' \) is generated by taking the original failing assertion and applying one or more predefined modifications, or mutations, defined by the mutation model. This model defines the ability of the methodology to handle different assertion languages as well as different types of errors. We define a practical model for SVA in the next section but different models based on user experience are also possible.

The second step of the methodology quickly rules out invalid properties in \( P' \) through simulation with the failing counter-example. A counter-example in this context is a simulation trace that shows one way for the assertion to fail. The intuition here is that since the counter-example causes the original assertion to fail, it will also provide a quick filter for related properties in \( P' \). It accomplishes this by evaluating each property in \( P' \) for each cycle in the counter-example through simulation. If any of the properties in \( P' \) fail for any of the evaluations, they are removed from \( P' \). The resulting set of properties is denoted by \( P'' \).

The final step of the methodology uses an existing verification flow to filter out the remaining invalid properties in \( P'' \). This is the most time-consuming step in this process, which is the reason for generating \( P'' \). The existing verification flow can either be a high coverage simulation testbench or a formal property checker. In the case of the testbench, the properties in \( P' \) will have a high confidence of being true. While with the formal flow, \( P \) is a set of proven properties for the design. In most verification environments, both these flows are automated resulting in no wasted manual engineering time. The final set of filtered properties \( P \) are verified by the environment and can be presented to the user for analysis.

![Fig. 1. Example 1: Timing Specification](image-url)
IV. SYSTEMVERILOG ASSERTION MUTATION MODEL

This section describes a practical mutation model for SystemVerilog assertions to be used with the methodology described in Section III. These mutations are designed to model common industrial errors [2], [19] as well as misinterpretations of SVA [17]. This model is created based on our discussions with industrial partners, our experiences writing assertions, as well as common errors cited in literature. Note that other mutation models can be developed based on user experience.

Each mutation modifies the assertion either by adding operators, changing operators or changing parameters to operators. Each new property is generated by applying a fixed number of these mutations to the failing assertion. The number of mutations is defined to be the cardinality of the candidate and depends on the number of additions or changes to the assertion. In some cases, multiple or complex errors may require higher cardinality to model. The rest of the section will describe the various types of mutations in this model.

The first group of mutations involves modifying Boolean expressions. SVA provides operators for signal transitions across a pair of clock cycles and previous values. These operators can frequently be misused. For example, a common error occurs when interpreting the word “active”. It is ambiguous whether the intent is a rising edge ($rose(sig)$), or a level sensitive sensitive (sig). Similarly with the $past(sig, k)$ operator, the number of cycles ($k$) to evaluate in the past is a common source of errors. The mutation can model this error by adding or subtracting an integer $i$. The following table gives the mapping from Boolean operators to the set of possible mutations where <$s$> is a given signal.

<table>
<thead>
<tr>
<th>Name</th>
<th>Operator/Expression</th>
<th>Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boolean Expressions</td>
<td>$&lt;s$, $past$, $rose$, $fell$, $stable$, $changed$</td>
<td>Replace with {&lt;s&gt;, $past(&lt;s&gt;)$, $rose(&lt;s&gt;)$, $fell(&lt;s&gt;)$, $stable(&lt;s&gt;)$, $changed(&lt;s&gt;)$}</td>
</tr>
</tbody>
</table>

The next group of mutations involves commonly used arithmetic, logical and bitwise operators. They replace such operators with their related counter-parts. For example, if $&$ is mutated into $|$, it may relax a condition to generate a passing property. This can provide insight into possible places where the error may have occurred. Another example might be replacing $>$ with $>=$. These mutations for these types are operators are in the table below.

<table>
<thead>
<tr>
<th>Name</th>
<th>Operator/Expression</th>
<th>Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logical</td>
<td>$l$, $&amp;$, $</td>
<td>$, $!$</td>
</tr>
<tr>
<td>Bitwise</td>
<td>$=$, $k$, $\land$</td>
<td>Replace with {$\land$, $\lor$}</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>$+$, $-$</td>
<td>Replace with {+, -}</td>
</tr>
<tr>
<td>Equality</td>
<td>$&lt;, &lt;=, ==, &gt;=, &gt;, !=$</td>
<td>Replace with {$&lt;, &lt;=, ==, &gt;=, &gt;, !=$}</td>
</tr>
</tbody>
</table>

SVA sequential binary operators make up the next group of mutations. These operators have subtle differences that are easy to misinterpret. For example, intersect is similar to and except with the added restriction that the lengths of the matched sequences must be the same. Another example is replacing intersect with within which might produce a passing property that can give insight into the underlying error. The following table shows the possible mutations.

<table>
<thead>
<tr>
<th>Name</th>
<th>Operator/Expression</th>
<th>Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequential Binary Operators</td>
<td>$and$, $intersect$, $or$, $within$</td>
<td>Replace with $and$, $intersect$, $or$, $within$</td>
</tr>
</tbody>
</table>

Sequential repetition operators are the next group of mutations. These operators allow repetition of signals in consecutive or non-consecutive forms with subtle differences. For example, the non-consecutive goto repetition $sig[-->k]$ is frequently confused with the $sig[->k]$ operator. The difference between the two is whether the sequence ends with strictly $k$ matches, or if multiple cycles are allowed before the next occurrence. The mutations may either change the repetition operator or the number of repetitions performed using $i$ and $j$. The following table shows the mutations.

<table>
<thead>
<tr>
<th>Name</th>
<th>Operator/Expression</th>
<th>Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repetition</td>
<td>$[-&gt;k]$, $[-&gt;k;k]$</td>
<td>Replace op with {[$i$, $=$, $</td>
</tr>
</tbody>
</table>

The last group of mutations involves the implication operators. This family of operators are often used because most properties are evaluations based on a condition. The first type of mutation is a change between the overlapping ($->$) and non-overlapping ($=>$) implication. This accounts for when there is extra or missing delay between the antecedent and consequent. The next mutation extends this idea by allowing a multiple cycle delay after the antecedent with the addition of the $##$ operator. The third type of mutation in this group involves adding the first_match operator to the antecedent of the implication. This addresses a subtlety of SVA where
the consequent of each matching antecedent thread must be satisfied. The first_match operator handles this subtlety by allowing only the first matching sequence to be used. The following table outlines the possible mutations.

<table>
<thead>
<tr>
<th>Implication</th>
<th>Operator/Expression</th>
<th>Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>=</td>
<td>&gt;</td>
<td></td>
</tr>
<tr>
<td>{</td>
<td>=</td>
<td>&gt;</td>
</tr>
<tr>
<td>Append</td>
<td>=</td>
<td>&gt;</td>
</tr>
<tr>
<td>Add first match on antecedent</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The following example shows how mutations can be applied in the context of an entire property.

**Example 2** Consider the assertion described in Example 1. We can generate properties in $P'$ by mutating the failing assertion with each of the applicable rules. Here we show several examples of mutated properties with different cardinals.

```c
// Cardinality 1
P1: start => data[-2] ##2 stop ##1 !stop
P2: start |- data[-2] ##1 stop ##1 !stop
// Cardinality 2
P3: $rose(start) => data[=2] ##1 stop ##1 !stop
P4: $rose(start) => data[-2] ##2 stop ##1 !stop
// Cardinality 3
P5: start |_|!data[-2] ##1 stop ##3 !stop
```

V. PRACTICAL CONSIDERATIONS AND EXTENSIONS

The methodology outlined in Section III along with the model in Section IV generates a set of closely related properties, $P$. However practically speaking, they are only useful if the number of properties returned by the methodology is small enough to be analyzed by an engineer. Two techniques that greatly reduce the number of properties are discussed here.

The first technique deals with vacuous assertions. Assertions that are vacuous typically are considered erroneous since their intended behavior is not exercised. Similarly, all verified properties that are found to be vacuous for all evaluations are removed from $P$. This reduces its size significantly as seen in the experimental results.

The second technique deals with multiple cardinalities. As the cardinality increases, the size of the mutated properties, $P'$, increases exponentially. This may become unmanageable at higher cardinalities. To deal with this, $P'$ can be reduced by eliminating properties with mutations that have been verified at lower cardinalities. For example if the property $P2$ from Example 2 is found to be a verified property, it would remove $P4$ from consideration since it contains the same mutation from $P1$. The intuition here is that the removed properties do not add value because they are more difficult to contrast with the original assertion. This proves to be very effective in reducing the size of $P'$ for higher cardinalities by removing these inessential properties.

VI. EXPERIMENTS

This section presents the experimental results for the proposed work. All experiments are run on a single core of a dual-core AMD Athlon 64 4800+ CPU with 4 GB of RAM. A commercial simulator or property checker is used for all simulation and verification steps. All instances are generated from Verilog RTL designs from OpenCores [20] and our industrial partners. SVA is written for all designs based on their specifications.
The results of these experiments are presented in columns five and six of Table II. Column five describes the number of operators or variables that are potentially erroneous which we denote as \( P \). Column six shows the suspects divided by the total number of operators and variables as a percentage.

From the table, we see that path tracing returns suspects covering a large part of the assertion for all instances. This ranges from 50% for \( wb1 \) to 100% for \( usb1 \). On average, 71% of the assertion operators and variables are returned confirming the ineffectiveness of circuit debugging techniques for use with assertions. This motivates the need for the mutation based debugging methodology presented in this work.

B. Assertion Debugging Methodology

The experimental results from implementing the proposed methodology, mutation model and enhancements are presented in this section. For each instance, two sets of experiments are run. The first uses the accompanying simulation testbench to generate the initial counter-example and the final verification step. The counter-example is generated by stopping the simulation at the point of first failure, while the verification runs through the entire testbench. We refer to these experiments as testbench. The second set instead uses a formal property checker for these tasks and we refer to them as formal. Note that not all instances have both a testbench and formal environment, thus some entries will be not available and are denoted by N/A. In addition for the same instance, different environments may produce different results due to assumption constraints in formal which are not respected in testbench. Each instance is run through the methodology varying the cardinality from one to three. Table III shows the results of these experiments.

The first two columns show the instance name and cardinality. The next eight columns show the testbench experiments. Columns three and four show the size of the initial candidates, \( P' \), without and with the multiple cardinality enhancement from Section V, respectively. Columns five to seven show the number of passing assertions from \( P' \) after simulation with the failing counter-example (\( P'' \)), the number of vacuous properties found in \( P'' \), and the final number of verified properties (\( P \)). Column eight shows the percent reduction of the cardinality and vacuity enhancements from the unoptimized \( P' \). Column nine and ten show the counter-example simulation time followed by the total testbench verification time. The last eight columns show same respective results for the formal experiments. The last three columns of Table II present supplemental data regarding the number of clock cycles in the testbench counter-example, testbench final verification step and formal counter-example respectively. Run-times to create the mutated properties (\( P' \)) take less than two seconds and are not shown in the table due to space considerations. Time-outs for each step of the methodology are set at 3600 seconds and are indicated by TO.

The applicability of the proposed technique is apparent when analyzing the final number of verified properties in \( P \). Despite the large number of initial properties in \( P' \), the methodology successfully filters properties to a manageable size. This is important for debugging because if \( P \) is too large then it becomes impractical to use. Moreover we see that in the case of formal, most instances are able to return proven properties that precisely specify the behavior of the design.

\[
\begin{align*}
\text{rose}(rd) & \& \text{tim_cas} \#(1) : rd[*4] \rightarrow \text{rose}(wr_safe).
\end{align*}
\]

The mutation here changed the repetition operator from \([*7]\) to \([*4]\). This directly corresponds to the RTL error which changed the timing between the read and write phases. Showing a different case for \( sp12 \), a cardinality one solutions is:

\[
\text{wfe} \rightarrow \text{stable}(rwe) \rightarrow 1 \text{ within } \text{rose}(state=1) \text{ } \#0 \text{ (state=0) } \rightarrow 1.
\]

The mutation added \( \text{stable} \) to \( rwe \). This gives insight into the design behavior where the error in the RTL is that \( rwe \) does not toggle in the correct state.

Figure 3 shows the size of each set of properties on a log-scale for several sample instances. From the figure, we see that the multiple cardinality technique from Section V helps to reduce the size of \( P' \) from the original unoptimized \( P' \). This results in an average reduction across all instances of 23\%. Next, we see that simulation with the counter-example efficiently reduces the size of \( P' \) to \( P'' \) by an average of 59\% across all instances. This is critical to ensure that the run-time of the final verification step is minimized.

For columns six and fourteen in Table III, we see that the number vacuous solutions also contributes to the reduced size of \( P \) from \( P'' \). As a percentage of \( P'' \), these verified vacuous properties represent an average of 34\% of the set. In addition, the enhancements from Section V, shown in columns eight and sixteen, reduce the size of the unoptimized properties in \( P'' \) by 27\% across all instances.

Finally from Table II, we see that the run-time for the counter-example simulation and final verification step increases with the number of properties. For cardinality one, this does not significantly impact performance. For higher cardina- nalities, this becomes more costly causing time-out conditions in certain formal cases. This degraded run-time is the trade-off for being able to model more complicated types of errors. However, this can be avoided with more precise input mutation models so that a costly increase in cardinality is not needed.

VII. CONCLUSION

In this work, we present a methodology, mutation model and additional techniques to automatically debug errors in SystemVerilog assertions. The methodology works by using the assertion, counter-example and mutation model to generate a set of alternative properties that have been verified against the design. These properties serve as a basis for possible corrections as well as provide insight into the design behavior and failing assertion. Experimental results show the methodology is effective in generating high quality alternative properties for all empirical instances.
REFERENCES