

SystemVerilog, Batteries Included:

A Programmer's Utility Library for SystemVerilog

Jonathan Bromley
jonathan.bromley@verilab.com

André Winkelmann
andre.winkelmann@verilab.com

Verilab Inc., Austin, Texas

Abstract—As the language of choice for many verification engineers, SystemVerilog is expected to act not only as a specialist verification language, but also as a hardware description language and a general purpose programming language. Although SystemVerilog's object oriented programming features and rich set of native data types provide excellent support for general purpose programming, many users including the authors have been frustrated by its lack of utility features that would be taken for granted in other languages. In this paper we present the results of our efforts to develop a comprehensive, consistent, easy-to-use utility library for SystemVerilog. We believe this to be the first announcement of such a library that is vendor-independent and carefully tailored to the needs of SystemVerilog users.

Keywords—SystemVerilog, DPI, svLib, utility library, document object model

I. INTRODUCTION

The project described in this paper, provisionally named svLib, is a collection of SystemVerilog utility packages. It is freely available under a permissive open-source licence, along with source code, documentation, simple build scripts for popular SystemVerilog simulators, and examples.

svLib complements SystemVerilog's existing capabilities by providing a variety of utility functions that the authors have found lacking in the base language. It aims to be applicable to a wide variety of programming problems, including – but not limited to – SystemVerilog's traditional "home territory" of digital design verification. Nothing in svLib is verification specific; those specialized features are already well supported by methodology toolkits such as the UVM.

As an example taken from the authors' experience, consider a DUT whose configuration is described by a file in JSON[1] or YAML[2] format¹. The location of this file is specified by an operating system environment variable. The testbench should check that this file is at least as recent as some specified date, and then read configuration information from the file. Locating the file requires that the testbench be able to read an environment variable. Checking the file is up-to-date requires knowledge of the current date and time, and the ability to inspect directory contents and determine a file's modification time. Finally, to read the configuration we must parse information from the chosen file format, probably with further string processing to get the required information from values so

obtained. The library features of SystemVerilog do not readily support such tasks, comparing poorly with general purpose languages such as C and Java, and even less favorably with script languages such as Python, Ruby, Tcl and Perl. Users of SystemVerilog are forced into creating their own implementations, perhaps with the help of external software via DPI, because SystemVerilog lacks a sufficiently comprehensive library. Some facilities of this kind can be found bundled with major vendors' simulators, but these vendor-specific libraries are inherently non-portable and are not necessarily as complete as one might wish. There are also a few useful functions (notably regular-expression string matching) provided with the Universal Verification Methodology codebase, but once again the available offering lacks the completeness that users might reasonably expect.

The paper will describe the motivations that led to the creation of this library, and our experiences in designing, implementing and testing it. Particular attention will be given to our aim of providing a SystemVerilog API that is as natural and convenient as possible. Code examples and performance measurements will be presented, allowing potential users to evaluate the suitability of the library for their own applications. The source code and build scripts will be made freely available, at the time of the conference, under a permissive open-source license. The authors hope that the user community will provide further input and refinement, making this a truly standard library.

II. OVERVIEW OF SVLIB

svLib features fall into a few broad categories, with some inevitable interaction between them. For example, functions that manipulate file pathnames have been grouped together with other file system functions, although they are in truth nothing more than specialized string operations. Each category is provided as a SystemVerilog **package**, as follows:

- **svLib_Str_pkg**: general-purpose string manipulations, including regular expressions
- **svLib_File_pkg**: functions for manipulating file and directory names; directory lookup, directory listing, and file property inquiries
- **svLib_Sys_pkg**: functions for interacting with the operating system including environment variables, command-line arguments, and wall-clock time and date
- **svLib_Utils_pkg**: miscellaneous utility features including convenience functions for manipulating enumerated types

¹ JSON and YAML are both file formats that provide textual representation of native data structures in a form that is both human- and machine-readable. Many other such file formats exist, for example XML.

- `svLib_Cfg_pkg`: toolkit for reading and writing configuration data stored in files using `.ini` or YAML representation

Finally, a header file `svLib_macros.svh` defines some SystemVerilog macros providing automation of object serialization, and some miscellaneous operations that could not otherwise be offered in a convenient way.

Full details of the complete contents of the packages can be found in the distribution, but the following brief descriptions illustrate a few highlights and provide some simple examples of library usage.

A. String manipulations and regular expressions

`svLib_Str_pkg` supports full-featured regular expression matching. It offers capture of matches and sub-matches, automated search-and-replace including sub-match placeholders, and global matching and replacement (find all possible matches in a source string). A regular expression can be saved as an object, making it faster and more convenient to use for subsequent match attempts. These features of `svLib` are underpinned by the POSIX-compliant regular expression processor of the standard C library, providing a robust and well-documented implementation as described by the Linux/Unix document `man 7 regex`. We are aware that more modern and advanced regular expression processors exist, notably the Perl-compatible library `pcre`[3], and we plan to add support for this regex dialect in a future version of `svLib`.

The package also provides a range of string manipulation functions including finding a substring within a string, insertion of one string into another, splitting a string into a queue of strings on every occurrence of a separator character, and the string join function familiar from many script languages. Right, left and center justify functions simplify the preparation of data for display in tabular format.

Finally, functions are provided that can read and write numeric values in the full Verilog numeric literal syntax such as `16'hF6_Ex`. This feature plugs a minor gap in the existing formatted I/O support.

B. Interacting with the file system

Traditionally it has been impractical for a SystemVerilog program to make file system inquiries, for example to discover the last modification time of a file, or to get a list of all files in a directory. `svLib_File_pkg` provides easy access to these operations directly from SystemVerilog. For example, to get a queue of strings containing the names of all files in the current directory whose names end in `sv`, the following code fragment could be used:

```
string dirList [$];
dirList = svLib_File_pkg::glob("*.sv");
```

`svLib_File_pkg` also includes functions to manipulate file path names, making it easy (for example) to extract the filename alone from a full pathname. Files can also be renamed, moved, and deleted directly by SystemVerilog code.

C. Interacting with the operating system

`svLib_Sys_pkg` has functions to find the current working directory, work with environment variables, and find the complete command line used to launch the simulator (which provides a solution to the long-standing problem of how to handle multiple plusargs with the same name).

This package also contains functions to read wall-clock time and date, and convert it to or from human-readable text.

D. Configuration data files

`svLib_Cfg_pkg` provides the ability to read a configuration file, in `ini` or YAML format, making the file's contents available so that your SystemVerilog code can read any named item from it. This will typically be used to populate configuration objects at the start of a simulation run. Likewise, arbitrary user data can be written out to a file in the same formats. Support for Comma Separated Value (CSV) files is under consideration but has not yet been implemented. The configuration file features are described in more detail in section VIII.

III. DIVIDING RESPONSIBILITY BETWEEN LANGUAGE AND LIBRARIES

A. There's Something Missing

Verification engineers face difficult problems daily, and it is not surprising that sometimes they express frustration that SystemVerilog does not provide all the support they need. The evident success of the UVM [4] and other verification methodology base class libraries (BCLs) clearly indicates strong user appetite for ready-to-use packages of domain-specific functionality. Consequently it was surprising, to the authors at least, that there does not appear to be any comprehensive and readily available library of utility functions for SystemVerilog. Users working in a general-purpose programming language such as C or Java, or a script language such as Python or Perl, would expect as a matter of course to have access to a huge and well-established collection of functions for string manipulation, file access and file system exploration, operating system interface, data structure manipulation, and similar utilities. SystemVerilog's provision of such functions is patchy: for example, text file I/O is quite well supported, but string manipulation is provided through a limited and somewhat idiosyncratic repertoire of operations, and operating system interface support is vestigial.

One possible reason for SystemVerilog's weakness in this area is that the core language, as specified by its LRM [5], already provides a useful selection of functions that in a more general-purpose language might be provided as library functions. For example, extensive formatting and I/O of textual information is available through the `$display`, `$scan` and `$sformat` families of system functions. However, having such functionality built into the core language makes it difficult to add new features, as they must be added through the protracted process of language standardization. The authors' perception is that the SystemVerilog user base has become accustomed to the idea that utility operations should be built in, and new operations can be added only by extending the language.

This situation is at least in part to blame for the often-heard criticism that SystemVerilog is bloated and contains too many special features. One of the authors has previously argued [6] that this criticism is largely unjust, and that it is entirely appropriate for SystemVerilog to have an unusually rich set of specialized features. Statically elaborated instance hierarchy, constrained randomization, temporal assertions, and functional coverage are good examples of EDA domain-specific requirements that are better provided as core language constructs with their own dedicated syntax, rather than as library functions presented to the user through the somewhat restricted syntax of calling a task or function. On the other hand, the huge selection of library functions that is conveniently available to any C programmer has no counterpart in SystemVerilog, and many of those functions would be very welcome additions to a typical verification engineer's toolkit.

Each of the major simulators ships with some kind of utility library offering some of these functions. However, in every case the feature set is patchy and incomplete, and – more important – many users prefer to avoid vendor-specific offerings because commercial or technical imperatives require them to be able to use more than one tool with a common code base.

B. Does the DPI make a library superfluous?

One particularly interesting feature of SystemVerilog is its comprehensive and well-standardized C-language interfaces, the VPI and DPI, described in clauses 35 and 36 of [5]. One well-known consequence of the DPI is that users can now directly call any function in the usual C library if its argument and return data types are among the set of types that are natively supported by the DPI interface mechanism. The following code example² shows how a standard C math library function can be made directly available in SystemVerilog using the DPI:

```
module test_dpi_sqrt;
  import "DPI-C" function real sqrt(real x);
  initial
    $display("sqrt(9.0)=%f", sqrt(9.0));
endmodule
```

No further coding effort is required. The DPI argument passing mechanism automatically ensures that the `real` argument and return type appear as `double` values in C, and it is merely necessary for the user to ensure that the SystemVerilog simulation has access to the necessary C library object code using appropriate command-line options.

Given this very convenient access to the existing wide repertoire of C functions, it might seem unnecessary to provide any kind of utility library in SystemVerilog. However, further consideration soon shows that this is not the case. Many of the

C library functions of interest depend heavily on data structures that cannot reliably be mapped directly into SystemVerilog, so some indirection is required. Furthermore, as svLib is ported to other platforms, some underlying C functions may change but we must preserve an unchanged SystemVerilog API.

For example, svLib provides a function `svLib_File_pkg::mtime`, which returns the last-modified timestamp of a file. In svLib, wall-clock time/dates are consistently represented as seconds since the beginning of 1970, using a `longint`. In the C library, a file's timestamp information (along with other status) is returned in a `struct stat` that must be allocated by user code. We do not wish to trouble the SystemVerilog programmer with these details, which are specific to the C library, so it is preferable to provide wrapper functions. Some of this wrapper functionality is implemented in C and some in private SystemVerilog code; the precise split is opaque to users, and may change in future revisions of svLib, but the user-facing `mtime` function can remain unchanged.

It is also important that the library as a whole should not exhibit memory leaks, i.e. situations in which memory is allocated but never released. This can be challenging to achieve in svLib because some C library functions must allocate data structures in C memory before returning that information back to SystemVerilog. After the return to SystemVerilog, the C code is no longer running and has no way to deallocate the memory that it needed. In svLib we can handle this problem by careful management of the shared responsibility between SystemVerilog and C, so that the C code is always given a chance to deallocate any memory it has used. But this can only be done if a strict protocol is established between the C and SystemVerilog parts of the library, and this in its turn implies that there must be some parts of the library infrastructure that remain hidden from users so that their integrity is assured. The overall library architecture then appears as in Figure 1.

² Since the 2009 revision of the language standard SystemVerilog has supported a wide selection of math routines, including `$sqrt`, as predefined system functions. Consequently this example is no longer useful, but it serves to illustrate the principles involved.

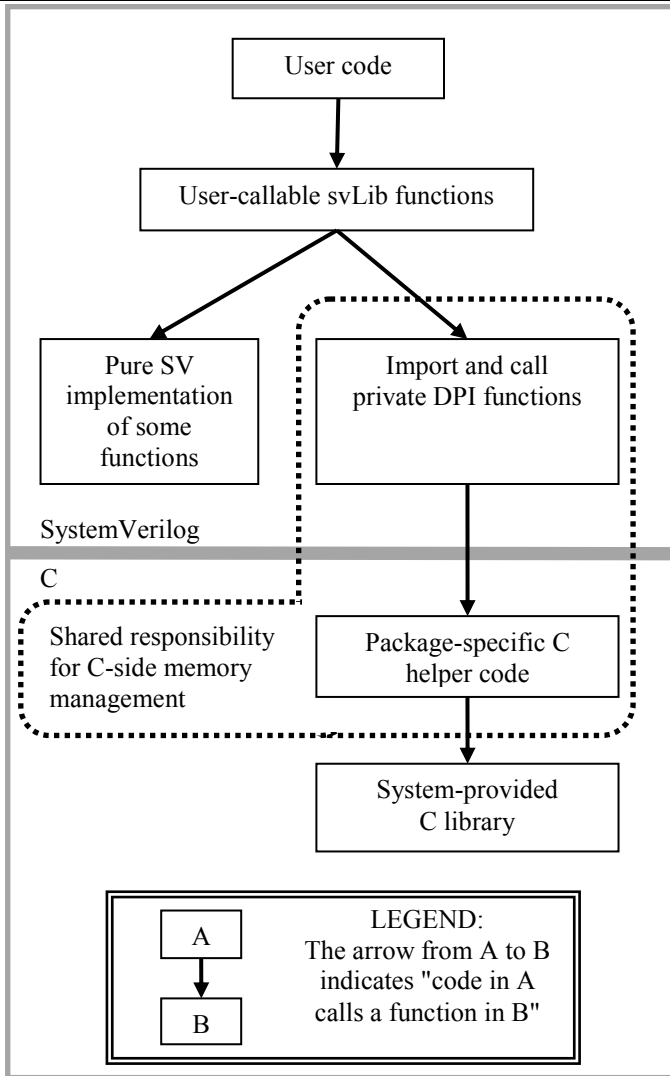


Figure 1: SV/DPI/C architecture

IV. SPECIFYING THE LIBRARY

Our SystemVerilog utility library, svLib, aims to provide a wide selection of useful functions for general programming. We had several design principles in mind from the outset. We found that adhering to these principles when specifying svLib was by far the most challenging part of our task. By contrast, the implementation and testing described in section X was comparatively straightforward, and fitted readily into our usual coding practices.

Our core goals, outlined in more detail in later parts of this section, were as follows:

- Convenient API for SystemVerilog programmers, with no C or DPI programming required
- Consistent and uniform handling of error conditions
- Consistent naming conventions
- Portability across simulators, operating systems, and host system architecture

- No dependencies on any other SystemVerilog packages
- Compatibility with all popular verification methodologies
- All deliverables to be open-source, unencumbered by any restrictive licence agreements

In addition, we naturally paid attention to ensuring that the library is as free of bugs as possible, does not impose any undue performance penalty, and is well documented. These latter goals, however, can be revisited and improved in future versions of the svLib code base. By contrast, it was essential to meet our core goals from the outset in order to avoid the severe inconvenience to users that might arise if svLib's user-facing specification were to change in the future.

A. Convenient API for SystemVerilog programmers

We have never expected svLib to offer unique functionality that cannot be found elsewhere. Every operation provided by svLib can be found in some other language or library. svLib is valuable not for any novelty, but by offering useful and powerful facilities in a way that is completely natural and straightforward in use for any SystemVerilog programmer³. To make use of svLib, no extra effort beyond importing a package should be required of the programmer.

Beyond this simple use model, though, we spent a great deal of effort in considering the most appropriate functions to include and, especially, what form they should take. Two issues gave us particular difficulty: the handling of error conditions as described in subsection IV.B below, and the choice of whether to use an object-oriented approach. The question of whether to present svLib as an OO (class-based) package was sufficiently troublesome that we have sidestepped it by providing both an object-oriented API and a straightforward function-call API for many of its features. Users are free to choose whichever approach better suits their purposes and coding style. This issue is discussed more fully in section V.

B. Consistent and uniform handling of error conditions

In any utility library, some of the functions may give rise to error conditions. Sometimes these errors are purely internal (for example, a string overflowing some pre-allocated buffer space), some are genuine usage errors that the programmer probably wants to know about (for example, attempting to discover the modification date of a file that does not exist), and some errors might or might not be considered to be user errors (for example, attempting to examine a file in a directory for which the user does not have read or execute permission). Library users must be given the opportunity to handle any errors gracefully.

³ There is an important exception to this general aim. From the outset we anticipated that svLib would never be appropriate for RTL design. svLib makes heavy use of non-synthesizable features such as the DPI, VPI, and dynamic data types. Consequently it is unsuitable for use by RTL designers writing synthesizable code.

1) Error code in an output argument

We briefly considered the possibility of having every function pass its error code through an output argument. However, we quickly dismissed this approach as clumsy and impractical. It forces the user to declare a suitable result variable and explicitly pass it to each function call. Default output arguments, introduced in the 2009 revision of the SystemVerilog standard, could make this approach tractable. Unfortunately this valuable language feature is not yet implemented in some commercially available simulators at the time of writing.

2) Error code as a function return value

Our first attempt to achieve uniformity was based on having every svLib function, without exception, return an integer error code as its result. In most cases there are many possible error conditions, but only one success code is required. Consequently we chose to have such functions return zero (false) on success, and an appropriate integer code for any kind of error.

Consider, for example, the function that reports the current working directory as a string. Using this approach, the function prototype and an example of its use would be:

```
// declaration in svLib_Sys_pkg:
function int getcwd(output string dirname);

// user code:
import svLib_Sys_pkg::*;
int err;
string cwd;
err = getcwd(cwd);
if (err)
    $display("svLib error %0d occurred", err);
else
    $display("working directory is %s", cwd);
```

Although attractive for its consistency, this approach was very inconvenient in practice. It imposed a significant burden on the programmer to capture every function's result in an error variable or by means of a test, when for many of the library's functions it is more natural simply to return a result value.

However, this function is essentially free of possible error conditions and would be much more natural if expressed thus:

```
// declaration in svLib_Sys_pkg:
function string getcwd();

// user code:
import svLib_Sys_pkg::*;
string cwd;
cwd = getcwd();
$display("working directory is %s", cwd);
```

In addition, the need to use zero (false) for success made for user code that was error-prone and hard to read, and introduced serious difficulty with the naming of some

functions. For example, we have a function `regex_match` that tests a string against a regular expression. But that name was completely inappropriate when the return value was to be zero (false) on success.

Although it was a hard decision to reject a solution that offered such complete consistency, we felt that this approach was just too clumsy to be comfortable, and we abandoned it.

3) Our chosen solution

Following our avowed aim of making svLib as painless as possible for SystemVerilog programmers to use, we finally settled on a different approach.

As far as possible, functions are designed so that in the event of user error they return a result that is sensible and – most importantly – not misleading. For those few functions where the user clearly has a responsibility to anticipate and handle an error condition, we provide an output argument that the user must populate with an error-code variable. Other functions are unencumbered by error reporting machinery.

This leaves the problem of how to handle errors that the user would not normally anticipate. In these cases we detect the error in procedural code, but then use a SystemVerilog immediate assertion to throw an error, providing detailed diagnostics at the console. Using an immediate assertion in this way makes it possible to suppress the error using `$assertoff` or some equivalent tool command. We also provide package configuration functions that allow users to do the same thing without needing to know the names of the assertions in question.

It is appropriate to observe here that errors in svLib are invariably reported on the SystemVerilog side. Many such errors in fact originate in C code that was called through the DPI, but such errors are always handled defensively in our C support code and passed back to SystemVerilog for presentation to the programmer in a SystemVerilog-friendly form.

For users who wish to handle errors for themselves, we also provide a means to test error status of the most recent svLib function call. This information is recorded separately for each SystemVerilog process, so that there is no risk of some other process disturbing the error information between occurrence of an error and the user's code testing for it.

The per-process error recording mechanism is described in section VI. It provides a useful compromise for the conflicting demands of robustness, convenience and flexibility:

- By default, library errors cause an assertion-style error message, with the function that caused the error taking defensive default actions and returning some default result.
- Users can easily suppress these error messages if required, accepting the risk of occasional unexpected behavior if they do not test for such errors explicitly.
- Advanced users retain the ability to test explicitly for svLib errors and handle the consequences in any way they choose.

C. Consistent naming conventions

svLib provides various kinds of named items that a user may need to mention in their code: packages, classes, functions and their arguments, and enumerated types with their enumerator⁴ names. Users value consistent naming to make item names not only memorable and somewhat descriptive, but also easy to find in the documentation.

As with many matters of style, this is sure to involve some compromise. Descriptive names are inevitably longer and therefore more cumbersome. For each major group of functionality in svLib we have attempted to prepare practical examples of its use. That experience has sometimes led us to change our opinion about what names are most suitable. Generally we have chosen the following guidelines:

- `camelCase` for names that are formed from more than one word, such as `findFirst`;
- underscore separator between a prefix or suffix (identifying a group of related ideas or features) and the remainder of a name, such as `regex_match` and `svLib_File_pkg`;
- short names wherever possible for class methods, which will always be prefixed by an object name;
- names that are unique across all packages that make up svLib, so that it is feasible to wildcard-import the whole library (several packages) if you so wish;
- compact names for the packages, so that it remains reasonable to use explicit package prefixes if you wish to avoid wildcard importing of all names in a package;
- long and very descriptive names for svLib internal functions and other items such as DPI imported function names that should never be referenced by users, to minimize any risk of collisions with ordinary user-defined names.

D. Portability across simulators and platforms

From the outset we were committed to ensuring that svLib is fully portable across all major simulators. SystemVerilog simulators from three major vendors have now reached a high level of maturity and very little tool-specific treatment was required to achieve compatible behavior with all three tools.

However, these tools are now widely deployed on a variety of platforms, often but not always using the Linux operating system where the distinction between 32-bit and 64-bit platforms is important. We have tested the current release of svLib on a range of platforms but we cannot be sure of having covered all possibilities. Instead, we have taken great care to ensure that the implementation is as robust as possible against such things as endianness or 32/64-bit differences, and we hope to engage the community in helping to extend the set of fully supported platforms.

⁴ The word *enumerator* denotes one of the named values that make up an enumerated type's set of permissible values. The authors rather like the portmanteau *enumerel* to mean the same thing, but that latter word appears to have been hijacked by the GCC documenters as a synonym for "enumerated type".

Microsoft Windows® is also used with some SystemVerilog simulators, but at the time of writing we have been unable to extend our testing to that platform. Differences between the Windows and Linux C libraries may be expected to present some challenges.

E. No dependencies on any other SystemVerilog packages

svLib aims to provide services to any SystemVerilog programmer. As such it must be a truly stand-alone library, capable of deployment in any user environment regardless of what other packages that environment uses.

svLib does have some dependencies, of course. It makes heavy use of functions from the standard C library, which we assume to exist on any practical simulation platform. It also makes use of one third-party package, the `libyaml` open-source YAML parser/writer. A copy of this package is bundled with the svLib distribution. Neither of these dependencies is likely to be burdensome for the SystemVerilog programmer.

Although there is always a possibility of name collisions between a new package and users' existing code, such collisions can always be sidestepped if necessary by using fully qualified package names, or explicit import of individual names, rather than wildcard package import.

F. Compatibility with all popular verification methodologies

svLib makes no attempt to intrude on the territory of verification methodology libraries such as the UVM. It aims to be complementary to them, providing functionality that will be as useful in a UVM-based testbench as in any other verification code. svLib can be used alongside any published or proprietary verification methodologies with no special considerations required. As described in section IV.B, svLib's built-in error messages can easily be suppressed, allowing the programmer to detect svLib error conditions and report them through a methodology-specific reporting mechanism if desired.

Inevitably there is some overlap of features between svLib and verification base class libraries such as the UVM. For example, the UVM already provides regular expression matching and interrogation of the simulator command line (although in both these cases svLib's support is considerably more comprehensive and flexible). Although there are no fundamental problems caused by this overlap, we have attempted to choose names for any such functions in a way that minimizes any confusion with comparable UVM features.

There is, however, one specific area where there is an interesting possible synergy between the UVM and svLib. As discussed in section VIII, `svLib_Cfg_pkg` offers convenience macros to automate the serialization and deserialization of user-specified objects to and from a configuration file holding a textual representation of such objects. These macros have a very similar look and feel to the field automation macros of the UVM. If, at some future time, svLib were to be distributed with the UVM as a matter of course, the UVM field automation macros could easily be augmented to provide this serialization support, rendering the svLib serialization macros unnecessary in a UVM environment.

Similar considerations apply to the OVM[7], which has close similarities to the UVM and is still in active use.

G. All deliverables to be open-source

svLib has initially been developed by a small group within our company. Its ambition, though, is to offer a utility package that will be widely adopted among the SystemVerilog user base. This goal can be met only if it is fully open-source. We have applied the rather liberal Apache open-source licence [8] to it, just as Accellera has done with the UVM. It is hoped that this will present no obstacles either to adoption or to community involvement with its development.

The third-party `libyaml` package used by svLib is available under the MIT open source licence [9], which is similarly permissive.

V. CLASS METHODS OR PACKAGE FUNCTIONS?

Our initial instinct was to present svLib to the user as a set of SystemVerilog classes. We imagined that we would have a class to represent a regex, another class to represent a filename, and so on. From the library writer's point of view, packaging user data in this way is extremely attractive because it allows us to associate arbitrary hidden data with each object. In particular, this hidden data could allow our SystemVerilog code to keep track of what was happening in the C world, allowing many optimizations because we could often avoid passing information across the DPI boundary thanks to this local knowledge of the C state.

As soon as we began to sketch out the user-facing API, though, we immediately encountered a serious problem of usability. SystemVerilog users will typically expect library operations to accept input arguments that are native SystemVerilog types (particularly strings), process them in various ways, and return their results also as native data types. A good example is provided by the string `trim` function, which removes leading and/or trailing whitespace from a string. There are at least four plausible ways to provide this functionality:

- as a simple function that accepts a string argument and returns a string result;
- as an operation on an object representing the string, changing the object's contents to represent the trimmed string;
- as an operation on an object representing the string, leaving the object's original contents untouched, and returning a new object containing the trimmed string;
- as an operation on an object representing the string, leaving the object's original contents untouched, and returning a SystemVerilog string containing the trimmed value.

Of those four approaches, only the first is likely to be attractive to SystemVerilog users who have become familiar with the behavior of the language's native string data type. But the advantages to the library writer of an object-based representation are too great to ignore, especially for functions that must do significant work on the C side of the DPI.

We finally decided to adopt a hybrid approach in which classes, defined by svLib, are used to represent almost all user

data. The methods of these classes are available for any user who prefers the explicit class-based representation. However, we anticipate that the majority of users will prefer a straightforward function-call interface at the package level. We have therefore implemented the classes with more concern for efficiency and ease of library management than for usability, and we have provided a set of package-level convenience functions around those classes. In this way we hope to get the best of both worlds, with a robust and extensible class-based infrastructure that nevertheless can be used in a very straightforward way through the convenience functions for users who prefer not to adopt the OO approach. Careful implementation of the convenience functions minimizes the amount of data copying performed.

An interesting consequence of this decision is that each user-facing function requires one or more new objects to be brought into use. If this were done in the obvious way by constructing the objects on demand, it would probably create an unreasonable burden on SystemVerilog's memory allocation and garbage collection. To mitigate this, the library maintains a pool of recyclable objects that can be used internally by the convenience functions. These objects are used only briefly by each call to a convenience function, and are then returned to the pool for use by future function calls.

To illustrate the way this model is presented to users, here are some excerpts from the svLib declarations of the string object class and its trim function, together with examples of how a user might apply them.

```
package svLib_Str_pkg;
// svLib object that holds a string
class Str extends svLib_base;
typedef enum {NONE,LEFT,RIGHT,BOTH} side_e;
extern function Str create(string contents="");
extern function string get();
extern function void trim(side_e side=BOTH);
...
endclass
// Convenience wrapper for the trim function
function string Str_trim(
    string s, Str::side_e side=Str::BOTH);
...
```

The `create` method of class `Str` acts as a constructor, but has some useful additional properties that are discussed in more detail in section VII. Users are discouraged from directly calling the constructor of any svLib class.

The `get` method, not surprisingly, returns the object's string contents as a native SystemVerilog string.

It is noteworthy that we have chosen to put the definition of enumeration type `side_e` into the `Str` class, rather than exposing it at package level. If it were a package-level definition, then the short names `NONE`, `LEFT`, `RIGHT`, `BOTH` would be injected into the user's namespace by a wildcard import of the package. This would be very likely to lead to name collisions. By concealing its typedef in a class, we keep

the names short and descriptive while enforcing the use of a simple and mnemonic `Str::` prefix when they are used.

Equipped with this set of declarations, we can now perform string trimming in two different ways. A user who has a simple one-off need to remove surplus space from the beginning or end of a string can use the convenience function directly:

```
import svLib_Str_pkg::*;
...
string s = "  needs trimming  ";
$display("s=%s", Str_trim(s, Str::RIGHT));
// displays "s= needs trimming="
```

By contrast, a user who needs to perform many successive operations on a string, or who wishes to be able to pass string objects around easily by reference, may find it preferable to use the `Str` object representation:

```
import svLib_Str_pkg::Str;
...
Str s = create("  needs trimming  ");
s.trim(Str::RIGHT); // do the trim operation
... // possibly do further operations on 's'
$display("s=%s", s.get());
// displays "s= needs trimming="
```

VI. PER-PROCESS ERROR HANDLING

As described earlier, we chose to report errors by capturing error information from the most recent failed library function call, making this available to users through a query function (which cannot itself yield any errors, and does not affect the library error status). To avoid any risk of corrupting the error information in one process by the activity of some other process, we record this error information independently for each process.

This is straightforward to implement by maintaining an associative array of error objects indexed by process handle. Unfortunately, one widely used commercial SystemVerilog simulator does not support associative arrays indexed by class or process handle. Fortunately the offending tool (alone among currently available simulators) provides a unique string name for any process, obtained by formatting the process handle itself as a string using the `%p` formatter. For that simulator only, we keep per-process information in an associative array with string index type. This imposes a slight performance penalty, measured as about 1.3 μ s of run time per library function call on a modest 32-bit Linux workstation. Although this performance impact is unfortunate, we believe it to be an acceptable price to pay for the convenience and robustness of per-process error tracking.

We also chose to keep the error information distinct for each of the five packages that make up `svLib`. Although this gives slightly better granularity of error management, it was mainly motivated by implementation convenience.

VII. EFFECTS ON RANDOM STABILITY

In SystemVerilog, the creation of any new object (of class type) disturbs the random number seed of the creating process, as described in clause 18.14 of [5]. This is problematic for `svLib` because, although `svLib` itself does not participate in random number generation in any way, it creates objects on the fly in a way that is very hard for users to predict or control. Since `svLib` is much concerned with string and file processing, it seems likely that it will be heavily used in diagnostics that may be inserted or removed from the user's codebase as development and debugging proceeds. If the introduction of debug code affects random stability, the debugging task becomes much more difficult because it is impossible to reproduce a problem case after adding the instrumentation.

To mitigate this problem, we have implemented `svLib` in a way that guarantees it will have no impact on random stability, regardless of how many objects it creates. This is achieved by outlawing use of the constructor of any `svLib` class, instead delegating construction to a static `create` method provided by each class. The relevant parts of the `create` method of some imaginary class `svLib_C` is shown here.

```
class svLib_C extends svLib_base;
  static function svLib_C create(...);
    std::process p = std::process::self();
    string rs = p.get_randstate();
    create = new(...);
    ... // any other initialization activity here
    p.set_randstate(rs);
  endfunction
  ...
```

By saving and restoring the randomization state of the calling process across object construction, we guarantee that `svLib` cannot disturb the user's randomization and therefore users are free to add or remove `svLib`-based debugging code with confidence.

A compile-time macro can be defined to defeat this mechanism, removing its (rather small) performance overhead but sacrificing the guarantee of random stability during debug. This option is intended for confident users who have a fully debugged environment that they wish to use for a regression, with many different random seeds, where the highest possible performance is considered to be more valuable than random stability.

VIII. WORKING WITH CONFIGURATION FILES

As outlined in the introduction, an important driver for the development of `svLib` was our desire to be able to read and write configuration data files in structured formats such as YAML. In our experience it is common for such files to be created, typically by scripts, to reflect the desired DUT and testbench setup for a given simulation run. The verification environment should be able to read such a file and use its contents to populate one or more configuration data objects.

At first glance this requirement seems to be centered around the problem of reading and writing the desired file

format. However, interpreting the file is not sufficient. It is also necessary to provide some automation of the conversion between native data structures and a representation that maps easily on to the file.

As an example, consider these SystemVerilog classes, and objects of those classes with some values stored in their data members. Note the hierarchical instantiation of a `simpleConfig` object within a `largerConfig`.

```
class simpleConfig;
  int scalarInt;
  string scalarString;
  int arrayInt[3];
endclass
class largerConfig;
  int scalarInt;
  simpleConfig objectSC;
endclass
...
simpleConfig sc = new;
sc.scalarInt = 1234;
sc.scalarString = "string in sc";
largerConfig lc = new;
lc.scalarInt = 5678;
lc.objectSC = sc;
```

If we were to write the `largerConfig` object out to a YAML file, we would expect the file's contents to appear approximately as follows:⁵

```
%YAML 1.1
---
# Contents of largerConfig object
scalarInt: 5678
objectSC:
  # contents of simpleConfig object
  scalarInt: 1234
  scalarString: string in sc
```

Although the structure of the YAML data is clear enough, there is no straightforward way to map from it to the contents of our two SystemVerilog objects. For more complex objects, the problem would be even more troublesome. Furthermore, it is likely that not all data members of an object should be represented in an external configuration file. For all these reasons, it became clear that we needed an intermediate representation. Borrowing from other software environments, we chose to think of this representation as a tree-structured *Document Object Model* (DOM).

⁵ The YAML file format is quite flexible and allows many detailed variations in the way data is represented in the text file. The example shown here is legal YAML, but is by no means the only possible representation. Lines beginning with a `#` character are comments and do not influence interpretation of the data.

A. Characteristics of the svLib DOM

In svLib a DOM is a tree structure (a directed acyclic graph) with a single root node. A node in the tree can have one of three basic forms:

- a *map*, whose value is an un-ordered collection of nodes, each node identified by a string name (key)
- a *sequence*, whose value is an ordered list of nodes with each node identified by its position in the list
- a *scalar*, whose value is a "scalar value" object

Clearly, scalars are the leaves of the node tree. We chose to have each scalar's value be represented by a further object in order to make it easy to add new scalar data types in future versions. Currently only integral and string scalars are supported; we have yet to encounter a situation in which this is truly inadequate, but further data types are planned for increased flexibility and convenience. Scalar value classes are parameterized by the type of raw data that they hold, in much the same way as the UVM's resource database uses parameterized SystemVerilog classes to provide resource storage for any user-defined data type.

The two structured node types – sequence and map – are sufficient to represent almost any reasonable data structure, but they also map naturally on to native SystemVerilog types. A DOM map is simply an associative array of nodes indexed by string; a DOM sequence is a queue of nodes. In this way, we were able to create a DOM in SystemVerilog with only a few new classes, all derived from two base types:

- `cfgNode` and its subclasses `cfgNodeMap`, `cfgNodeSequence`, `cfgNodeScalar` represent the different kinds of DOM node. A `cfgNodeScalar` contains a `cfgScalar` object representing the scalar value at the node.
- `cfgScalar` and its subclasses such as `cfgScalarString` are containers for scalar values.

Users of svLib typically are not expected to interact directly with a DOM. Instead, they will use built-in svLib functionality to copy a DOM to or from any of the supported file types (currently YAML and the more restricted INI file format).

B. Representing user-defined data as a DOM, and vice versa

Although the DOM structure is very flexible and provides a consistent way to represent any required data structure, it is too clumsy for ordinary use. Users of course expect to create their own configuration data objects (typically using classes) in whatever way is most convenient. There remains, therefore, a problem of how to transfer data between the user's classes and a DOM representation. Users could write their own methods to create and unpack a DOM representation from the object, but this is tedious and error-prone, and automation is obviously desirable. Frustratingly, SystemVerilog lacks the introspection and reflection facilities that are needed to make this fully automatic.⁶ To handle this problem we took our cue from the

⁶ At first we hoped to use SystemVerilog *attributes*. VPI code would scan the class definition, finding data members with the

UVM and provided field automation macros to achieve the same result. Using this approach, our `largerConfig` object becomes:

```
class largerConfig;
  int scalarInt;
  simpleConfig objectSC;
  `SVLIB_DOM_UTILS_BEGIN(largerConfig)
    `SVLIB_DOM_FIELD_INT(scalarInt)
    `SVLIB_DOM_FIELD_OBJECT(objectSC)
  `SVLIB_DOM_UTILS_END
endclass
```

The other class `simpleConfig` is handled in a very similar way.

These macros automate the code for two new methods of the class:

```
cfgNode toDOM(string name);
void fromDOM(cfgNode dom);
```

`toDOM` constructs and returns a new DOM map node containing a representation of the object's contents, *including those of the lower-level object*. This DOM can then be written to a file of any supported format, simply by passing it to the `serialize` method of an appropriate `cfgFile` object; in the current version of `svLib`, subclasses `cfgFileYAML` and `cfgFileINI` are provided.

`fromDOM` uses the given DOM's contents to populate the object's data members, again descending as necessary into the inner `objectSC` object.

The following example shows how a `largerConfig` object can be populated from a source file `src.yaml`, modified, and then written out to another file `dst.ini` in a different file format:

```
largerConfig cfg; // The object to be populated
cfgNode dom;      // The DOM root node
cfgFileYAML yf;    // File reader/writer object

// Read the YAML file into a DOM
yf = new;
dom = yf.read("src.yaml");

// Populate cfg from the DOM
cfg = new;
cfg.fromDOM(dom); // also constructs inner object

// Modify the cfg objects
cfg.scalarInt = 42;
cfg.objectSC.scalarString = "new value";

// Make a new DOM from objects. Call it "NewCfg"
dom = cfg.toDOM("NewCfg");

// Write the new DOM to a file
yf.write("dst.yaml", dom);
```

Given the YAML file provided above, the resulting output in `dst.ini` would be as follows:

```
scalarInt=42

[objectSC]
scalarInt=1234
scalarString=new value
```

We can see that the INI file format uses `[section]` to denote first-level map entries that are themselves maps. Deeper nesting is not supported by INI files; similarly, sequence nodes cannot be written to INI files. YAML and other formats are much more flexible and expressive, and can represent the full range of DOM features.

Full details of these and other features of the DOM and file reader/writer classes can be found in the `svLib` documentation.

C. Ambiguities of the DOM

We have chosen a very lightweight DOM for `svLib`, emphasizing flexibility and simplicity. Consequently there are some aspects of typical user data structures that cannot be fully represented. In particular, map nodes are used for two quite distinct purposes: to represent an associative array indexed by string, and to represent the various named fields of a struct or the named data members of an object. In practice this ambiguity does not present any real difficulty, as the target object should have been designed to match the anticipated data structure. A key advantage of the automation macros is that various syntactic and semantic checks are performed as a DOM is copied to or from an object, so any mismatches will be identified and must be corrected.

IX. SIMULATOR RESTART, CHECKPOINT, AND RESTORE

Management of simulator reset/restart, checkpoint, and restore to a saved checkpoint is typically a troublesome problem for the implementer of any VPI/DPI-based package. If

(* `svLibSerialize` *) attribute and automatically adding them to the DOM. Unfortunately, simulator VPI support for attributes on class members is very patchy and immature, and we were forced reluctantly to abandon this approach.

the package maintains any state in C memory, that state must be cleared on restart (not especially difficult) and saved and restored across simulator checkpoint/restore (much more difficult). Our solution is to manage all package state in SystemVerilog where it is automatically saved and restored by the simulator's built-in mechanisms.

However, keeping all state on the SystemVerilog side can impose a performance penalty because it may mean that various data objects need to be copied repeatedly across the DPI boundary. To mitigate this cost, we have implemented cached state in C, shadowing the SystemVerilog package state. Because all interaction across the DPI boundary is performed by hidden private functions and not by user-callable code, we can do this with confidence.

On any simulator restart – whether to time-zero reset or to a saved checkpoint – all memory managed by our C code is deallocated, effectively wiping our state cache. Subsequent library function calls will then suffer the overhead of rebuilding the cache, but in every other way the memory management is completely invisible to the user.

Because this memory caching is completely invisible to users, we are able to use it more and more aggressively over successive implementations without risk of breaking user-visible functionality. Our initial implementation has minimal caching, and consequently had suboptimal performance, but is fully functional and has allowed us to conduct usability tests. In due course, some use of the cache mechanism will be implemented for those functions (notably regular expression processing and the YAML file reader) where it is obviously beneficial.

X. IMPLEMENTATION, TESTING AND DOCUMENTATION

As already noted, the implementation of svLib was generally trouble-free; by far the most challenging aspect of its development was to establish a convenient user-facing API that is natural for SystemVerilog programmers to use. The API and other features were established iteratively over extensive discussions with many colleagues participating.

A. Testing

We aimed to use a test-driven code development methodology. The SVUnit framework [10] is ideally suited to the testing of classes and packages containing many self-contained functions that can be tested individually, and we used it from the outset. In most cases we were able to construct at least a simple test harness for each new feature before it was developed, and indeed the SVUnit test cases provided a very useful sanity check that we had a sensible API in each case. As an example, here is a fragment of our unit test for the string "find substring" functions `first` and `last`. These methods have the prototype

```
int first(string sought, int ignore=0)
```

where `sought` is the string to search for, and `ignore` is the number of characters at the start of the string (or at the end, for `last()`) that should be skipped-over by the search. Their result is the character position, in the original string, of the leftmost character of the sought substring.

```
string sought;
my_Str.set("012345678901234567890");
// Seek something that's not in the string
sought = "A";
`FAIL_UNLESS_EQUAL(my_Str.first(sought), -1)
sought = "1235";
`FAIL_UNLESS_EQUAL(my_Str.first(sought), -1)

// Find occurrences of a single character
sought = "0";
`FAIL_UNLESS_EQUAL(my_Str.first(sought), 0);
`FAIL_UNLESS_EQUAL(my_Str.first(sought, 1), 10);
`FAIL_UNLESS_EQUAL(my_Str.first(sought, 11), 20);
`FAIL_UNLESS_EQUAL(my_Str.last(sought), 20);
`FAIL_UNLESS_EQUAL(my_Str.last(sought, 1), 10);

// Find occurrences of a longer string
sought = "0123";
`FAIL_UNLESS_EQUAL(my_Str.first(sought), 0);
`FAIL_UNLESS_EQUAL(my_Str.last(sought), 10);
```

1) Rapid feedback on implementation success

This approach gave us very rapid feedback on trivial implementation bugs that might otherwise have been hard to locate. Each unit test runs very rapidly, giving a success/failure indication on each code change in less than twenty seconds even when run on all three popular simulators. With this ability to test every new change on all supported platforms in less time than it takes to get a cup of coffee, we were highly motivated to keep the codebase working correctly at all times. Implementation errors were very quickly pinned down to a single test case by SVUnit's simple and clear reporting mechanism.

2) Continuous sanity checking of the library API

Unit testing also provided us with an early indication of whether the user-facing API was convenient and sensible, because we were obliged to write test code using that API before it was implemented. Poor design decisions were quickly and effectively highlighted. For example, one of our early attempts had some functions with very large numbers of optional arguments (with default values). It quickly became clear that this was extremely error-prone for users, as it was very easy to supply a value to the wrong optional argument and get surprising behavior.

3) Unit tests as a documentation pool

Finally, the ever-growing suite of tests continues to provide a valuable collection of simple usage examples for the various functions, with the benefit that the result of each example is clearly documented by the test itself.

B. Documentation

It has become popular to use automatic document generators such as Doxygen [11] or NaturalDocs [12] for the documentation of packages of this kind. They have the obvious and important benefit that the documentation is automatically kept up to date with respect to changes in the code. However,

we have chosen to buck this trend and revert to traditional manually generated documentation. This seemingly retrogressive step was taken only after careful thought. Documentation generated from code and its comments tends to focus exclusively on the properties of individual code fragments down to the function level, whereas we believe a human reader is more likely to be interested in the overall rationale for the package, the relationships among its components, and any guiding principles that it follows as a whole. This information we believe is better captured in human-written material.

The overwhelming advantage of autogenerated documentation is found when the code base is changing rapidly, and complex relationships among (for example) classes in a hierarchy must be kept under review. For a package such as svLib whose user-facing behavior should be very stable, this advantage is much less telling.

XI. PERFORMANCE

Although svLib provides very convenient access to functionality that previously was hard to reach from SystemVerilog, that convenience would be useless if its performance was unacceptably slow. We have measured the performance cost of selected svLib features and we believe that it is acceptable, because performance-hungry operations such as file access and manipulation of large strings are likely to be used only infrequently.

As a convenient benchmark we tried a search-and-replace in which a 2000-character string was scanned using a regular expression that found 100 instances of a 10-character pattern, performing a replacement on each instance. The runtime cost of this operation, on a modest 64-bit Linux machine, was between 5 and 12 milliseconds for three different simulators (with default optimization in force in each case). This appears to be about a factor of 10 slower than the same operation using script languages such as Tcl or Perl. Approximately half of the time was attributable to the regular expression compilation and execution in the C library, and about half to the string substitution which is performed in native SystemVerilog.

In a second test we timed lookup of a file's modification time, as an example of an operation where most of the work is done by the C library and svLib overhead should be expected to be insignificant. This indeed proved to be the case, with svLib being competitive with script languages on this operation.

It is clear from our measurements that there are many performance optimizations that could be applied to svLib's string manipulations, particularly regular expressions, and we plan to investigate this in more detail when other aspects of svLib have matured.

XII. FUTURE WORK

We have an extensive "shopping list" of future work that we would like to add to svLib as resources allow, and if there appears to be any demand for them. Among them we should mention:

- Full support for the Microsoft Windows® platform

- Support for Comma Separated Value (CSV) tabular data files
- Support for SQL access to a database
- Some way to associate an already-open C file handle with a Verilog file identifier. With this mechanism in place, a wide range of interesting new functionality would be available including better support for temporary files, and communication through TCP/IP sockets and other network mechanisms.
- A huge range of exciting possibilities opens up if it is possible to establish a tight link between SystemVerilog and a script language such as Python. This idea has already been explored in a slightly different way by the PyHVL project [13] and we have hardly begun to consider the possible interactions that are possible in this area.

XIII. CONCLUSIONS

An initial release of svLib is available for download at the time of publication, at **URL to be added before final publication**

We regard this as a beta-quality release: it has a full feature set, and we have made considerable use of it ourselves but it has not yet been stressed by heavy use in project work and therefore may lack maturity. At this time we welcome feedback from potential users, and we hope to use DVCon as an opportunity to gauge user appetite for such a library.

If svLib is received enthusiastically, we will begin to seek ways to engage the wider community in supporting it and carrying it forward. The authors welcome any feedback, both on the technical content of svLib and its usefulness in practice.

ACKNOWLEDGMENTS

The authors wish to thank the following for their contributions and support:

- Our employer Verilab consistently provides opportunity, support and encouragement for the development of ideas of this kind.
- Cliff Cummings of Sunburst Design, Inc. provided supportive and thorough reviews.
- We have many Verilab colleagues whose constructive reviews, practical support, deep expertise, high expectations, and relentless teasing are a source of inspiration and delight. You know who you are.
- Dr Gordon McGregor is a former colleague now at Nitero Inc. Gordon's commitment to the highest standards of quality and good sense in software development, his impatience with some of SystemVerilog's shortcomings, and his wise and insightful advice, were instrumental in giving svLib its current shape.
- The libyaml package[14], written by Kirill Simonov, saved us many days of implementation work.

- Bryan Morris of Verilab and Neil Johnson of XtremeEda Inc wrote and maintain the SVUnit testing framework, which encouraged us to adopt a test-driven development approach that greatly eased our implementation effort.

REFERENCES

- [1] D. Crockford, "JavaScript Object Notation," [Online]. Available: <http://tools.ietf.org/html/rfc4627>.
- [2] C. C. Evans, "The Official YAML Web Site," [Online]. Available: <http://www.yaml.org/>.
- [3] P. Hazel, "PCRE: Perl Compatible Regular Expressions," [Online]. Available: <http://pcre.sourceforge.net/>.
- [4] Accellera Systems Initiative, "UVM (Universal Verification Methodology)," [Online]. Available: <http://www.accellera.org/downloads/standards/uvm>.
- [5] IEEE, Standard 1800-2012 for SystemVerilog Hardware Design and Verification Language, New York, NJ: IEEE, 2012.
- [6] J. Bromley, "If SystemVerilog Is So Good, Why Do We Need the UVM?," in *FDL*, Paris, 2013.
- [7] Mentor Graphics, Inc., "OVM and UVM," [Online]. Available: <https://verificationacademy.com/topics/verification-methodology>.
- [8] Apache Software Foundation, "Apache License, version 2.0," 2004. [Online]. Available: <http://www.apache.org/licenses/LICENSE-2.0>.
- [9] Massachusetts Institute of Technology, "The MIT License," [Online]. Available: <http://opensource.org/licenses/MIT>.
- [10] N. Johnson and B. Morris, "AgileSoC SVUnit," [Online]. Available: <http://www.agilesoc.com/open-source-projects/svunit/>.
- [11] D. van Heesch, "Doxygen," [Online]. Available: <http://www.stack.nl/~dimitri/doxygen/index.html>.
- [12] G. Valure, "NaturalDocs," [Online]. Available: <http://www.naturaldocs.org>.
- [13] T. Loftus, J. Greene and B. Smith, "PyHVL," [Online]. Available: <http://pyhvl.sourceforge.net/>.
- [14] K. Simonov, "LibYAML," [Online]. Available: <http://pyyaml.org/wiki/LibYAML>.