

Prolog Unit Tests

HOME DOWNLOAD DOCUMENTATION TUTORIALS COMMUNITY USERS

WIKI

Documentation

Reference manual

Packages

SWI-Prolog Semantic Web Library Constraint Query Language A high I SWI-Prolog binding to GNU readlin SWI-Prolog ODBC Interface

SWI-Prolog binding to libarchive

Transparent Inter-Process Communi JPL: A bidirectional Prolog/Java into Pengines: Web Logic Programming SWI-Prolog SSL Interface

Google's Protocol Buffers Library SWI-Prolog Natural Language Proce

Prolog Unit Tests

Introduction

A Unit Test box

Using separate test files

Running the test-suite

Tests and production systems

Controlling the test suite

Auto-generating tests

Coverage analysis

Portability of the test-suite

Motivation of choices

SWI-Prolog Unicode library

SWI-Prolog YAML library

SWI-Prolog HTTP support

SWI-Prolog Regular Expression libr

Managing external tables for SWI-P

A C++ interface to SWI-Prolog

SWI-Prolog SGML/XML parser

SWI-Prolog binding to zlib

Paxos -- a SWI-Prolog replicating ke

SWI-Prolog Source Documentation SWI-Prolog C-library

SWI-Prolog binding to BSD libedit

SWI-Prolog RDF parser

Prolog Unit Tests

Jan Wielemaker University of Amsterdam VU University Amsterdam The Netherlands E-mail: jan@swi-prolog.org

Abstract

This document describes a Prolog unit-test framework. This framework was initially developed for SWI-Prolog. The current version also runs on SICStus Prolog, providing a portable testing framework. See section 9.1.

Table of Contents

- 1 Introduction
- 2 A Unit Test box
 - 2.1 Test Unit options
 - 2.2 Writing the test body
 - 2.2.1 Testing deterministic predicates
 - 2.2.2 Testing semi-deterministic predicates
 - 2.2.3 Testing non-deterministic predicates
 - 2.2.4 Testing error conditions
 - 2.2.5 One body with multiple tests using assertions
- 3 Using separate test files
- 4 Running the test-suite
- 5 Tests and production systems
- 6 Controlling the test suite
- 7 Auto-generating tests
- 8 Coverage analysis
- 9 Portability of the test-suite
 - 9.1 PlUnit on SICStus
- 10 Motivation of choices

1 Introduction

There is really no excuse not to write tests!

Automatic testing of software during development is probably the most important Quality Assurance measure. Tests can validate the final system, which is nice for your users. However, most (Prolog) developers forget that it is not just a burden during development.

- Tests document how the code is supposed to be used.
- Tests can validate claims you make on the Prolog implementation. Writing a test makes the claim explicit.
- Tests avoid big applications saying 'No' after modifications. This saves time during development, and it saves *a lot* of time if you must return to the application a few years later or you must modify and debug someone else's application.

2 A Unit Test box

Tests are written in pure Prolog and enclosed within the directives begin tests/1,2 and end tests/1. They

can be embedded inside a normal source module, or be placed in a separate test-file that loads the files to be tested. Code inside a test box is normal Prolog code. The entry points are defined by rules using the head test (Name) or test (Name, Options), where *Name* is a ground term and *Options* is a list describing additional properties of the test. Here is a very simple example:

The optional second argument of the test-head defines additional processing options. Defined options are:

blocked(+Reason:atom)

The test is currently disabled. Tests are flagged as blocked if they cannot be run for some reason. E.g. they crash Prolog, they rely on some service that is not available, they take too much resources, etc. Tests that fail but do not crash, etc. should be flagged using fixme (Fixme).

fixme(+*Reason:atom*)

Similar to blocked (Reason), but the test it executed anyway. If it fails, a - is printed instead of the . character. If it passes a + and if it passes with a choicepoint, !. A summary is printed at the end of the test run and the goal test_report(fixme) can be used to get details.

condition(:Goal)

Pre-condition for running the test. If the condition fails the test is skipped. The condition can be used as an alternative to the setup option. The only difference is that failure of a condition skips the test and is considered an error when using the setup option.

cleanup(:Goal)

Goal is always called after completion of the test-body, regardless of whether it fails, succeeds or throws an exception. This option or call_cleanup/2 must be used by tests that require side-effects that must be reverted after the test completes. Goal may share variables with the test body.

setup(:Goal)

Goal is run before the test-body. Typically used together with the cleanup option to create and destroy the required execution environment.

forall(:Generator)

Run the same test for each solution of *Generator*. Each run invokes the setup and cleanup handlers. This can be used to run the same test with different inputs. If an error occurs, the test is reported as name (forall bindings = $\langle vars \rangle$), where $\langle vars \rangle$ indicates the bindings of variables in *Generator*.

true(AnswerTerm Cmp Value)

Body must succeed deterministically. AnswerTerm is compared to Value using the comparison operator Cmp. Cmp is typically one of =/2, ==/2, =:=/2 or =@=/2, but any test can be used. This is the same as inserting the test at the end of the conjunction, but it allows the test engine to distinguish between failure of copy_term/2 and producing the wrong value. Multiple variables must be combined in an arbitrary compound term. E.g. A1-A2 == v1-v2

AnswerTerm Cmp Value(*E*)

quivalent to true (AnswerTerm Cmp Value) if Cmp is one of the comparison operators given above

fail

Body must fail.

throws(Error)

Body must throw Error. The error is verified using subsumes_chk(Error, Generated). I.e. the generated error must be more specific than the specified Error.

error(Error)

Body must throw error (Error, Context). See throws for details.

all(AnswerTerm Cmp Instances)

Similar to true (AnswerTerm Cmp Values), but used for non-deterministic predicates. Each element is compared using *Cmp*. Order matters. For example:

```
test(or, all(X == [1,2])) :-
( X = 1 ; X = 2 ).
```

set(AnswerTerm Cmp Instances)

Similar to all (AnswerTerm Cmp Instances), but before testing both the bindings of *AnswerTerm* and *Instances* are sorted using sort/2. This removes duplicates and places both sets in the same order.²

nondet

If this keyword appears in the option list, non-deterministic success of the body is not considered an error.

sto(Terms)

Declares that executing body is subject to occurs-check (STO). The test is executed with *Terms*. *Terms* is either rational_trees or finite_trees. STO programs are not portable between different kinds of terms. Only programs *not* subject to occurs-check (NSTO) are portable³. Fortunately, most practical programs are NSTO. Writing tests that are STO is still useful to ensure the robustness of a predicate. In case sto4 and sto5 below, an infinite list (a rational tree) is created prior to calling the actual predicate. Ideally, such cases produce a type error or fail silently.

Programs that depend on STO cases tend to be inefficient, even incorrect, are hard to understand and debug, and terminate poorly. It is therefore advisable to avoid STO programs whenever possible.

SWI's Prolog flag occurs check must not be modified within plunit tests.

2.1 Test Unit options

begin_tests(+Name)

Start named test-unit. Same as begin_tests(Name, []).

begin_tests(+Name, +Options)

Start named test-unit with options. Options provide conditional processing, setup and cleanup similar to individual tests (second argument of test/2 rules).

Defined options are:

```
blocked(+Reason)
```

Test-unit has been blocked for the given Reason.

```
condition(:Goal)
```

Executed before executing any of the tests. If *Goal* fails, the test of this unit is skipped.

setup(:Goal)

Executed before executing any of the tests.

cleanup(:Goal)

Executed after completion of all tests in the unit.

sto(+Terms)

Specify default for subject-to-occurs-check mode. See section 2 for details on the sto option.

2.2 Writing the test body

The test-body is ordinary Prolog code. Without any options, the body must be designed to succeed *deterministically*. Any other result is considered a failure. One of the options fail, true, throws, all or set can be used to specify a different expected result. See section 2 for details. In this section we illustrate typical test-scenarios by testing SWI-Prolog built-in and library predicates.

2.2.1 Testing deterministic predicates

Deterministic predicates are predicates that must succeed exactly once and, for well behaved predicates, leave no choicepoints. Typically they have zero or more input- and zero or more output arguments. The test goal supplies proper values for the input arguments and verifies the output arguments. Verification can use test-options or be explicit in the body. The tests in the example below are equivalent.

```
test(add) :-
    A is 1 + 2,
    A =:= 3.

test(add, [true(A =:= 3)]) :-
    A is 1 + 2.
```

The test engine verifies that the test-body does not leave a choicepoint. We illustrate that using the test below:

```
test(member) :-
member(b, [a,b,c]).
```

Although this test succeeds, member/2 leaves a choicepoint which is reported by the test subsystem. To make the test silent, use one of the alternatives below.

```
test(member) :-
    member(b, [a,b,c]), !.

test(member, [nondet]) :-
    member(b, [a,b,c]).
```

2.2.2 Testing semi-deterministic predicates

Semi-deterministic predicates are predicates that either fail or succeed exactly once and, for well behaved predicates, leave no choicepoints. Testing such predicates is the same as testing deterministic predicates. Negative tests must be specified using the option fail or by negating the body using \+/1.

2.2.3 Testing non-deterministic predicates

Non-deterministic predicates succeed zero or more times. Their results are tested either using findall/3 or setof/3 followed by a value-check or using the all or set options. The following are equivalent tests:

```
test(member) :-
    findall(X, member(X, [a,b,c]), Xs),
    Xs == [a,b,c].

test(member, all(X == [a,b,c])) :-
    member(X, [a,b,c]).
```

2.2.4 Testing error conditions

Error-conditions are tested using the option throws (Error) or by wrapping the test in a catch/3. The following tests are equivalent:

```
test(div0) :-
    catch(A is 1/0, error(E, _), true),
    E = @ = evaluation_error(zero_divisor).

test(div0, [error(evaluation_error(zero_divisor))]) :-
    A is 1/0.
```

2.2.5 One body with multiple tests using assertions

PlUnit is designed to cooperate with the assertion/1 test provided by library(debug). If an assertion fails in the context of a test, the test framework reports this and considers the test failed, but does not trap the debugger. Using assertion/1 in the test-body is attractive for two scenarios:

- Confirm that multiple claims hold. Where multiple claims about variable bindings can be tested using the == option in the test header, arbitrary boolean tests, notably about the state of the database, are harder to combine. Simply adding them in the body of the test has two disadvantages: it is less obvious to distinguish the tested code from the test and if one of the tests fails there is no easy way to find out which one.
- Testing scenarios or sequences of actions. If one step in such a sequence fails there is again no easy way to find out which one. By inserting assertions into the sequence this becomes obvious.

Below is a simple example, showing two failing assertions. The first line of the failure message gives the test. The second reports the location of the assertion.⁵ If the assertion call originates from a different file this is reported appropriately. The last line gives the actually failed goal.

```
:- begin_tests(test).

test(a) :-
    A is 2^3,
    assertion(float(A)),
    assertion(A == 9).
```

```
?- run_tests.

% PL-Unit: test
ERROR: /home/jan/src/pl-devel/linux/t.pl:5:
    test a: assertion at line 7 failed
    Assertion: float(8)
ERROR: /home/jan/src/pl-devel/linux/t.pl:5:
    test a: assertion failed
    Assertion: 8==9
. done
% 2 assertions failed
```

3 Using separate test files

Test-units can be embedded in normal Prolog source-files. Alternatively, tests for a source-file can be placed in another file alongside the file to be tested. Test files use the extension .plt. The predicate load_test_files/1 can load all files that are related to source-files loaded into the current project.

4 Running the test-suite

At any time, the tests can be executed by loading the program and running run_tests/0 or run_tests(+Unit).

run_tests

Run all test-units.

run tests(+Spec)

Run only the specified tests. *Spec* can be a list to run multiple tests. A single specification is either the name of a test unit or a term *<Unit>:<Tests>*, running only the specified test. *<Tests>* is either the name of a test or a list of names. Running particular tests is particularly useful for tracing a test:⁶

```
?- gtrace, run_tests(lists:member).
```

To identify nonterminating tests, interrupt the looping process with *Control-C*. The test name and location will be displayed.

5 Tests and production systems

Most applications do not want the test-suite to end up in the final application. There are several ways to achieve this. One is to place all tests in separate files and not to load the tests when creating the production environment. Alternatively, use the directive below before loading the application.

```
:- set_test_options([load(never)]).
```

6 Controlling the test suite

set_test_options(+Options)

Defined options are:

load(+Load)

Determines whether or not tests are loaded. When never, everything between begin_tests/1 and end_tests/1 is simply ignored. When always, tests are always loaded. Finally, when using the default value normal, tests are loaded if the code is not compiled with optimisation turned on.

run(+Run)

Specifies when tests are run. Using manual, tests can only be run using run_tests/0 or run_tests/1. Using make, tests will be run for reloaded files, but not for files loaded the first time. Using make (all) make/0 will run all test-suites, not only those that belong to files that are reloaded.

silent(+Bool)

When true (default is false), send informational messages using the 'silent' level. In practice this means there is no output except for errors.

 $\mathsf{sto}(+Bool)$

When true (default false), assume tests are not subject to occurs check (non-STO) and verify this if the Prolog implementation supports testing this.

load_test_files(+Options)

Load .plt test-files that belong to the currently loaded sources.

running tests

Print all currently running tests to the terminal. It can be used to find running thread in multi-

threaded test operation or find the currently running test if a test appears to be blocking.

test report(+What)

Print report on the executed tests. What defines the type of report. Currently this only supports fixme, providing details on how the fixme-flagged tests proceeded.

Auto-generating tests

Prolog is an interactive environment. Where users of non-interactive systems tend to write tests as code, Prolog developers tend to run queries interactively during development. This interactive testing is generally faster, but the disadvantage is that the tests are lost at the end of the session. The test-wizard tries to combine the advantages. It collects toplevel queries and saves them to a specified file. Later, it extracts these queries from the file and locates the predicates that are tested by the queries. It runs the query and creates a test clause from the query.

Auto-generating test cases is experimentally supported through the library library (test wizard). We briefly introduce the functionality using examples. First step is to log the queries into a file. This is accomplished with the commands below. Queries.pl is the name in which to store all queries. The user can choose any filename for this purpose. Multiple Prolog instances can share the same name, as data is appended to this file and write is properly locked to avoid file corruption.

```
:- use_module(library(test_wizard)).
  set_prolog_flag(log_query_file,
```

Next, we will illustrate using the library by testing the predicates from library library (lists). To generate test cases we just make calls on the terminal. Note that all queries are recorded and the system will select the appropriate ones when generating the test unit for a particular module.

```
?- member(b, [a,b]).
?- reverse([a,b], [b|A]).
A = [a];
```

Now we can generate the test-cases for the module list using make tests/3:

```
?- make tests(lists, 'Queries.pl', current_output).
:- begin_tests(lists).
test(member, [nondet])
member(b, [a, b]).

test(reverse, [true(A==[a])]):

reverse([a, b], [b|A]).
:- end_tests(lists).
```

8 Coverage analysis

An important aspect of tests is to know which parts of program is used (covered) by the tests. An experimental analysis is provided by the library library (test cover).

show_coverage(:Goal)

Run Goal and write a report on which percentage of the clauses in each file are used by the program and which percentage of the clauses always fail.

show coverage(:Goal, +Modules)

As show coverage/1, but only report on modules in the list *Modules*.

We illustrate this here using CHAT, a natural language question and answer application by David H.D. Warren and Fernando C.N. Pereira.

```
1 ?- show coverage(test chat).
Chat Natural Language Question Answering Test
______
                                                              Coverage by File
File
                                                                                                                                         %Cov %Fail
                                                                                                              Clauses
/staff/jan/lib/prolog/chat/xgrun.pl
/staff/jan/lib/prolog/chat/newg.pl
/staff/jan/lib/prolog/chat/newg.pl
/staff/jan/lib/prolog/chat/newdic.pl
/staff/jan/lib/prolog/chat/slots.pl
/staff/jan/lib/prolog/chat/slots.pl
/staff/jan/lib/prolog/chat/templa.pl
/staff/jan/lib/prolog/chat/tplan.pl
/staff/jan/lib/prolog/chat/talkr.pl
/staff/jan/lib/prolog/chat/ndtabl.pl
/staff/jan/lib/prolog/chat/world0.pl
/staff/jan/lib/prolog/chat/rivers.pl
/staff/jan/lib/prolog/chat/rivers.pl
/staff/jan/lib/prolog/chat/cities.pl
/staff/jan/lib/prolog/chat/countr.pl
                                                                                                                                                             0.0
                                                                                                                                          100.0
                                                                                                                          186
                                                                                                                                            89.2
                                                                                                                                                          18.3
                                                                                                                         28
275
                                                                                                                                            89.3
                                                                                                                                                             0.0
                                                                                                                                            35.6
74.2
                                                                                                                                                             0.0
                                                                                                                          128
                                                                                                                                                             1.6
                                                                                                                                            70.5
                                                                                                                                                             3.0
                                                                                                                          67
106
                                                                                                                                            55.2
75.5
                                                                                                                                                             1.5
                                                                                                                             60
                                                                                                                                            20.0
                                                                                                                             42
                                                                                                                                            59.5
                                                                                                                                                             0.0
                                                                                                                                            48.9
                                                                                                                                             71.8
                                                                                                                                                             1.5
                                                                                                                             41
                                                                                                                                          100.0
                                                                                                                                                             0.0
```

43.4

100.0

0.0

Using ?- show_coverage(run_tests)., this library currently only shows some rough quality measure for test-suite. Later versions should provide a report to the developer identifying which clauses are covered, not covered and always failed.

9 Portability of the test-suite

One of the reasons to have tests is to simplify migrating code between Prolog implementations. Unfortunately creating a portable test-suite implies a poor integration into the development environment. Luckily, the specification of the test-system proposed here can be ported quite easily to most Prolog systems sufficiently compatible to SWI-Prolog to consider porting your application. Most important is to have support for term_expansion/2.

In the current system, test units are compiled into sub-modules of the module in which they appear. Few Prolog systems allow for sub-modules and therefore ports may have to fall-back to inject the code in the surrounding module. This implies that support predicates used inside the test unit should not conflict with predicates of the module being tested.

9.1 PlUnit on SICStus

The directory of plunit.pl and swi.pl must be in the library search-path. With PLUNITDIR replaced accordingly, add the following into your .sicstusrc or sicstus.ini.

```
:- set_prolog flag(language, iso). % for maximal compatibility library_directory('PLUNITDIR').
```

The current version runs under SICStus 3. Open issues:

- Some messages are unformatted because SICStus 3 reports all ISO errors as instantiation errors.
- Only plunit.pl. Both coverage analysis and the test generation wizard currently require SWI-Prolog.
- The load option normal is the same as always. Use set_test_options(load, never) to avoid loading the test suites.
- The run option is not supported.
- Tests are loaded into the enclosing module instead of a separate test module. This means that predicates in the test module must not conflict with the enclosing module, nor with other test modules loaded into the same module.

10 Motivation of choices

Easy to understand and flexible

There are two approaches for testing. In one extreme the tests are written using declarations dealing with setup, cleanup, running and testing the result. In the other extreme a test is simply a Prolog goal that is supposed to succeed. We have chosen to allow for any mixture of these approaches. Written down as test/1 we opt for the simple succeeding goal approach. Using options to the test the user can choose for a more declarative specification. The user can mix both approaches.

The body of the test appears at the position of a clause-body. This simplifies identification of the test body and ensures proper layout and colouring support from the editor without the need for explicit support of the unit test module. Only clauses of test/1 and test/2 may be marked as non-called in environments that perform cross-referencing.

Index

```
assertion/1
2.2.5 2.2.5
begin_tests/1
2 6
begin_tests/2
call_cleanup/2
2
catch/3
2.2.4
copy_term/2
```

```
2
end tests/1
      26
findall/3
      2.2.3
load test files/1
make/0
make tests/3
      7
member/2
      2.2.1
run tests/0
      4 6
run_tests/1
      6
running tests/0
set_test_options/1
setof/3
      2.2.3
show coverage/1
      8
show coverage/2
sort/2
term expansion/2
test/1
      10 10
test/2
      2.1 10
test report/1
variant/2
      2
```

Tag confusing pages with doc-needs-help|Tags are associated to your profile if you are logged in|Report abuse

Tags:

doc-needs-help ×

Robert Laing said (2018-11-23T15:17:03):

 $\triangle_1 \nabla$

I got stuck at =load test files (+Options) = trying to use the plt filename as the argument.

 $I\ discovered\ I\ should\ use = \verb|load_test_files|\ ([make(all)]). = thanks\ to\ Google\ leading\ me\ to\ a\ nice\ tutorial\ at\ http://www.paulbrownmagic.com/blog/swi_prolog_unit_testing_env$

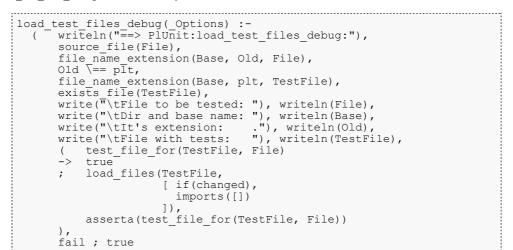


ktp said (2019-11-11T10:24:16):



Hello!

I suggest changing the 'load_test_files / 1' function in the 'plunit' library or adding a new one - e.g. 'load_test_files_debug / 1' with the body:



The only change is the addition of 5 lines with write/writeln predicates which write to the stdout name(s) of tested `.pl' file(s) and used `.plt' file(s).



It helped me a lot while debugging my scripts.

(of course when new function is created it should be added to exported functions list)

Hans Nowak said (2017-12-03T21:25:50):





"In the current system, test units are compiled into sub-modules of the module in which they appear."

It is not stated here explicitly, but apparently this affects the behavior of assertz/1, retractall/1, and the lookup of any predicates defined using assert{a,z}? I ran into issues with calls to retractall/1 inside tests that (seemingly) didn't actually retract anything, and predicates defined with assertz/l inside tests that were then not found by other predicates (defined elsewhere). Such behavior makes more sense though if you know that tests run in their own (sub)module. Which is indeed stated here, but IMHO it's not clear what it actually means in practice.

(The retractall/1 inside the test didn't remove predicates defined in the enclosing module (in this case, user); similarly, predicates defined with assertz/1 inside the test, were not found by a predicate defined in user. This did not happen when I ran the exact same code outside of the tests.)

LogicalCaptain said (2016-12-16T02:03:35):





Genius!

The index is messed up though



login to add a new annotation post.

Powered by SWI-Prolog 8.1.28