On Improving Run-time Checking in Dynamic Languages

Nataliia Stulova\footnote{Research supported in part by projects EU FP7 318337 ENTRA, Spanish MINECO TIN2012-39391 StrongSoft, TIN2015-67522-C3-1-R TRACES and TIN2008-05624 DOVES, and Comunidad de Madrid TIC/1465 PROMETIDOS-CM and S2013/ICE-2731 N-Greens Software, and Madrid Region program M141047003 N-GREENS.} \footnote{Supervised by José F. Morales and Manuel V. Hermenegildo.}

1 IMDEA Software Institute, Madrid, Spain
nataliia.stulova@imdea.org
2 Universidad Politécnica de Madrid (UPM), Madrid, Spain

Abstract

In order to detect incorrect program behaviors, a number of approaches have been proposed, which include a combination of language-level constructs (procedure-level annotations such as assertions/contracts, gradual types, etc.) and associated tools (such as static code analyzers and run-time verification frameworks). However, it is often the case that these constructs and tools are not used to their full extent in practice due to a number of limitations such as excessive run-time overhead and/or limited expressiveness. This issue is especially prominent in the context of dynamic languages without an underlying strong type system, such as Prolog. In our work we propose several practical solutions for minimizing the run-time overhead associated with assertion-based verification while keeping the correctness guarantees provided by run-time checks. We present the solutions in the context of the Ciao system, where a combination of an abstract interpretation-based static analyzer and run-time verification framework is available, although our proposals can be straightforwardly adapted to any other similar system.

1998 ACM Subject Classification D.3.3 Language Constructs and Features, D.2.4 Software/Program Verification, D.2.5 Testing and Debugging, F.3.1 Specifying and Verifying and Reasoning about Programs, F.3.3 Studies of Program Constructs, F.3.2 Semantics of Programming Languages

Keywords and phrases Runtime Verification, Assertions, Prolog, Logic Programming


1 Introduction

Detecting incorrect program behaviors is an important part of the software development life cycle. It is also a complex and tedious one, in which dynamic languages bring special challenges.

A number of techniques have been proposed to aid in the process, among which we center our attention on the use of language-level constructs to describe expected program behavior, and of associated tools to compare actual program behavior against expectations, such as static code analyzers/verifiers and run-time verification frameworks.

Approaches that fall into this category are the assertion-based frameworks used in (Constraint) Logic Programming \cite{10, 24, 16, 19}, soft/gradual typing approaches in functional...
programming [5, 34, 32] and contract-based extensions in object-oriented programming [17, 18, 11]. These tools are aimed at detecting violations of the expected behavior or certifying the absence of any such violations, and often involve a certain degree of run-time testing, specially for non-trivial properties.

In practice, however, run-time overhead often remains impractically high, specially for complex properties, such as, for example, deep data structure tests. This reduces the attractiveness of run-time checking to programmers, which may allow sporadic checking of very simple conditions, but tend to turn off run-time checking for more complex properties. Some approaches even opt for limiting the expressiveness of the assertion language in order to reduce the overhead.

Our research objective is twofold:

1. First, we aim for enhancing the expressiveness of the assertion language to reflect all the features of the related programming language, including, e.g., higher-order constructs, and to do so in a way that allows the programmer to write precise program specifications while not imposing a learning or programming burden on them.

2. At the same time, our goal is to efficiently check such specifications, mitigating the associated run-time overhead as much as possible without compromising the safety guarantees that the checks provide.

While our work is general and system-independent, we present it for concreteness in the context of the Ciao run-time checking framework. The Ciao model [13, 24] is well understood, and different aspects of it have been incorporated in popular (C)LP systems, such as Ciao, SWI, and XSB [14, 31, 21].

### 2 Current Research Results

#### 2.1 Supporting Higher-Order Properties

Higher-order programming is a widely adopted programming style that adds flexibility to the software development process. Within the (Constraint) Logic Programming ((C)LP) paradigm, Prolog has included higher-order constructs since the early days, and there have been many other proposals for combining the first-order kernel of (C)LP with different higher-order constructs, e.g., [35, 23, 4]). Many of these proposals are currently in use in different (C)LP systems and have been found very useful in programming practice, inheriting the well-known benefits of code reuse (templates), elegance, clarity, and modularization.

When higher-order constructs are introduced in the language it becomes necessary to describe properties of arguments of predicates/procedures that are themselves also predicates/procedures. While the combination of contracts and higher-order has received some attention in functional programming [12, 9], within (C)LP the combination of higher-order with the previously mentioned assertion-based approaches has received comparatively little attention to date. Current Prolog systems simply use basic atomic types (i.e., stating simply that the argument is a `pred`, `callable`, etc.) to describe predicate-bearing variables (see Fig. 1). Other approaches [1] are more oriented instead towards meta programming, describing meta-types but there is no notion of directionality (modes), and only a single pattern is allowed per predicate.

Our proposal [26] contributes towards filling this gap between higher-order (C)LP programs and assertion-based extensions for error detection and program validation. Our starting point is the Ciao assertion model, which we have enhanced with a new class of properties, “predicate properties” (or `predprops`), for which we have proposed a syntax and
\[ \text{min}(X,Y,P, \text{Min}) : \text{P}(R,X,Y), R \leq 0, \text{Min} = X. \]

less(0,A,A). \quad \text{lt}(\text{=},A,A).

less(-1,A,B) :- A < B. \quad \text{lt}(\text{<},A,B) :- A < B.

less(1,...). \quad \text{lt}(\text{>},...,...).

\[
\text{test_min} :- \text{min}(4,2,\text{lt},2). \quad \% \text{lt}/3 \text{ is passed, but less}/3 \text{ is expected}
\]

\[ \text{comparator}(\text{Cmp}) \{
\text{pred Cmp}(\text{Res}, M, N) : (\text{num}(M), \text{num}(N)) \Rightarrow \text{between}(-1,1,\text{Res}).
\}\]

\[ \text{pred min}(X,Y,\text{Cmp}, \text{Min}) : \text{callable}(\text{Cmp}). \]

\[
\text{min}(X,Y,P, \text{Min}) : \text{callable}(\text{Cmp}).
\]

\text{lt}(\text{=},A,A).

\text{lt}(\text{<},A,B) :- A < B.

\text{lt}(\text{>},...,...).

\[ \text{test_min} :- \text{min}(4,2,\text{lt},2). \quad \% \text{lt}/3 \text{ is passed, but less}/3 \text{ is expected} \]

\[ \text{comparator}(\text{Cmp}) \{
\text{pred Cmp}(\text{Res}, M, N) : (\text{num}(M), \text{num}(N)) \Rightarrow \text{between}(-1,1,\text{Res}).
\}\]

\[ \text{pred min}(X,Y,\text{Cmp}, \text{Min}) : \text{comparator}(\text{Cmp}). \]

\[ \text{comparator}(\text{Cmp}) \{
\text{pred Cmp}(\text{Res}, M, N) : (\text{num}(M), \text{num}(N)) \Rightarrow \text{between}(-1,1,\text{Res}).
\}\]

\text{lt}(\text{=},A,A).

\text{lt}(\text{<},A,B) :- A < B.

\text{lt}(\text{>},...,...).

\[ \text{test_min} :- \text{min}(4,2,\text{lt},2). \quad \% \text{lt}/3 \text{ is passed, but less}/3 \text{ is expected} \]

\text{comparator}(\text{Cmp}) \{
\text{pred Cmp}(\text{Res}, M, N) : (\text{num}(M), \text{num}(N)) \Rightarrow \text{between}(-1,1,\text{Res}).
\}\]

\[ \text{pred min}(X,Y,\text{Cmp}, \text{Min}) : \text{comparator}(\text{Cmp}). \]

**Figure 1** A simple program with a higher-order predicate \textit{min}/4 that accepts a custom comparator predicate.

**Figure 2** A \textit{predprop} \textit{comparator} example and an \textit{anonymous} assertion in its definition.

The new properties can be used in assertions for higher-order predicates to describe the properties of the higher-order arguments. An example of a \textit{predprop} is provided in Fig. 2, where an \textit{anonymous} assertion (note the variable symbol \text{Cmp} in place of a predicate symbol) is used to describe a comparison predicate, that is not known at compilation time. By reusing the original assertion language syntax to describe call and success conditions of predicate-bearing arguments we allow both for better integration of these new constructs into the verification framework and at the same time lessening the burden on a programmer, who needs to provide such annotations.

Our \textit{predprop} properties specify conditions for predicates that are independent of the usage context. This corresponds in functional programming to the notion of \textit{tight} contract satisfaction [9], and it contrasts with alternative approaches such as \textit{loose} contract satisfaction [12]. In the latter, contracts are attached to higher-order arguments by implicit function wrappers. The scope of checking is local to the function evaluation. Although this is a reasonable and pragmatic solution, we believe that our approach is more general and more amenable to combination with static verification techniques. For example, avoiding wrappers allows us to remove checks (e.g., by static analysis) without altering the program semantics. Moreover, our approach can easily support \textit{loose} contract satisfaction, since it is straightforward in our framework to optionally include wrappers as special \textit{predprops}.

### 2.2 Trading Memory for Speed

While having become an integral part of software development process, run-time testing can generally incur a high penalty in execution time and/or space over the standard, test-less program execution. A number of techniques have been proposed to date to reduce this overhead, including simplifying the checks at compile time via static analysis [2, 13] or reducing the frequency of checking, including for example testing only at a reduced number of points [19, 20]. Our proposal of [27] describes an approach to run-time testing that is efficient while being minimally obtrusive and remaining exhaustive. It is based on the use of memoization to cache intermediate results of check evaluation in order to avoid repeated checking of previously verified properties over the same data structure.

Memoization has of course a long tradition in (C)LP in uses such as tabling resolution [33,
Figure 3 A minimalistic tree data structure implementation as a regular type tree/1

Memoization has also been used in program analysis [36, 22], where tabling resolution is performed using abstract values. However, in tabling and program analysis it is call-success patterns that are usually tabled, whereas in our case the aim is to cache the results of test execution.

We concentrate our attention on checks for conformance of run-time heap structures to regular types [7], a useful subset of properties that are often used in assertions. An example in Fig. 3 shows a binary tree (left) and its possible implementation as a regular type together with an instance of that type describing the tree (right).

Our approach is based on the observation that run-time checks of regular types are monotonic instantiation checks, i.e., terms only become more and more instantiated with every subsequent state that the program enters.\(^1\) We extend the Ciao run-time checking framework with a common cache, accessible to run-time checks, that stores tuples \((x, t)\), where \(x\) is a term address and \(t\) is an identifier of a regular type. After the initial check on the term is performed, the corresponding tuple is added to the cache and for all the subsequent checks on the term, the check is replaced by a (computationally cheaper) cache lookup operation.

While studying the resulting performance of the enhanced verification framework we have observed that a straightforward use of a cache does not necessarily lead to a significant reduction in checking cost. An important issue that must be taken into account is the character of cache rewrites: as the terms grow in size cache collisions happen more often, which leads to element eviction. This in turn cancels out the advantage that using a cache gives: the ability to just lookup the regular type of some term without actually performing the check of it. However, this effect can be remedied by limiting the depth of terms that are stored in the cache (e.g., not caching terms with depth more than some \(n\)).

The idea of using memoization techniques to speed up checks has attracted some attention recently [15]. Their work (developed independently from ours) is based on adding fields to data structures to store the properties that have been checked already for such structures. In contrast, our approach has the advantage of not requiring any modifications to data structure representation, or to the checking code, program, or core run-time system.

2.3 Intertwining Compile- and Run-time Checks

A complementary approach to run-time overhead reduction consists in using static analysis to minimize the number and cost of the run-time checks that need to be placed in the program to detect incorrect program behaviors. This idea was pioneered by the Ciao system where a number of (abstract interpretation-based) static analyses are combined in order to verify

---

\(^1\) This is not the case of course if the language allows mutable variables and they are used in the program, but in those cases variable inmutability is tracked by the compiler / preprocessor.
assertions to the largest extent possible at compile time, and for simplifying and reducing
the number of remaining properties that need to be introduced in the program as runtime
checks. However, while there has been evidence from use, there has been little systematic
experimental work presented to date measuring the actual impact of analysis on reducing
run-time checking overhead.

In our work [28] we generalize the existing practices as four assertion checking modes,
each of which represents a trade-off between code annotation depth, execution time slowdown,
and program behavior safety guarantees:

- **Unsafe**: no run-time checks are generated from program assertions, program execution
  is fast but incorrect program behaviors may stay undetected; the run-time overhead is
  nonexistent.
- **Client-Safe**: run-time checks are generated from the assertions of the program’s interface
  (providing behavior guarantees for its the clients), yet internal program assertions remain
  unchecked; the run-time overhead is taken as a minimal unavoidable one.
- **Safe-RT**: run-time checks are generated from all program assertions; this mode of checking
  is characterized with the strongest behavior safety guarantees yet is at the same time
  associated with highest check costs;
- **Safe-CT-RT**: a variation of Safe-RT checking mode with an additional static analysis
  phase, during which some of program assertions may be proven to always hold and
  generating run-time checks from them can be omitted; depending on the kind of analysis
  and the program the overhead generated by the checks from remaining assertions may be
  closer to that of Client-Safe Safe-RT modes.

We also define a transformation-based approach in order to implement each one of these
modes. The differences between the transformations used in different modes are illustrated
in Fig. 4. Starting with the source of the program and its assertions in the bottom-left
corner, a sequence of preprocessing and compilation steps (denoted by arrows) is applied.
Respective components of the verification framework that perform the source transformations
are represented as hexagons.

We then concentrate on the reduction of the number of run-time tests via (abstract
interpretation-based) program analysis. To this end we propose a technique that enhances
analysis precision by taking into account that any assertions that cannot be proved statically
will be the subject of run-time testing. In practice it means that it is safe to make the two
following assumptions for any predicate that has an assertion specifying properties that
should hold on calls and success:

- The calls conditions hold after the analysis has entered the predicate definition, since
  either the checks for these calls conditions have already succeeded or the program has
  exited with error.
- The relevant success conditions hold after the predicate has exited, since, again, at this
  point either these success conditions have already succeeded or the program has exited
  with error.

2.4 Benefiting from Information Hiding

While dynamic languages offer programmers great flexibility in term creation and manip-
ulation, for the very same reason the need for exhaustive run-time checks arises in order
to guarantee the data manipulation safety and correctness. Reusable libraries, i.e., library
modules that are pre-compiled independently of the client, pose special challenges in this
context. The key issue here is that there is virtually no control on how and where valid
terms can be created, and thus it is quite common in the client-library interaction that any
of these modules can create any data shape and pass it. As a consequence, (often expensive) run-time checks on the module boundaries become a necessary evil. In our work [29] we propose a possible solution to overhead reduction for run-time checks on module boundaries based on the information hiding principle (which is adopted in many other systems in form of encapsulation or opaque data types).

Currently, most mature Prolog implementations adopt some flavor of a module system, predicate-based in SWI [37], SICStus [30], YAP [25], ECLiPSe [6], and atom-based in XSB [31]. The difference between the two systems is the strictness of the term visibility rules: in an atom-based system local to the module terms are not visible outside it if they are not a part of the module interface. The Ciao approach [3] has until now been closer to a predicate-based module system.

We propose an extension of the predicate-based module system, that allows to specify only a subset of module terms as local (hidden) ones. Our argument is that in this setup we have the guarantee of the module terms’ structural homogeneity, as only one module is allowed to construct/deconstruct some particular data shapes. With this there is no need to perform thorough checks of properties that verify the correctness of the data term structure at the module boundaries. Instead, it would suffice to perform checks of shallow versions of data shape properties: the weakened forms of the original properties that are semantically equivalent to them in the context of the possible program executions. These versions typically require asymptotically fewer execution steps and in some cases of the calls across module boundaries allow us to achieve constant run-time overhead.

To illustrate the idea behind the approach, let us consider that the tree/1 regular type from Fig. 3 is defined in a following module bintrees:

```
1 :- module(bintrees, [tree/1, add/3, del/3, find/2]). % exported predicates
2 :- hide e/0. % \_ hidden
3 :- hide t/3. % / functors
```
If we can assure by static analysis that none of the exported predicates of `bintrees` passes `t(_,_,_)` terms outside the module, which could allow module clients to construct arbitrary terms with this functor (e.g., `t([...],_,[...])`), then we can safely substitute the full `tree/1` property in the precondition check for `add/1` (and also in precondition checks for other exported predicates!) by its shallow version:

This way in all calls across module `bintrees` boundaries the cost of precondition checks goes down from linear in the size of the term to constant (functor correctness check).

Preliminary experimental results and details about the algorithms that perform the inference of shallow properties and conditions under which it is safe to use them in run-time checks are provided in a technical report, to which [29] refers to.

### 3 Conclusions and Future Work

The specification-based runtime verification approach has attracted significant interest in recent decades, both from academia and industry. As it is the case with any other open problem, finding an approach that would suit each and every system is hard, and thus opting for tailored partial solutions is more practical.

In our work we have concentrated on the peculiarities of the run-time verification task in the context of dynamic languages and their use in (C)LP systems. We have proposed an enhancement for the Ciao assertion language that allows us to capture the concrete execution contexts of higher-order terms and thus adapts the verification framework to this new use case. We have also proposed several solutions for reducing the overhead associated with run-time checks, that treat the issue from several different angles.

While for concreteness of presentation our work was carried out within the Ciao language and combined static/dynamic verification framework, our results are general and system-independent. We believe they can be straightforwardly transferred to the contexts of other declarative languages. In addition, given the advances in verification of a wide class of programming languages, including imperative languages, by translation into Horn clauses and proving properties at this level, and the fact that this approach is fully supported in the Ciao system, we argue that our results can easily be adapted to a much broader spectrum of languages.

For future work we plan to concentrate first on fully incorporating the optimization techniques described here into Ciao’s run-time verification framework, as now they are available as separate system bundles. Among the issues we would also like to address more profoundly are further developments on blame assignment in the case of higher-order assertion checking, scalability evaluation of the combination of all optimization techniques, and providing on-line demos and documentation.
On Improving Run-time Checking in Dynamic Languages

Acknowledgements. We would like to thank the anonymous reviewers for providing valuable comments and suggestions that had helped to improve this paper.

References


Edison Mera and Jan Wielemaker. Porting and refactoring Prolog programs: the PROSYN case study. TPLP, 13(4-5-Online-Supplement), 2013.


On Improving Run-time Checking in Dynamic Languages


