An algorithmic mitigation of large spurious interprocedural cycles in static analysis

Hakjoo Oh*† and Kwangkeun Yi

School of Computer Science and Engineering, Seoul National University, 599 Gwanak-ro Gwanak-gu, Seoul 151-742, Korea

SUMMARY
We present a simple algorithmic extension of the approximate call-strings approach to mitigate substantial performance degradation caused by spurious interprocedural cycles. Spurious interprocedural cycles are, in a realistic setting, key reasons for why approximate call-return semantics in both context-sensitive and -insensitive static analysis can make the analysis much slower than expected.

In the approximate call-strings-based context-sensitive static analysis, because the number of distinguished contexts is finite, multiple call-contexts are inevitably joined at the entry of a procedure and the output at the exit is propagated to multiple return-sites. We found that these multiple returns frequently create a single large cycle (we call it "butterfly cycle") covering almost all parts of the program and such a spurious cycle makes analyses very slow and inaccurate.

Our simple algorithmic technique (within the fixpoint iteration algorithm) identifies and prunes these spurious interprocedural flows. The technique’s effectiveness is proven by experiments with a realistic C analyzer to reduce the analysis time by 7%-96%. Since the technique is algorithmic, it can be easily applicable to existing analyses without changing the underlying abstract semantics, it is orthogonal to the underlying abstract semantics’ context-sensitivity, and its correctness is obvious.

KEY WORDS: Static Analysis; Interprocedural Analysis; Abstract Interpretation; Spurious Cycles; Fixpoint Algorithm
1. Introduction

In the approximate call-strings approach, which was proposed by Sharir and Pnueli [21], it is inevitable to follow some spurious (unrealizable or invalid) return paths. When the analysis uses a limited context information in which the number of distinguished contexts is finite, multiple call-contexts are inevitably joined at the entry of a procedure and the output at the exit are propagated to multiple return-sites. For example, in a conventional way of avoiding invalid return paths by distinguishing a finite \( k \geq 0 \) call-sites to each procedure [21], the analysis is doomed to still follow spurious paths if the input program’s nested call-depth is larger than the \( k \). Increasing the \( k \) to remove more spurious paths quickly hits a limit in practice because of the increasing analysis cost in memory and time.

In this article, which is an extended version of [15], we present the following:

- in a realistic setting, these multiple returns often create a single large flow cycle (we call it “butterfly cycle”) covering almost all parts of the program,
- such a big spurious cycle makes the approximate call-strings method [21] that distinguishes the last \( k \) call-sites very slow and inaccurate,
- this performance problem can be relieved by a simple extension of the call-strings method,
- our extension is an algorithmic technique within the worklist-based fixpoint iteration routine, without redesigning the underlying abstract semantics part
- the algorithmic technique works regardless of the underlying abstract semantics’ context-sensitivity (the \( k \)), and
- the technique also works regardless of the existing worklist ordering strategies of the fixpoint algorithm. The technique consistently saves the analysis time, without sacrificing (or with even improving) the analysis precision.

1.1. Problem: Large Performance Degradation By Inevitable, Spurious Interprocedural Cycles

Static analysis’ spurious paths make spurious cycles across procedure boundaries in global analysis. For example, consider the semantic equations in Figure 1 that (context-insensitively (\( k = 0 \))) abstract two consecutive calls to a procedure. The system of equations says to evaluate equation (4) and (6) for every return-site after analyzing the called procedure body (equation (3)). Thus, solving the equations follows a cycle: \( (2) \rightarrow (3) \rightarrow (4) \rightarrow (5) \rightarrow (2) \rightarrow \cdots \). Spurious cycles can also be created when \( k \geq 1 \). The following example describes how spurious cycles are created during the analysis for \( k = 1 \).

Example 1. The \( k \) length suffix method can be understood by applying intraprocedural analysis algorithm to the extended supergraph [12]. We first describe how an extended supergraph is created from the program. Assume that a program is represented by a supergraph [17] \( G = (N,E) \), which is a directed graph in which control flow graphs for procedures are connected according to the calling relationships between procedures. The extended supergraph \( G_E = (N_E,E_E) \) is a directed graph with \( N_E = \{ (n,c) \mid n \in N \text{ and } c \in pcs(n) \} \) where \( pcs(n) \) represents the set of possible call-strings for node \( n \). \( ((n_1,c_1),(n_2,c_2)) \in E_E \) iff \( (n_1,n_2) \in E \) and \( c_2 \) is the updated call string from \( c_1 \). In other words, \( G_E \) is a directed graph whose nodes
AN ALGORITHMIC MITIGATION OF LARGE SPURIOUS CYCLES IN STATIC ANALYSIS

Figure 1. Spurious dependence cycles because of abstract procedure calls and returns. The right-hand side is a system of equations for \( k = 0 \) and the left-hand side shows the dependences between the equations. Note a dependence cycle \((2) \rightarrow (3) \rightarrow (4) \rightarrow (5) \rightarrow (2) \rightarrow \cdots\).

are defined by pairs of nodes and their possible contexts and the edges explicitly shows the propagation paths of abstract values in context-sensitive manner.

Figure 2(a) shows an example of supergraph where procedure \( f \) is called twice from procedure \( m \) and \( g \) is called once from \( f \). Figure 2(b) shows its extended supergraph for \( k = 1 \). In Figure 2(b), since \( f \) is called by two times, each node of \( f \) has two separate contexts. But, since \( g \) is called only once, each node of \( g \) has only one context. Note that, even though the procedure \( g \) returns to a single return node (node 9), there are two paths which flow to the two different contexts, \( k \) and \( l \); these two contexts are due to the two different call sites (node 2 and 4). So the analysis follows a spurious cycle \( m \rightarrow c \rightarrow d \rightarrow h \rightarrow j \rightarrow o \rightarrow p \rightarrow k \rightarrow m \rightarrow \cdots\).

Such spurious cycles degrade the analysis performance both in precision and speed. Spurious cycles exacerbate the analysis imprecision because they model spurious information flow. Spurious cycles degrade the analysis speed too because solving cyclic equations repeatedly applies the equations in vain until a fixpoint is reached.

The performance degradation becomes dramatic when the involved interprocedural spurious cycles cover a large part of the input program. This is indeed the case in reality. In analyzing real C programs, we observed that the analysis follows (Section 2) a single large cycle that spans almost all parts of the input program. Such spurious cycles size can also be estimated by just measuring the strongly connected components (scc) in the “lexical”\(^\dagger\) control flow graphs. Table I shows the sizes of the largest scc in some open-source programs.\(^\dagger\) In most programs, such cycles cover most (80-90\%) parts of the programs. Hence, globally analyzing a program is likely to compute a fixpoint of a function that describes almost all parts of the input program.

---

\( \dagger \)One node per lexical entity, ignoring function pointers.

\( \dagger \)We measured the sizes of all possible cycles in the flow graphs. Note that interprocedural cycles happen because of either spurious returns or recursive calls. Because recursive calls in the test C programs are immediate or spans only a small number of procedures, large interprocedural cycles are likely to be spurious ones.
Even when we do the call-strings-based context-sensitive analysis ($k > 0$), large spurious cycles are likely to remain (Section 2).

1.2. Solution: An Algorithmic Mitigation Without Redesigning the Analysis (Abstract Semantics)

We present a simple algorithmic technique inside a worklist-based fixpoint iteration procedure that, without redesigning the abstract semantics part, can effectively relieve the performance degradation caused by spurious interprocedural cycles in both call-strings-based context-sensitive ($k > 0$) and -insensitive ($k = 0$) analysis. For the moment, we consider context-insensitive case only. We extend it to context-sensitive analysis in Section 3.

While solving flow equations, the algorithmic technique simply forces procedures to return to their corresponding called site, in order not to follow the last edge (edge (3) → (4) in Figure 1) of the “butterfly” cycles. In order to enforce this, we control the equation-solving orders so that each called procedure is analyzed exclusively for its one particular call-site. To be safe, we apply our algorithm to only non-recursive procedures.
Table I. The sizes of the largest strongly-connected components in the “lexical” control flow graphs of real C programs. In most cases, most procedures and nodes in program belong to a single cycle.

<table>
<thead>
<tr>
<th>Program</th>
<th>Procedures in the largest cycle</th>
<th>Basic-blocks in the largest cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>spell-1.0</td>
<td>24/31 (77%)</td>
<td>751/782 (95%)</td>
</tr>
<tr>
<td>gzip-1.2.4a</td>
<td>100/135 (74%)</td>
<td>5,988/6,271 (95%)</td>
</tr>
<tr>
<td>sed-4.0.8</td>
<td>230/294 (78%)</td>
<td>14,559/14,976 (97%)</td>
</tr>
<tr>
<td>tar-1.13</td>
<td>205/222 (92%)</td>
<td>10,194/10,800 (94%)</td>
</tr>
<tr>
<td>wget-1.9</td>
<td>346/434 (80%)</td>
<td>15,249/16,544 (92%)</td>
</tr>
<tr>
<td>bison-1.875</td>
<td>410/832 (49%)</td>
<td>12,558/18,110 (69%)</td>
</tr>
<tr>
<td>proftpd-1.3.1</td>
<td>940/1,096 (85%)</td>
<td>35,386/41,062 (86%)</td>
</tr>
<tr>
<td>apache-2.2.2</td>
<td>1,364/2,075 (66%)</td>
<td>71,719/95,179 (75%)</td>
</tr>
</tbody>
</table>

Consider the equation system in Figure 1 again and think of a middle of the analysis (equation-solving) sequence, \( \cdots \rightarrow (5) \rightarrow (2) \rightarrow (3) \), which indicates that the analysis of procedure \( f \) is invoked from (5) and is now finished. After the evaluation of (3), a classical worklist algorithm inserts all the equations, (4) and (6), that depend on (3). But, if we remember the fact that \( f \) has been invoked from (5) and the other call-site (1) has not invoked the procedure until the analysis of \( f \) finishes, we can know that continuing with (4) is useless, because the current analysis of \( f \) is only related to (5), but not to other calls like (1). So, we process only (6), pruning the spurious sequence \( (3) \rightarrow (4) \rightarrow \cdots \).

We demonstrate the effectiveness of our technique in a realistic setting. We implemented the algorithm inside an industry-strength abstract-interpretation-based C static analyzer [6, 7, 8] and tested its performance on open-source benchmarks. We have saved 7%-96% of the analysis time for context-insensitive or -sensitive global analysis.

1.3. Contributions

- We present an extension of the approximate call-strings approach, which effectively reduces the inefficiency caused by large, inevitable, spurious interprocedural cycles. We prove the effectiveness of the technique by experiments with an industry-strength C static analyzer [6, 7, 8] in globally analyzing medium-scale open-source programs.
- The technique is meaningful in three ways.
  1. The technique aims to alleviate one major reason (spurious interprocedural cycles) for substantial inefficiency in global static analysis.
  2. It is purely an algorithmic technique inside the worklist-based fixpoint iteration routine. So, it can be directly applicable without changing the analysis’ underlying abstract semantics, regardless of whether the semantics is context-sensitive or not. The technique’s correctness is obvious enough to avoid the burden of a safety proof that would be needed if we newly designed the abstract semantics.
3. The technique not only reduces the analysis time but also improves the analysis precision. This is because (1) our technique removes some (worklist-level) computations that occur along invalid return paths (Section 3.3.1); (2) when the underlying analysis uses widenings, the technique reduces the number of widening points (Section 3.3.2).

- We report one key reason (spurious interprocedural cycles) for why less accurate context-sensitivity actually makes the analyses very slow. Though it is well-known folklore that less precise analysis does not always have less cost [13, 18, 20], there haven’t been realistic experiments about their explicit reason.

1.4. Related Work

We compare, on the basis of their applicability to general semantic-based static analyzers§, our method with other approaches that eliminate invalid paths.

The approximate call-strings approach [21] is popular in practice but its precision is not enough to mitigate large spurious cycles. Sharir and Pnueli [21] presented an approximate call-strings approach in which the last $k$ call-sites are remembered for the calling contexts to each procedure. The $k$ length suffix method is an approximation of the full call-strings approach [21, 9, 11] and has been used as a feasible alternative in practice [1, 12, 13]. Moreover, it is actually one of very few options available for semantic-based global static analysis that uses infinite domains and non-distributive flow functions (e.g., [1, 7]). However, the $k$ length suffix method induces a large spurious cycle because it permits multiple returns of procedures. Our algorithm is an extension of the $k$ length suffix method and adds extra precision that relieves the performance problem from spurious interprocedural cycles.

Another approximate call-strings method that uses full context-sensitivity for non-recursive procedures has been shown to be practical for points-to analysis [22, 23] but, the method is too costly for more general semantic-based analyses. The method is approximate because it does not distinguish the calling contexts for recursive calls. Whaley and Lam [23] used BDDs to efficiently encode the calling contexts and showed that full context-sensitivity is feasible for non-recursive procedures. Their analysis is fully context-sensitive for non-recursive procedures and does not suffer from large spurious cycles caused by non-recursive procedures. Sridharan and Bodík [22] presented an approximation, called regular-reachability, of the CFL(context-free language)-reachability [17]. They transform the analysis problem into graph reachability problem [16] and only consider execution paths where calls and returns are properly matched for programs without recursive procedures. Since the set of calling contexts that they consider is finite (because they do not consider recursion), the set of calling contexts can be described by a regular language instead of context-free languages. Though these approaches are more precise than $k$ length suffix method, it is unknown whether the BDD-based method [23] or regular-
reachability [22] are also applicable in practice to general semantic-based analyzers rather than pointer analysis. Our algorithm can be useful for analyses for which these approaches hit a cost limit in practice and k length suffix method should be used instead.

Full call-strings approaches [21, 9, 11] and functional approaches [21] do not suffer from spurious cycles but are limited to restricted classes of data flow analysis problems. The original full call-strings method [21] prescribes the domain to be finite and its improved algorithms [9, 11] are also limited to bit-vector problems or finite domains. For infinite domains, these algorithms can possibly generate infinite number of call-strings and hence may not terminate. Khedker et al.’s algorithm [11] supports infinite domains only after unfolding cyclic call chains by a fixed number. A functional approach [21] builds the summary flow functions for each procedure in a context independent way and these functions are used as flow functions of call statements. Because using the summary functions does not require traversing the called procedure’s bodies, functional approaches also does not suffer from spurious cycles problem. However, computing summary flow functions requires efficient representation of function compositions and meets and hence is applicable to only a restricted data flow analysis problems.

Reps et al.’s algorithms [17, 19] to avoid unrealizable paths are limited to analysis problems that can be expressed only in their graph reachability framework. These algorithms are variants of the iterative functional approach [21] that require the flow functions to be distributive. So, their algorithm cannot handle prevalent yet non-distributive analyses. For example, our analyzer that uses the interval domain [5] with non-distributive flow functions does not fall into either their IFDS [17] or IDE [19] problems. Meanwhile, our algorithm is independent of the underlying abstract semantic functions. The regular-reachability [22], which is a restricted version of Reps et al.’s algorithm [17], also requires the analysis problem to be expressed in graph reachability problem.

Chambers et al.’s technique [4] is similar to ours but entails a relatively large change to an existing worklist order. Their technique analyzes each procedure intraprocedurally, and at call-sites continues the analysis of the callee. It returns to analyze the nodes of the caller only after finishing the analysis of the callee. Our worklist prioritizes the callee only over the call nodes that invoke the callee, not the entire caller, which is a relatively smaller change than Chambers et al.’s. In addition, they assume worst case results for recursive calls, but we do not degrade the analysis precision for recursive calls.

The idea of remembering immediate calling context is first proposed by Myers [14] and we extend it to call-strings method. By remembering the immediate calling context only, Myers’ algorithm is context-sensitive for bit-vector frameworks [10]. Unfortunately, Myers’ formulation is applicable only to bit-vector problems and hard to extend it to general call-strings-based analysis. This paper can be understood as an extension of Myers’ algorithm for general call-strings-based static analysis.

1.5. Organization

Section 2 discusses the performance problem of the traditional call-strings-based context-sensitive or -insensitive interprocedural analysis. Section 3 presents our solution to mitigate the problem. We first describe the approximate call-strings approach and then present our
2. Performance Problems by Large Spurious Cycles

In this section, we show that large spurious cycles are frequently created during (both context-insensitive and -sensitive) global static analysis, and that they drastically degrade the analysis performance. The approximate call-strings-based context-sensitive abstract semantics cannot effectively eliminate such large spurious cycles.

2.1. Interprocedural Spurious Cycles Reach Far In Real C Programs

If a spurious cycle is created by multiple calls to a procedure \( f \), then all the procedures that are reachable from \( f \) or that reach \( f \) via the call-graph belong to the cycle because of call and return flows. For example, consider a call-chain \( \cdots f_1 \rightarrow f_2 \rightarrow \cdots \). If \( f_1 \) calls \( f_2 \) multiple times, creating a spurious butterfly cycle \( f_1 \rightleftharpoons f_2 \) between them, then fixpoint-solving the cycle involves all the nodes of procedures that reach \( f_1 \) or that are reachable from \( f_2 \). This situation is common in C programs. For example, in GNU software, the \texttt{xmalloc} procedure, which is in charge of memory allocation, is called from many other procedures, and hence generates a butterfly cycle. Then every procedure that reaches \texttt{xmalloc} via the call-graph is trapped into a fixpoint cycle.

In conventional context-sensitive analysis that distinguishes the last \( k \) call-sites [21], if there are call-chains of length \( l \) (\( > k \)) in programs, it’s still possible to have a spurious cycle created during the first \( l - k \) calls. This spurious cycle traps the last \( k \) procedures into a fixpoint cycle by the above reason.

One spurious cycle in a real C program can trap as many as 80-90\% of basic blocks of the program into a fixpoint cycle. Figure 3 shows this phenomenon. In the figures, the x-axis represents the execution time of the analysis and the y-axis represents the procedure name, which is mapped to unique integers. During the analysis, we draw the graph by plotting the point \((t, f)\) if the analysis’ worklist algorithm visits a node of procedure \( f \) at the time \( t \). For brevity, the graph for sed-4.0.8 is shown only up to 100,000 iterations among more than 3,000,000 total iterations. From the results, we first observe that similar patterns are repeated and each pattern contains almost all procedures in the program. And we find that there are much more repetitions in the case of a large program (sed-4.0.8, 26,807 LOC) than a small one (spell-1.0, 2,213 LOC): more than 150 repeated iterations were required to analyze sed-4.0.8 whereas spell-1.0 needed about 30 repetitions.

3. Our Algorithmic Mitigation Technique

In this section, we present our extension of the approximate call-strings-based approach, aiming to mitigate performance problems caused by the large spurious cycles. Our technique is purely
algorithmic: the technique does not depend on the underlying abstract semantics but is a simple addition to the existing worklist-based fixpoint algorithm.

We first describe the traditional call-strings-based analysis algorithm (section 3.2) as well as the representation of programs (section 3.1). Then we present our algorithmic extension of the classical algorithm (section 3.3).

3.1. Graph Representation of Programs

We assume that a program is represented by a supergraph [17]. A supergraph consists of control flow graphs of procedures with interprocedural edges connecting each call-site to its callee. Each node \( n \in \text{Node} \) in the graph has one of the five types:

\[
\text{entry}_f | \text{exit}_f | \text{call}^{g,r}_f | \text{rtn}^c_f | \text{cmd}_f
\]

The subscript \( f \) of each node represents the procedure name enclosing the node. \( \text{entry}_f \) and \( \text{exit}_f \) are entry and exit nodes of procedure \( f \). A call-site in a program is represented by a call node and its corresponding return node. A call node \( \text{call}^{g,r}_f \) indicates that it invokes a procedure \( g \) and its corresponding return node is \( r \). We assume that function pointers are resolved (before the analysis).\(^4\) Node \( \text{rtn}^c_f \) represents a return node in \( f \) whose corresponding call node is \( c \). Node \( \text{cmd}_f \) represents a general command statement. Edges are assembled by a function, \( \text{succof} \), which maps each node to its successors. \( \text{CallNode} \) is the set of call nodes in a program.

\(^4\)We use an efficient, flow-insensitive pointer analysis for resolving function pointers.
3.2. Normal\(_k\): A Normal Call-Strings-Based Analysis Algorithm

Call-strings are sequences of call nodes. To make them finite, we only consider call-strings of length at most \(k\) for some fixed integer \(k \geq 0\). We write \(\text{CallNode}^{\leq k} \equiv \Delta\) for the set of call-strings of length \(\leq k\). We write \([c_1, c_2, \ldots, c_i]\) for a call-string of call sequence \(c_1, c_2, \ldots, c_i\). Given a call-string \(\delta\) and a call node \(c, [\delta, c]\) denotes a call-string obtained by appending \(c\) to \(\delta\). In the case of context-insensitive analysis \((k = 0)\), we use \(\Delta = \{\epsilon\}\), where the empty call-string \(\epsilon\) means no context-information.

Figure 4. (a) shows the worklist-based fixpoint iteration algorithm that performs call-strings(\(\Delta\))-based context-sensitive (or insensitive, when \(k = 0\)) analysis. The algorithm computes a table \(T \in \text{Node} \rightarrow \text{State}\) which associates each node with its input state \(\text{State} = \Delta \rightarrow \text{Mem}\), where \(\text{Mem}\) denotes abstract memory, which is a map from program variables to abstract values. That is, call-strings are tagged to the abstract memories and are used to distinguish the memories propagated along different interprocedural paths, to a limited extent (the last \(k\) call-sites). The worklist \(\mathcal{W}\) consists of node and call-string pairs. The algorithm chooses a work-item \((n, \delta) \in \text{Node} \times \Delta\) from the worklist and evaluates the node \(n\) with the flow function \(\mathcal{F}\). Next work-items to be inserted into the worklist are defined by function \(\mathcal{N} \in \text{Node} \times \Delta \rightarrow 2^{\text{Node} \times \Delta} : \)

\[
\mathcal{N}(n, \delta) = \begin{cases} 
\{(r, \delta') \mid \delta = [\delta', \text{call}^{\mathcal{F}, r}_f]_k \land \delta' \in \text{dom}(T(\text{call}^{\mathcal{F}, r}_f))\} & \text{if } n = \text{exit}_g \\
\{(\text{entry}_g, [\delta, n]_k)\} & \text{if } n = \text{call}^{\mathcal{F}, r}_f \\
\{(n', \delta) \mid n' \in \text{succof}(n)\} & \text{otherwise}
\end{cases}
\]

where \(\text{dom}(f)\) denotes the domain of map \(f\) and \([\delta, c]_k\) denotes the call-string \([\delta, c]\) but possibly truncated so as to keep at most the last \(k\) call-sites.

The algorithm can follow spurious return paths if the input program’s nested call-depth is larger than the \(k\). The mapping \(\delta'\) to \([\delta', \text{call}^{\mathcal{F}, r}_f]_k\) is not one-to-one and \(\mathcal{N}\) possibly returns many work-items at an exit node. The following example illustrates this situation.

**Example 2.** Let \(k = 2\) and suppose call-strings \([c_1, c_3]\) and \([c_2, c_3]\) are tagged to a call node \(\text{call}^{\mathcal{F}, r}_f\). Suppose \(\text{call}^{\mathcal{F}, r}_f\) calls \(g\) under the call-string \([c_1, c_3]\). By the definition of \(\mathcal{N}\), the call-string at \(\text{entry}_g\) is \([c_1, c_3, \text{call}^{\mathcal{F}, r}_f]_2 = [c_3, \text{call}^{\mathcal{F}, r}_f]_2\). After the analysis of \(g\), the call-string at \(\text{exit}_g\) is also \([c_3, \text{call}^{\mathcal{F}, r}_f]_2\). When \(g\) returns, since the call-string at \(\text{exit}_g\) equals to both \([c_1, c_3, \text{call}^{\mathcal{F}, r}_f]_2\) and \([c_2, c_3, \text{call}^{\mathcal{F}, r}_f]_2\), \(\mathcal{N}\) returns two work-items \((r, [c_1, c_3])\) and \((r, [c_2, c_3])\). The return to \((r, [c_1, c_3])\) is spurious because \(g\) was called under the context \([c_1, c_3]\). □

We call the above analysis algorithm \(\text{Normal}_k(k = 0, 1, 2, \ldots)\). \(\text{Normal}_0\) performs context-insensitive analysis, \(\text{Normal}_1\) performs context-sensitive analysis that distinguishes the last 1 call-site, and so on.

3.3. Normal\(_k\)/(RSS): Our Algorithm

Before discussing our technique, we define the call-context that will be used throughout this section.
Definition 1. When a procedure $g$ is called from a call node $\text{call}^g_{r_f}$ under context $\delta$, we say that $(\text{call}^g_{r_f}, \delta)$ is the call-context for that procedure call. Since each call node $\text{call}^g_{r_f}$ has a unique return node, we interchangeably write $(r, \delta)$ and $(\text{call}^g_{r_f}, \delta)$ for the same call-context.

Our return-site-sensitive (RSS) technique is simple. When calling a procedure at a call-site, the call-context for that call is remembered until the procedure returns. The bookkeeping cost is limited to only one memory entry per procedure. This is possible by the following strategies:

1. **Single return**: Whenever the analysis of a procedure $g$ is started from a call node $\text{call}^g_{r_f}$ in $f$ under call-string $\delta$, the algorithm remembers its call-context $(r, \delta)$, consisting of the corresponding return node $r$ and the call-string $\delta$. And upon finishing analyzing $g$’s body, after evaluating $\text{exit}_g$, the algorithm inserts only the remembered return node $(r, \delta)$ into the worklist. Multiple returns are avoided. For correctness, this single return should be allowed only when the other call nodes that call $g$ are not analyzed until the analysis of $g$ from $(\text{call}^g_{r_f}, \delta)$ completes.

Example 3. Consider the situation of Example 2 again. When $g$ is called from $\text{call}^g_{r_f}$ under the context $[c_1, c_3]$, our algorithm remembers $g$’s call-context $(r, [c_1, c_3])$. And at $\text{exit}_g$, under its context $[c_3, \text{call}^g_{r_f}]$, our algorithm inserts only the remembered $(r, [c_1, c_3])$ into the worklist. The spurious return to $(r, [c_2, c_3])$ is avoided.

2. **One call per procedure, exclusively**: We implement the single return policy by using one memory entry per procedure to remember the call-context. This is possible if we can analyze each called procedure exclusively for its one particular call-context. If a procedure is being analyzed from a call node $c$ with a call-string $\delta$, processings of other call-sites that call the same procedure should wait until the analysis of the procedure from $(c, \delta)$ is completely finished. This one-exclusive-call-per-procedure policy is enforced by not selecting from the worklist call nodes that (directly or transitively) call the procedures that are currently being analyzed.

Example 4. Suppose procedure $g$ was called from $\text{call}^g_{r_f}$ under the context $[c_1, c_3]$ and our algorithm has remembered the call-context $(r, [c_1, c_3])$. Suppose also the current worklist $W = \{(\text{call}^g_{r_f}, [c_2, c_3]), \cdots\}$ which contains a call-site that invokes $g$. In this situation, our algorithm does not select $(\text{call}^g_{r_f}, [c_2, c_3])$ as a next work-item unless the analysis of $g$ is completely finished.

3. **Recursion handling**: The algorithm gives up the single return policy for recursive procedures. This is because we cannot finish analyzing a recursive procedure’s body without considering another call (recursive call) in it. Recursive procedures are handled in the same way as the normal worklist algorithm.

The algorithm does not follow spurious return paths regardless of the program’s nested call-depth. While Normal$_k$ starts losing its power when a call chain’s length is larger than $k$, Normal$_k$/RSS does not. The following example shows this difference between Normal$_k$ and Normal$_k$/RSS.
Example 5. Consider a program that has the following call-chain (where $f_1 \xrightarrow{c_1,c_2} f_2$ denotes that $f_1$ calls $f_2$ at call-sites $c_1$ and $c_2$) and suppose $k = 1$:

$$f_1 \xrightarrow{c_1,c_2} f_2 \xrightarrow{c_3,c_4} f_3$$

- **Normal**$_1$: The analysis results for $f_2$ are distinguished by $[c_1]$ and $[c_2]$ hence no butterfly cycle happens between $f_1$ and $f_2$. Now, when $f_3$ is called from $f_2$ at $c_3$, we have two call-contexts $(c_3,[c_1])$ and $(c_3,[c_2])$ but analyzing $f_3$ proceeds with context $[c_3]$ (because $k = 1$). That is, Normal$_k$ forgets the call-context for procedure $f_3$. Thus the result of analyzing $f_3$ must flow back to all call-contexts with return site $c_3$, i.e., to both the call-contexts $(c_3,[c_1])$ and $(c_3,[c_2])$.

- **Normal**$_1$/RSS: The results for $f_2$ and $f_3$ are distinguished in the same way as Normal$_1$. But, Normal$_1$/RSS additionally remembers the call-contexts for every procedure call. If $f_3$ was called from $c_3$ under context $[c_1]$, our algorithmic technique forces Normal$_k$ to remember the call-context $(c_3,[c_1])$ for that procedure call. And finishing analyzing $f_3$’s body, $f_3$ returns only to the remembered call-context $(c_3,[c_1])$. This is possible by the one-exclusive-call-per-procedure policy.

We ensure the one-exclusive-call-per-procedure policy by prioritizing a callee over call-sites that (directly or transitively) invoke the callee. The algorithm always analyzes the nodes of the callee $g$ first prior to any other call nodes that invoke $g$: before selecting a work-item as a next job, we exclude from the worklist every call node $\text{call}_g^{h,r}$ to a node $g$ if the worklist contains any node of procedure $h$ that can be reached from $g$ along some call-chain $g \rightarrow \cdots \rightarrow h$, including the case of $g = h$. After excluding such call nodes, the algorithm chooses a work-item in the same way as a normal worklist algorithm, i.e., after the exclusion, our algorithm relies on the existing worklist ordering strategy in selecting the next work-item.

Example 6. Consider a worklist $\{(\text{call}_f^{g,r_1},\delta_1),(\text{call}_j^{h,r_2},\delta_2), (n_h,\delta_3), (\text{call}_h^{i,r_4},\delta_4)\}$ and assume there is a path $f \rightarrow g \rightarrow h$ in the call graph. When choosing a work-item from the worklist, our algorithm first excludes all the call nodes that invoke procedures now being analyzed: $\text{call}_h^{i,r_4}$ is excluded because $h$’s node $n_h$ is in the worklist. Similarly, $\text{call}_j^{g,r_1}$ is excluded because there is a call-chain $g \rightarrow h$ in the call graph and $h$’s node $n_h$ exists. So, the algorithm chooses a work-item from $\{(n_h,\delta_3), (\text{call}_h^{i,r_4},\delta_4)\}$. The excluded work-items $(\text{call}_j^{g,r_1},\delta_1)$ and $(\text{call}_h^{i,r_2},\delta_2)$ will not be selected unless there are no nodes of $h$ in the worklist.

Figure 4(b) shows a change in the technique that is applied to the normal worklist algorithm of Figure 4(a). To transform Normal$_k$ into Normal$_k$/RSS, only shaded lines are inserted; other parts remain the same. ReturnSite is a map to record a single return site information (return node and context pair) per procedure. Lines 15-16 are for remembering a single return when encountering a call-site. The algorithm checks if the current node is a call-node and its target procedure is non-recursive (the recursive predicate decides whether the procedure is recursive or not), and if so, it remembers its single return-site information for the call. Lines 17-21 handle procedure returns. If the current node is an exit of a non-recursive procedure, only the remembered return for that procedure is used as a next work-item, instead of all possible next (successor, context) pairs (line 23).
(01) \( \delta \in \text{Context} = \Delta \)  
(02) \( w \in \text{Work} = \text{Node} \times \Delta \)  
(03) \( W \in \text{Worklist} = 2^\text{Work} \)  
(04) \( N \in \text{Node} \times \Delta \rightarrow 2\text{Node} \times \Delta \)  
(05) \( \text{State} = \Delta \rightarrow \text{Mem} \)  
(06) \( T \in \text{Table} \rightarrow \text{Node} \rightarrow \text{State} \)  
(07) \( \tilde{F} \in \text{Node} \rightarrow \text{Mem} \rightarrow \text{Mem} \)  

\[
\begin{align*}
(08) & \quad \text{ReturnSite} \in \text{ProcName} \rightarrow \text{Work} \\
(09) & \quad \text{FixpointIterate} (W, T) = \\
(10) & \quad \text{Repeat} \\
(11) & \quad \text{while} \\
(12) & \quad S := \{(\text{call}^p \cdot \cdot) \in W \mid (n, \cdot) \in W \land \text{reach}(g, h) \land \lnot \text{recursive}(g)\} \\
(13) & \quad (n, \delta) := \text{choose}(W) \\
(14) & \quad m := \tilde{F} n (T(n)(\delta)) \\
(15) & \quad \text{if } n = \text{call}^p \cdot \cdot \land \lnot \text{recursive}(g) \text{ then} \\
(16) & \quad \text{ReturnSite}(g) := (r, \delta) \\
(17) & \quad \text{if } n = \text{exit} \cdot g \land \lnot \text{recursive}(g) \text{ then} \\
(18) & \quad (r, \delta_r) := \text{ReturnSite}(g) \\
(19) & \quad \text{if } m \not\in T(r)(\delta_r) \text{ then} \\
(20) & \quad W := W \cup \{(r, \delta_r)\} \\
(21) & \quad T(r)(\delta_r) := T(r)(\delta_r) \cup m \\
(22) & \quad \text{else} \\
(23) & \quad \text{for all } (n', \delta') \in N(n, \delta) \text{ do} \\
(24) & \quad \text{if } m \not\in T(n')(\delta') \text{ then} \\
(25) & \quad W := W \cup \{(n', \delta')\} \\
(26) & \quad T(n')(\delta') := T(n')(\delta') \cup m \\
(27) & \quad \text{until } W = \emptyset
\end{align*}
\]

(a) a normal worklist algorithm \textbf{Normal}_k  
(b) our algorithm \textbf{Normal}_k/RSS

Figure 4. A normal context-sensitive worklist algorithm \textbf{Normal}_k and its RSS modification \textbf{Normal}_k/RSS. The left-hand side shows a worklist algorithm for call-strings based context-sensitive analysis. The right-hand side shows the RSS algorithm for the same analysis. These two algorithms are the same except for shaded regions. For brevity, we omit the usual definition of \( \tilde{F} \), which updates the worklist in addition to computing the flow equation’s body.

delaying call nodes to procedures now being analyzed. To do this, in line 12-13, the algorithm excludes the call nodes \( \{(\text{call}^p \cdot \cdot) \in W \mid (n, \cdot) \in W \land \text{reach}(g, h) \land \lnot \text{recursive}(g)\} \) that invoke non-recursive procedures whose nodes are already contained in the current worklist. \text{reach}(g, h) \) is true if there is a path in the call graph from \( g \) to \( h \).
Example 7. Analyzing the program in the left-hand side of the figure below proceeds as shown in the right-hand side table. (Assume that \( k = 0 \), the \textit{choose} function in Figure 4 arbitrarily chooses an element from the given worklist, and the initial worklist is \( \{1, 4\} \)).

<table>
<thead>
<tr>
<th>Iter</th>
<th>Worklist</th>
<th>Excluded</th>
<th>Return Site</th>
<th>Updated Worklist</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( {1, 4} )</td>
<td>( {} )</td>
<td>( f \mapsto 4 )</td>
<td>( {2, 4} )</td>
</tr>
<tr>
<td>2</td>
<td>( {2, 4} )</td>
<td>( {} )</td>
<td>( f \mapsto 4 )</td>
<td>( {2.5} )</td>
</tr>
<tr>
<td>3</td>
<td>( {2, 5} )</td>
<td>( {5} )</td>
<td>( f \mapsto 4 )</td>
<td>( {3.5} )</td>
</tr>
<tr>
<td>4</td>
<td>( {3, 5} )</td>
<td>( {5} )</td>
<td>( f \mapsto 4 )</td>
<td>( {4, 5} )</td>
</tr>
</tbody>
</table>

For each iteration of the algorithm, the table shows the contents of the current worklist \( (W) \), call nodes that are excluded at this iteration \( (S) \), return site information \( (\text{ReturnSite}) \), and the updated worklist \( (\text{updated} \ W) \). \( n \) represents the chosen node for each iteration. When the algorithm processes call node 1 at the first iteration, \( f \) remembers its corresponding return-site 4. At the 3rd and 4th iterations, node 5 was excluded, because it is another call to \( f \) and the worklist contains the nodes of \( f \) at both iterations. At the exit of \( f \) (when processing node 3 at the 4th iteration), only \( \text{ReturnSite}(f) = 4 \) is inserted into the worklist instead of \( \text{succof}(f) = \{4, 6\} \). \( \square \)

3.3.1. Correctness & Precision

One noticeable thing of Normal\( k \)/RSS is that the result is not a fixpoint of the given flow equation system, but still a sound approximation of the program semantics. Since the algorithm prunes some computation steps during worklist algorithm (at exit nodes of non-recursive procedures), the result of the algorithm may not be a fixpoint of the original equation system. However, because the algorithm prunes only spurious returns that definitely do not happen in the real executions of the program, our algorithm does not miss any information flow of real executions. In other words, our algorithm does not necessarily produce a maximal fixpoint solution but something below it and still above the real semantics.

For any \( f \) and any arbitrary call-context \( (\text{call}_{f,r}^{f,r}, \delta) \), the single return to \( (r, \delta) \) after analyzing \( f \) is correct if the state from \( (\text{call}_{f,r}^{f,r}, \delta) \) is implied by the input state used in the analysis of \( f \) and its result is guaranteed to be returned to \( (r, \delta) \). The state from every call-context flows into \( f \) (abstract semantics). Our single-return policy does not miss returning \( f \)'s analysis result to its corresponding call-context\( \parallel \) because (1) we remember the context at each call; (2) for every different call, modulo the underlying context-sensitivity, we exclusively analyze \( f \). Because we cannot enforce this exclusivity for recursive calls, we do not apply the algorithm to recursive procedures.

\( \parallel \)Here, we ignore the cases where the callee never returns (e.g., it calls exit()). However, even though that happens, we can enforce the return of callee by always inserting the exit node of a procedure when inserting the entry node of the procedure into the worklist.
Normal	extsubscript{k}/RSS is always at least as precise as Normal	extsubscript{k}. Because Normal	extsubscript{k}/RSS prunes some (worklist-level) computations that occur along invalid return paths, it is likely to have an effect of avoiding propagations of information along invalid return paths. Hence, Normal	extsubscript{k}/RSS gives more precise (or at least the same) results than Normal	extsubscript{k}. The actual precision of Normal	extsubscript{k}/RSS varies depending on the existing worklist order of Normal	extsubscript{k}.

**Example 8.** Consider the program in Example 7 again, and suppose the current worklist is \{1,5\}. When analyzing the program with Normal	extsubscript{0}, the fixpoint-solving follows both spurious return paths, regardless of the worklist order,

\[
\begin{align*}
1 & \rightarrow 2 \rightarrow 3 \rightarrow 6 \\
5 & \rightarrow 2 \rightarrow 3 \rightarrow 4
\end{align*}
\]

because of multiple returns from node 3. When analyzing with Normal	extsubscript{0}/RSS, there are two possibilities, depending on the worklist order:

1. When Normal	extsubscript{0}/RSS selects node 1 first: Then the fixpoint iteration sequence may be 1; 2; 3; 4; 5; 2; 3; 6. This sequence involves the spurious path (1) (because the second visit to node 2 uses the information from node 1 as well as from node 5), but not (2). Normal	extsubscript{0}/RSS is more precise than Normal	extsubscript{0}.

2. When Normal	extsubscript{0}/RSS selects node 5 first: Then the fixpoint iteration sequence may be 5; 2; 3; 6; 1; 2; 3; 4; 5; 2; 3; 6. This computation involves both spurious paths (1) and (2). With this iteration order, Normal	extsubscript{0} and Normal	extsubscript{0}/RSS have the same precision.

\[\square\]

### 3.3.2. Less Widening Points

Widening [5] is a speed-up technique that have been designed to safely approximate least fixpoints of semantic function. In abstract-interpretation-based static analysis, programs invariants are characterized as least fixpoints of (abstract) semantic functions over abstract domain. For finite height domains, the fixpoints are computed by using a classical iterative algorithm. But the iterative algorithm does not terminate or has unacceptable costs for domains with infinite height or very large height. For infinite or very large height domains such as lattice of intervals, the widening technique [5] is used to guarantee or accelerate the analysis’ termination. With widening, the iterative algorithm does not necessarily compute least fixpoints but finds a safe (upper) approximation of the least fixpoint.

Our technique reduces cycles, hence obviously reduces the number of widening points. Because applying widening means losing analysis precision, the widening operation should be carefully applied to as small as possible subset of the entire program points. A common way of selecting such widening points is to apply widening to every heads of loops in program [3], including ones that are interprocedurally created by calling a procedure multiple times. Normal	extsubscript{k}/RSS can reduce the number of widening points more. Normal	extsubscript{k}/RSS need not apply widenings at interprocedural loop-heads that are created by non-recursive procedure calls. This is because Normal	extsubscript{k}/RSS does not follow such interprocedural cycles.
Example 9. Consider the following code and interval-domain-based analysis of the code.

```c
int g = 0;
int f() {
    g++;
}
int main() {
    f();
    f();
}
```

Since procedure \( f \) is called twice from procedure \( \text{main} \), a spurious interprocedural cycle \((5) \rightarrow (2) \rightarrow (3) \rightarrow (4) \rightarrow (5) \cdots\) will be created during the analysis. Iterating through the cycle continually increases the value of the global variable \( g \): \([0, 0] \rightarrow [0, 1] \rightarrow [0, 2] \rightarrow \cdots\). In order to terminate the analysis, a widening should be applied at the entry of procedure \( f \). Hence, \( \text{Normal}_k \) computes \( g = [0, +\infty] \) at the end of procedure \( \text{main} \). However, \( \text{Normal}_k/RSS \) does not apply the widening at the entry of procedure \( f \) (since \( f \) is non-recursive and \( \text{Normal}_k/RSS \) does not follow the spurious return paths \((5) \rightarrow (2) \rightarrow (3) \rightarrow (4)\)), computing \( g = [0, 2] \) at the end of procedure \( \text{main} \).

3.4. A Fast Implementation of \( \text{Normal}_k/RSS \)

In practice, if the worklist algorithm uses a particular worklist ordering strategy, the RSS algorithm can be implemented more easily.

Assume that the worklist algorithm uses a partial order \( \text{RevTop} \) between nodes in the supergraph and retrieves the node with the highest order from the worklist. The order \( \text{RevTop} \) between nodes is defined as a reverse topological order between procedures on the call graph: a node \( n \) of a procedure \( f \) precedes a node \( m \) of a procedure \( g \) if \( f \) precedes \( g \) in the reverse topological order in the call graph. If \( f \) and \( g \) are the same procedure, the order between the nodes is defined by the weak topological order \([3]\) on the control flow graph of that procedure. Note that there can be two or more nodes that have the highest order, for example of each branch of conditional statements. In this case, the algorithm arbitrarily chooses a node among them.

Without recursive procedures, the order \( \text{RevTop} \) guarantees the one-exclusive-call-per-procedure policy. This is because the order means that a callee is always analyzed first rather than its caller. For instance, think of two procedures \( f \) and \( g \) where \( f \) precedes \( g \) in reverse topological order on the call graph. It means that \( f \) is called by some call sites in \( g \). Then the worklist algorithm selects a node of \( f \) first from the worklist rather than nodes of \( g \) unless the worklist does not contain any node of \( f \), which means that all the other calls to \( f \) inside \( g \) wait until the analysis of \( f \) is completely finished. Recursive procedures are handled in the same way as the normal worklist algorithm.

To implement the technique inside the algorithm Figure 4.(b), lines 12 is removed and line 13 is replaced by the following:

\[
(n, \delta) := \text{chooseRevTop}(W)
\]
where \( \text{choose}_{\text{RevTop}} \) chooses the work-item that has the highest \( \text{RevTop} \) order from the worklist \( (W) \).

4. Experiments

We implemented our algorithm inside a realistic C analyzer \([6, 7, 8]\). Experiments with open-source programs show that \( \text{Normal}_k / \text{RSS} \) for any \( k \) is very likely faster than \( \text{Normal}_k \), and that even \( \text{Normal}_{k+1} / \text{RSS} \) can be faster than \( \text{Normal}_k \).

4.1. Setting Up

\( \text{Normal}_k \) is our underlying worklist algorithm, on top of which our industry-strength static analyzer \([6, 7, 8]\) for C is installed. The analyzer is an interval-domain-based abstract interpreter. The analyzer performs by default flow-sensitive and call-strings-based context-sensitive global analysis on the supergraph of the input program: it computes \( T = \text{Node} \rightarrow \text{State} \) where \( \text{State} = \Delta \rightarrow \text{Mem} \). \( \text{Mem} \) denotes abstract memory \( \text{Mem} = \text{Addr} \rightarrow \text{Val} \) where \( \text{Addr} \) denotes abstract locations that are either program variables or allocation sites, and \( \text{Val} \) denotes abstract values including \( \hat{\mathbb{Z}} \) (interval domain), \( 2^{\text{Addr}} \) (points-to set), and \( 2^{\text{AllocSite} \times \hat{\mathbb{Z}} \times \hat{\mathbb{Z}}} \) (array block, consisting of base address, offset, and size \([7]\)).

We evaluated our algorithm in two ways. First, we measured the net effects of avoiding spurious interprocedural cycles. Since our algorithmic technique changes the existing worklist order, performance differences between \( \text{Normal}_k \) and \( \text{Normal}_k / \text{RSS} \) could be attributed not only to avoiding spurious cycles but also to the changed worklist order. In order to measure the net effects of avoiding spurious cycles, we applied the same worklist order \( \text{RevTop} \), defined in Section 3.4, to both \( \text{Normal}_k \) and \( \text{Normal}_k / \text{RSS} \). Note that this ordering itself contains the “prioritize callees over call-sites” feature and we don’t explicitly need the delaying call technique (lines 12-13 in Figure 4.(b)) in \( \text{Normal}_k / \text{RSS} \). Hence the worklist order for \( \text{Normal}_k \) and \( \text{Normal}_k / \text{RSS} \) are the same. ** For this evaluation, we compare analysis time and precision between \( \text{Normal}_k \) and \( \text{Normal}_k / \text{RSS} \).

We also evaluated our algorithm when our technique interferes with the existing worklist order. Because our technique interferes with (i.e., changes) the existing worklist order of \( \text{Normal}_k \), it is necessary to check whether our technique works well regardless of the existing worklist order strategies or not. To see what happens in this case, we applied our technique to \( \text{Normal}_k \) that uses the following worklist order, called \( \text{Arbitrary} \); the order between nodes in different procedures is determined by a random order that is fixed before the analysis and the order between nodes in the same procedure is defined by the weak topological order. Note that the worklist order does not contain the “prioritize callees over call-sites” because the order randomly chooses a procedure regardless of call relationship.

** In fact, the order described here is the one our analyzer uses by default, which consistently shows better performance than naive worklist management scheme (BFS/DFS) or simple “wait-at-join” techniques (e.g., \([7]\)).
Table II. Benchmark programs and their raw analysis results when using RevTop worklist order. Lines of code (LOC) are given before preprocessing. The number of nodes in the supergraph (#nodes) is given after preprocessing. \( k \) denotes the size of call-strings used for the analysis. Entries with \( \infty \) means missing data because of our analysis running out of memory.

<table>
<thead>
<tr>
<th>Program</th>
<th>LOC</th>
<th>#nodes</th>
<th>k-call-strings</th>
<th>#iterations</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Normal/RSS</td>
<td>Normal/RSS</td>
</tr>
<tr>
<td>spell-1.0</td>
<td>2,213</td>
<td>782</td>
<td>0</td>
<td>33,864</td>
<td>60.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>31,933</td>
<td>55.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>57,083</td>
<td>102.28</td>
</tr>
<tr>
<td>barcode-0.96</td>
<td>4,460</td>
<td>2,634</td>
<td>0</td>
<td>22,040</td>
<td>93.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>33,808</td>
<td>144.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>40,176</td>
<td>183.49</td>
</tr>
<tr>
<td>httptunnel-3.3</td>
<td>6,174</td>
<td>2,757</td>
<td>0</td>
<td>412,159</td>
<td>2020.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>267,291</td>
<td>1525.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>669,029</td>
<td>5983.27</td>
</tr>
<tr>
<td>gzip-1.2.4a</td>
<td>7,327</td>
<td>6,271</td>
<td>0</td>
<td>653,063</td>
<td>621.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>991,135</td>
<td>1217.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1,174,632</td>
<td>1116.25</td>
</tr>
<tr>
<td>jwhois-3.0.1</td>
<td>9,344</td>
<td>5,147</td>
<td>0</td>
<td>417,529</td>
<td>1273.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>272,377</td>
<td>1222.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>594,090</td>
<td>1631.07</td>
</tr>
<tr>
<td>parser</td>
<td>10,900</td>
<td>9,298</td>
<td>0</td>
<td>3,452,248</td>
<td>3270.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>( \infty )</td>
<td>( \infty )</td>
</tr>
<tr>
<td>bc-1.06</td>
<td>13,093</td>
<td>4,924</td>
<td>0</td>
<td>1,964,396</td>
<td>3644.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3,038,986</td>
<td>12557.88</td>
</tr>
<tr>
<td>less-290</td>
<td>18,449</td>
<td>7,754</td>
<td>0</td>
<td>3,149,284</td>
<td>20196.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>( \infty )</td>
<td>( \infty )</td>
</tr>
<tr>
<td>twolf</td>
<td>19,700</td>
<td>14,610</td>
<td>0</td>
<td>3,028,814</td>
<td>1395.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>( \infty )</td>
<td>( \infty )</td>
</tr>
<tr>
<td>tar-1.13</td>
<td>20,258</td>
<td>10,800</td>
<td>0</td>
<td>4,748,749</td>
<td>9973.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>( \infty )</td>
<td>( \infty )</td>
</tr>
<tr>
<td>make-3.76.1</td>
<td>27,304</td>
<td>11,061</td>
<td>0</td>
<td>4,613,382</td>
<td>44853.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>( \infty )</td>
<td>( \infty )</td>
</tr>
</tbody>
</table>

We have analyzed 11 open-source and SPEC2000 software packages. Table II shows our benchmark programs. All experiments were done on a Linux 2.6 system running on a Pentium4 3.2 GHz box with 4 GB of main memory. parser and twolf are from SPEC2000 benchmarks and the others are open-source software.

We use two performance measures: (1) \#iterations is the total number of iterations during the worklist algorithm. The number directly indicates the amount of computation; (2) time is the CPU time spent during the analysis. Though time is roughly proportional to \#iterations, it is subject to change because of different implementations and test environments.
4.2. The Net Effects of Avoiding Spurious Cycles

4.2.1. Reduced Analysis Time

Figure 5(a) compares \#iterations between Normal\(k\)/RSS and Normal\(k\) for \(k = 0, 1, 2\) using RevTop worklist order, which shows the net effects of avoiding spurious cycles. In this comparison, Normal\(k\)/RSS reduces the number of iterations of Normal\(k\) by on average 72%.

When \(k = 0\) (context-insensitive), Normal\(0\)/RSS has reduced \#iterations by, on average, about 72% against Normal\(0\). For most programs, the analysis time has been reduced by more than 50%. There is one exception: barcode. The amount of computation has been reduced by 11%. This is because barcode has unusual call structures: it does not call a procedure many times, but calls many different procedures one by one. So, the program contains few butterfly cycles.

When \(k = 1\), Normal\(1\)/RSS has reduced \#iterations by, on average, about 53% against Normal\(1\). Compared to the context-insensitive case (\(k = 0\)), for all programs, cost reduction ratios have been slightly decreased. As an example, for spell, the reduction ratio when \(k = 0\) is 83% and the ratio when \(k = 1\) is 68%. This is mainly because, in our analysis, Normal\(0\) costs more than Normal\(1\) for most programs (spell, httptunnel, jwhois). For httptunnel, in Table II, the analysis time (2020.10 s) for \(k = 1\) is less than the time (1525.26 s) for \(k = 0\). This means that performance problems by butterfly cycles is much more severe when \(k = 0\) than that of \(k = 1\), because by increasing context-sensitivity some spurious paths can be removed. However, by using our algorithm, we can still reduce the cost of Normal\(1\) by 53%.

When \(k = 2\), Normal\(2\)/RSS has reduced \#iterations by, on average, 60% against Normal\(2\). Compared to the case of \(k = 1\), the cost reduction ratio has been slightly increased for most programs. For example, the ratio for spell has changed from 68% to 73%. In the analysis of Normal\(2\), since the equation system is much larger than that of Normal\(1\), our conjecture is that the size of butterfly cycles is likely to get larger. Since larger butterfly cycles causes more serious problems (Section 2), our RSS algorithm is likely to greater reduce useless computation.

Figure 5(b) compares the performance of Normal\(k+1\)/RSS against Normal\(k\) for \(k = 0, 1\). The result shows that, for all programs except barcode, even Normal\(k+1\)/RSS is faster than Normal\(k\). Since Normal\(k+1\)/RSS can be even faster than Normal\(k\), if memory cost permits, we can consider using Normal\(k+1\)/RSS instead of Normal\(k\).

4.2.2. Increased Analysis Precision

Table III compares the precision between Normal\(0\) and Normal\(0\)/RSS.\(^\dagger\) In order to measure the increased precision, we first joined all the memories associated with each program point (Node). Then we counted the number of constant intervals (\#const, e.g., \([1, 1]\)), finite intervals (\#finite, e.g., \([1, 5]\)), intervals with one infinity (\#open, e.g., \([-1, +\infty)\) or \((-\infty, 1]\)), and intervals with

\(^\dagger\)We compared the precision for the case of \(k = 0\) and for the first five programs in Table II because we need more memory to do the precision comparison (we should keep two analysis results of Normal\(0\) and Normal\(0\)/RSS at the same time).
Figure 5. Net effects of avoiding spurious cycles

(a) Comparison of \#iterations between Normal$_k$ and Normal$_k$/RSS, for $k = 0, 1, 2$.

(b) Comparison of \#iterations between Normal$_k$ and Normal$_{k+1}$/RSS, for $k = 0, 1$. 

Copyright © 2000 John Wiley & Sons, Ltd.  
Softw. Pract. Exper. 2000; 00:1–7  
Prepared using speauth.cls
Table III. Comparison of precision between Normal\textsubscript{0} and Normal\textsubscript{0}/RSS.

<table>
<thead>
<tr>
<th>Program</th>
<th>Analysis</th>
<th>#const</th>
<th>#finite</th>
<th>#open</th>
<th>#top</th>
</tr>
</thead>
<tbody>
<tr>
<td>spell-1.0</td>
<td>Normal\textsubscript{0}</td>
<td>345</td>
<td>88</td>
<td>33</td>
<td>143</td>
</tr>
<tr>
<td></td>
<td>Normal\textsubscript{0}/RSS</td>
<td>345</td>
<td>89</td>
<td>35</td>
<td>140</td>
</tr>
<tr>
<td>barcode-0.96</td>
<td>Normal\textsubscript{0}</td>
<td>2136</td>
<td>588</td>
<td>240</td>
<td>527</td>
</tr>
<tr>
<td></td>
<td>Normal\textsubscript{0}/RSS</td>
<td>2136</td>
<td>589</td>
<td>240</td>
<td>526</td>
</tr>
<tr>
<td>httptunnel-3.3</td>
<td>Normal\textsubscript{0}</td>
<td>1337</td>
<td>342</td>
<td>120</td>
<td>481</td>
</tr>
<tr>
<td></td>
<td>Normal\textsubscript{0}/RSS</td>
<td>1345</td>
<td>342</td>
<td>120</td>
<td>473</td>
</tr>
<tr>
<td>gzip-1.2.4a</td>
<td>Normal\textsubscript{0}</td>
<td>1995</td>
<td>714</td>
<td>255</td>
<td>1214</td>
</tr>
<tr>
<td></td>
<td>Normal\textsubscript{0}/RSS</td>
<td>1995</td>
<td>716</td>
<td>255</td>
<td>1212</td>
</tr>
<tr>
<td>jwhois-3.0.1</td>
<td>Normal\textsubscript{0}</td>
<td>2740</td>
<td>415</td>
<td>961</td>
<td>1036</td>
</tr>
<tr>
<td></td>
<td>Normal\textsubscript{0}/RSS</td>
<td>2740</td>
<td>415</td>
<td>961</td>
<td>1036</td>
</tr>
</tbody>
</table>

two infinity (#\textsubscript{top}, (−\infty, +\infty)) from interval values (\(\hat{Z}\)) and array blocks (2\textsuperscript{AllocSite}×\(\hat{Z}\)×\(\hat{Z}\)) contained in the joined memory. The constant interval and top interval indicate the most precise and imprecise values, respectively. The results show that Normal\textsubscript{0}/RSS is more precise (spell, barcode, httptunnel, gzip) than Normal\textsubscript{0} or the precision is the same (jwhois).

4.3. Speed Up When Interfering the Existing Worklist Order

Figure 6.(a) compares #iterations between Normal\textsubscript{k} and Normal\textsubscript{k}/RSS for \(k = 0\) using Arbitrary worklist order. In the comparison, Normal\textsubscript{k}/RSS reduces the computation cost of Normal\textsubscript{k} by on average 53%. From this results, we can find that the interference does not significantly affect the overall performance differences: the reduction ratio has been decreased by 19% from the case of net effects of avoiding spurious cycles (72%). Hence, the technique is likely to relieve the problems of spurious cycles regardless of the existing worklist ordering strategies.

5. Conclusion

We have presented a simple algorithmic extension of the approximate call-strings approach to alleviate substantial inefficiency caused by large spurious interprocedural cycles. Such cycles are identified as a major reason for the folklore problem in static analysis that less precise analyses sometimes are slower. Although this inefficiency might not come to the fore when analyzing small programs, globally analyzing medium or large programs makes it outstanding. The proposed algorithmic technique reduces the analysis time by 7%-96% for open-source benchmarks.

Our technique is orthogonally applicable to context-sensitive analysis. It is a simple technique inside the worklist-based fixpoint iteration routine. It is directly applicable without changing
the analysis’ underlying abstract semantics, regardless of whether the semantics is context-
sensitive or not.

We have also shown, by experiments, that our technique works regardless of the existing
worklist ordering strategies. So, it is applicable without changing the underlying ordering
schemes of the fixpoint algorithm.

Our technique suggests the following implementation guideline in tuning a global semantic
analysis. Suppose we develop an analyzer that uses call-strings of size $k$ for context-sensitivity
with the $\text{Normal}_k$ algorithm. Suppose further that we cannot increase the call-strings size more
than $k$ because of either the time or memory cost. In this situation, our algorithmic technique
has the following usages.

- When we cannot increase the call-strings size more than $k$ because of the memory cost:
  then use $\text{Normal}_k/\text{RSS}$ instead of $\text{Normal}_k$. This is because (1) $\text{Normal}_k/\text{RSS}$ is empirically

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline
Program & $\#\text{iterations}$ & $\text{time}$ & $\#\text{iterations}$ & $\text{time}$ \\
& Normal & Normal/\text{RSS} & Normal & Normal/\text{RSS} \\
\hline
spell-1.0 & 36.272 & 20.377 & 99.19 & 43.66 \\
barcode-0.96 & 71.342 & 29.574 & 534.9 & 154.36 \\
htptunnel-3.3 & 591.030 & 132.668 & 4132.21 & 730.95 \\
gzip-1.2.4a & 804.240 & 204.553 & 6844.31 & 1299.36 \\
jwhois-3.0.1 & 777.867 & 761.117 & 5518.04 & 4664.2 \\
parser & 3,500,035 & 1,095,194 & 70248.32 & 4664.2 \\
bc-1.06 & 2,231,064 & 1,138,847 & 23136.25 & 14240.14 \\
less-290 & 5,310,745 & 2,334,886 & 70553.14 & 43381.18 \\
twolf & 5,310,745 & 2,334,886 & 70553.14 & 43381.18 \\
make-3.76.1 & 4,415,305 & 2,110,272 & 70553.14 & 43381.18 \\
\hline
\end{tabular}
\caption{Benchmark programs and their raw analysis results.}
\end{table}

Figure 6. The analysis results when using Arbitrary worklist order.
faster than Normal\(_k\) (Section 4.1 and Figure 5.(a),6); (2) Normal\(_k\)/RSS is in principle more accurate or at least does not sacrifice the precision of Normal\(_k\) (Section 3.3.1, 3.3.2 and Table III); (3) Normal\(_k\)/RSS requires in extra just as many memory entities as the number of procedures.

• When we cannot increase the call-strings size more than \(k\) because of the time cost: then, if memory permits, consider using Normal\(_{k+1}\)/RSS instead. This is because (1) Normal\(_{k+1}\)/RSS can be even faster than Normal\(_k\) (Section 4.1 and Figure 5.(b)); (2) it requires in extra just as many entities as the number of procedures.

Though tuning the accuracy of static analysis can in principle be controlled solely by redesigning the underlying abstract semantics, our algorithmic technique is a simple and orthogonal leverage to effectively shift the analysis cost/accuracy balance for the better. The technique’s correctness is obvious enough to avoid the burden of safety proof of otherwise a newly designed abstract semantics.

ACKNOWLEDGEMENTS

This work was supported by the Engineering Research Center of Excellence Program of Korea Ministry of Education, Science and Technology(MEST) / National Research Foundation of Korea (NRF) (Grant 2009-0063247) and the Brain Korea 21 Project, School of Electrical Engineering and Computer Science, Seoul National University.

REFERENCES


