SolQDebug: Debug Solidity Quickly for Interactive Immediacy in Smart Contract Development

Inseong Jeon[1], Sundeuk Kim[1], Hyunwoo Kim[1], Hoh Peter In[1*]

[1*]Department of Computer Science, Korea University, 145, Anam-ro, Seonbuk-gu, 02841, Seoul, Republic of Korea.

*Corresponding author(s). E-mail(s): hoh_in@korea.ac.kr;
Contributing authors: iwwyou@korea.ac.kr; sd_kim@korea.ac.kr; khw0809@korea.ac.kr;

Abstract

Debugging Solidity contracts remains cumbersome and slow. Even a simple inspection, such as tracking a variable through a branch, requires full compilation, contract deployment, preparatory transactions, and step-by-step bytecode tracing. Existing tools operate only after execution and offer no support while code is under construction. We present SolQDebug, the first interactive, source-level assistant for Solidity developers that provides millisecond feedback before compilation or chain interaction. SolQDebug extends the Solidity grammar with interactive parsing, incrementally maintains a dynamic control-flow graph, and performs interval-based abstract interpretation guided by inline test annotations, enabling developers to simulate symbolic inputs and inspect contract behavior as in traditional debugging environments. In an evaluation on real-world functions, SolQDebug enables low-latency, statement-level analysis during development without requiring compilation or deployment.

Keywords: Smart Contract Development, Solidity, Debugging, Abstract Interpretation

1 Introduction

Smart contracts are the backbone of decentralized applications, and Solidity has become the dominant language for writing them (3; 29). As contracts grow more complex and control more assets, developers must reason about correctness throughout the development cycle—not just at deployment. Large language models (LLMs) such as

ChatGPT (2) or Llama (17) can assist with code generation but offer no guarantees of correctness. Ultimately, developers remain responsible for understanding variable interactions, control flow, and numeric boundaries during authoring.

Unfortunately, the debugging workflow for Solidity lags far behind traditional programming environments. Even a single inspection requires full compilation, deployment, transaction-based state setup, and manual bytecode-level tracing. Tools like Remix IDE (23), Hardhat (12), and Foundry Forge (8) replicate this costly pipeline, providing no live feedback during edits. A prior study found that 88.8% of Solidity developers described debugging as painful, and 69% attributed this to the absence of interactive, source-level tooling Zou et al. (42). Despite this widely acknowledged pain point, we find no existing research or tooling that provides interactive feedback during Solidity code authoring—a gap that this paper aims to fill.

This paper presents SolQDebug, a source-level interactive Solidity debugger powered by abstract interpretation. Rather than replacing runtime debuggers, it complements them by enabling symbolic, per-statement inspection during code authoring—before compilation or deployment. It targets the Solidity pattern of single-contract, single-transaction execution, where each function is isolated and stateless—ideal for static reasoning but difficult to simulate manually. To support this, SolQDebug applies interval-based abstract interpretation, which generalizes over symbolic inputs, exposes edge-case behaviors, and provides sound results with low overhead. This approach gives developers immediate feedback and enables them to reason efficiently about how symbolic inputs influence variable behavior. Although these inputs enable generalization across multiple cases, certain input configurations or control structures may lead to wider output ranges. We evaluate these behaviors empirically and propose annotation strategies that help maintain interpretability across typical Solidity patterns.

To achieve this goal, SolQDebug builds on two core ideas. First, it extends the Solidity grammar with interactive parsing rules and dynamically updates the control-flow graph to reflect incremental edits, enabling keystroke-level structural changes during code authoring. Second, it performs abstract interpretation seeded by inline annotations. These annotations, written directly in the source code, allow developers to specify symbolic values for both parameters and storage variables, similar to how traditional debuggers let users configure initial states and explore control flow.

We evaluate SolQDebug on real-world functions from Zheng et al. (40), demonstrating millisecond-scale responsiveness under symbolic input. Beyond latency, we analyze how input interval structure affects interpretability in common Solidity patterns, such as division-normalized arithmetic.

This paper makes the following contributions:

- We identify the main barriers to interactive Solidity debugging: latency from compilation, deployment, and transaction setup, and EVM constraints that prevent lightweight re-execution.
- We design an interactive parser and dynamic control-flow graph (CFG) engine that supports live structural updates and syntactic recovery.
- We introduce an abstract interpreter that incorporates developer annotations as symbolic input, supporting fast, deployment-free debugging workflows.

- We implement and evaluate SolQDebug on real-world contracts, demonstrating its millisecond responsiveness and exploring annotation strategies that maintain interpretability under a range of symbolic input patterns.

2 Background

2.1 Structure of Solidity Smart Contract

Solidity smart contracts may declare contracts, interfaces, and libraries. Executable business logic typically resides in contracts, and functions serve as transaction entry points. Variables are usefully grouped as global (EVM metadata such as msg.sender or block. timestamp), state (persistent storage owned by a contract), and local (scoped to a call). Types include fixed-width integers, address, booleans, byte arrays, and user-defined structs; containers include arrays and mappings. A mapping behaves like an associative array with an implicit zero value for unseen keys and is not directly iterable. Storage classes (storage, memory, calldata) indicate lifetime and mutability; we mention them only to fix terminology. Visibility and mutability qualifiers (public, external, internal, private; pure, view, payable) exist but are not central to our single-contract, single-transaction setting. Control flow (if/else, while/for/do-while, break/continue, return) follows C/Java conventions.

Listing 1 Minimal example used to illustrate grammar elements relevant to our analysis

```
1  contract Example {
2      address public owner;
3      uint256 public totalSupply = 1000;
4      mapping(address => uint256) private balances;
5
6      modifier onlyOwner() {
7          require(msg.sender == owner, "not owner");
8          _;
9      }
10
11     function burn(uint256 amount) public onlyOwner {
12         uint256 bal = balances[msg.sender];
13         uint256 delta;
14         if (bal >= amount) {
15             balances[msg.sender] = bal - amount;
16             delta = amount;
17         }
18         else {
19             delta = 0;
20         }
21         totalSupply -= delta;
22     }
23 }
```

The example highlights the specific features we rely on later. State variables include general types (owner, totalSupply) and a mapping from addresses to balances; global variables appear implicitly in guards via msg.sender. The function burn introduces

parameters and a local variable (bal). The modifier onlyOwner performs a precondition check before the function body executes; the placeholder underscore marks where the original body is inserted when the modifier is inlined. In analysis, such modifiers are expanded at their precise positions around the function body in the control-flow graph.

These grammar elements connect directly to our semantics. Guards such as require narrow feasible ranges along taken branches. Modifiers are inlined so that their precondition checks are analyzed in sequence with the function body. Containers like mappings remain symbolic until a concrete key is accessed, at which point an abstract value is materialized for that access. This level of detail suffices for our abstract interpretation in the single-contract, single-transaction scope without introducing parts of the language that our evaluation does not exercise.

2.2 Solidity Execution Model

To execute a Solidity contract on the blockchain, it must first be deployed. Deployment occurs through a one-time transaction that stores the compiled bytecode on-chain and invokes the constructor exactly once. After deployment, all subsequent interactions are message-call transactions. In these, the caller specifies a public function along with encoded calldata. Once the transaction is mined into a block, the Ethereum Virtual Machine (EVM) jumps to the designated entry point and executes the corresponding function sequentially. At runtime, Solidity variables fall into three distinct storage classes (29):

- **Global variables** represent implicit, read-only metadata provided by the EVM, such as `block.timestamp`, `msg.sender`, and `msg.value`.
- **State variables** store persistent data within the contract and retain their values across transactions.
- **Local variables** include function parameters and temporary values scoped to a single execution context.

These three classes share a unified type system comprising primitive types like `uint` , `int`, `bool`, and `address`, as well as composite types such as arrays, mappings, and structs. Composite values can be nested to arbitrary depth using field access (`.`) or indexing (`[]`). Control flow follows familiar C-style constructs such as `if`/`else`, `while`, `for`, and `return`, alongside Solidity-specific statements like `emit` and `revert`.

As a result, debuggers must resolve potentially complex, multi-step expressions to analyze deeply nested elements within the contract state.

2.3 Root Causes of the Solidity-Debugging Bottleneck

Debugging Solidity programs remains significantly slower than traditional application development workflows due to two orthogonal obstacles.

(1) Environmental disconnect. Unlike conventional IDEs such as PyCharm (14) or Visual Studio (20), where the source editor and execution engine run in the same process, Solidity development involves external coordination with a blockchain node at every stage of the workflow. Even a single debugging cycle must pass through four sequential stages (see Fig.1). First, the contract must be compiled. Then, the bytecode

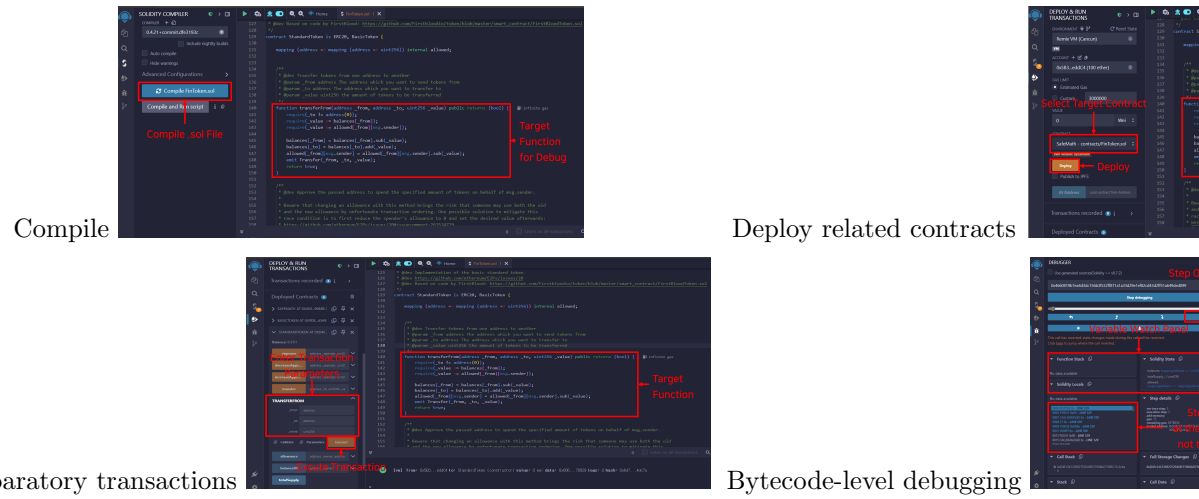

Fig. 1 Traditional Solidity debugging workflow

is deployed to a local or test chain. Next, developers must manually initialize the on-chain state by sending setup transactions. Finally, the target function is invoked, and its execution is traced step by step at the bytecode level.

This workflow introduces several seconds to minutes of latency per iteration, fundamentally breaking the fast "type-and-inspect" feedback cycle expected in modern development tools. To mitigate this friction, developers often rely on `emit` logs or event outputs to observe intermediate values. However, such instrumentation provides only runtime snapshots and lacks the structural insight needed to understand symbolic variation or control-flow behavior. Moreover, modifying the expression of interest typically requires recompilation and redeployment, compounding latency and disrupting iteration. The final stage—tracing raw EVM opcodes—is particularly costly, as developers are forced to mentally reconstruct source-level semantics. This not only adds execution overhead but also imposes significant cognitive burden during fault localization and fix validation.

(2) Architectural limitations of the EVM. The Ethereum Virtual Machine (EVM) is a state-based execution engine in which each transaction mutates a globally persistent storage. Once a function executes, its side effects are irreversible unless external intervention is performed. Re-executing the same function along the same control path is nontrivial: developers must either redeploy the entire contract to restore the initial state, or manually reconstruct the required preconditions via preparatory transactions—both of which incur significant overhead.

Additionally, if a function includes conditional guards that depend on the current state—such as account balances or counters—then any debugging session must first ensure that those conditions are satisfied. Fig. 2 illustrates this challenge: the debug target function enforces a check on `_balances[account]`, requiring developers to manually assign a sufficient balance before they can observe the downstream effects on `_totalSupply`. Without such setup, the function exits early, preventing inspection of the intended execution path.

In short, these constraints make repeated debugging iterations costly and fragile. According to a developer study (42), 88.8% of Solidity practitioners reported frustration with current debugging workflows, with 69% attributing this to the lack of interactive, state-aware tooling.

2.4 Proposed Methodology and Technical Challenges

SolQDebug addresses the two root causes of Solidity's debugging bottleneck—external latency from blockchain round trips, and internal opacity due to storage-based semantics—through a pair of lightweight but complementary techniques.

(1) Eliminating blockchain latency via in-editor interpretation. The traditional debugging workflow requires compilation, deployment, transaction-based state setup, and bytecode tracing—each incurring significant latency. SolQDebug replaces this round trip by performing both parsing and abstract interpretation directly inside the Solidity Editor. To support live editing, we extend the Solidity grammar with interactive parsing rules tailored for isolated statements, expressions, and control-flow blocks. When the developer types or edits code, only the affected region is reparsed using a reduced grammar.

Each parsed statement is inserted into a dynamic control-flow graph (CFG), and abstract interpretation resumes from the edit point. The interpreter uses an interval lattice, assigning each variable a conservative range $[l, h]$ to expose edge conditions (e.g., overflows or failing guards) and to approximate groups of concrete executions that follow the same path. This enables millisecond-scale feedback on code structure and control flow without compilation or chain interaction.

(2) Re-instantiating symbolic state without redeployment. The EVM does not support reverting to a prior state without redeploying the contract or replaying transactions—both of which disrupt iteration. SolQDebug introduces batch annotations as a lightweight mechanism for symbolic state injection. In essence, this reflects a core debugging activity: varying inputs or contract state to observe control-flow outcomes. Rather than reconstructing such conditions through live transactions, developers can write annotations at the top of the function to define initial abstract values. These values are injected before analysis begins and rolled back afterward, ensuring test-case isolation.

This approach brings the debugging workflow closer to the source by making state manipulation explicit and reproducible within the code itself. Developers can explore alternative execution paths by editing annotations alone—without modifying the contract logic or incurring compilation and deployment overhead. It effectively decouples symbolic input configuration from the analysis cycle, while preserving the intuitive debugging process developers already follow.

3 The design of SolQDebug

3.1 System Architecture

SolQDebug receives incremental edits as its primary input—typically a snippet of Solidity code or an inline debug annotation. Rather than expecting complete programs,

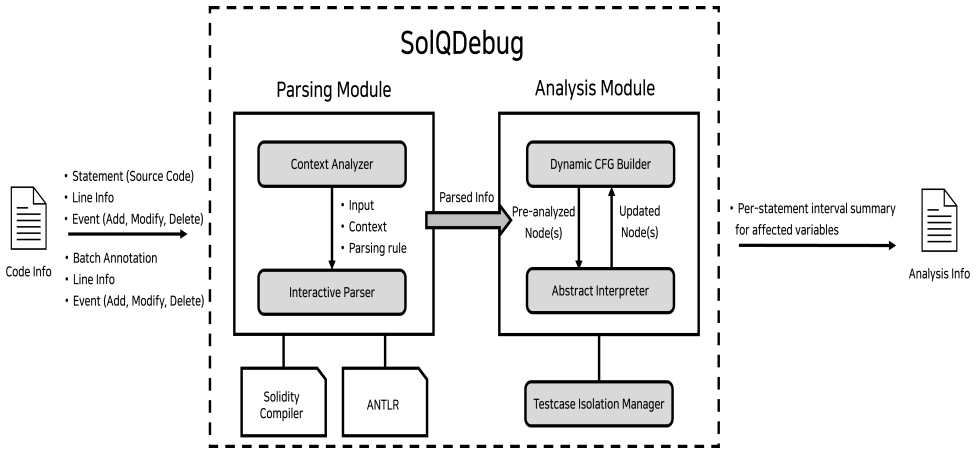

Fig. 2 SolQDebug architecture.

the system is designed to accept fragments ranging from full statements to partial control-flow constructs. These edits include partial Solidity fragments and batch annotations, which are processed in isolation without requiring recompilation or transaction replay. The structure of these inputs is described in Section 3.B; here, we outline the four-stage processing pipeline.

(1) Parsing Module. Each incoming edit first passes through the *Context Analyzer*, which reconstructs a source-level snapshot surrounding the modified lines, determines the enclosing contract or function, and selects the appropriate interactive grammar rule. Subsequently, the Interactive Parser, built atop ANTLR (1), applies an extended grammar that incorporates seven additional reduction rules to support isolated Solidity constructs such as expressions, statements, and definitions. A separate rule is dedicated to debug annotations, allowing single-line analysis directives to be parsed as valid units. To ensure syntactic integrity, the reconstructed source is also verified using the Solidity compiler before analysis proceeds. This allows the system to reject malformed fragments early and maintain consistency across the abstract syntax tree and control-flow graph. Debug annotations are parsed as valid syntactic units and forwarded for interpretation; their semantic effects are described in the analysis stage.

(2) Analysis Module. Each parsed statement is enriched with contextual metadata. This includes its enclosing contract and function, its semantic role (e.g., declaration or condition), and its static type. The statement is then forwarded to the Dynamic CFG Builder, which inserts a basic block at the precise edit point and rewires the surrounding control edges accordingly. Conditional branches merge incoming states using $\sqcup$, and loop headers are updated via localized fixpoint computation. The Abstract Interpreter propagates abstract values using a classic interval lattice. Types such as uintN and intN are interpreted as N-bit intervals. Boolean values are modeled as $\{0, 1\}$ intervals, and addresses as 160-bit unsigned intervals. Byte arrays and strings remain symbolic throughout. For composite containers such as structs, arrays, and mappings, the container itself is treated as symbolic until a specific field,

element, or key is accessed. At that point, the interpreter materializes a fresh abstract value. If the base type is elementary, it receives the corresponding interval; otherwise, a new symbolic placeholder is propagated. Before each batch run, a Test-case Isolation Manager snapshots the full abstract memory. Once execution completes, the snapshot is restored. This guarantees that consecutive test cases remain isolated and do not interfere with each other, even when global or local bindings are modified.

(3) Line-Level Output. After interpretation, the system emits a per-statement summary of relevant variable intervals. This includes:

- **Variable declarations and assignments:** the interval of the updated left-hand side variable.
- **If, else-if, and else branches:** true-path-pruned intervals for variables in single-variable conditions.
- **Require and assert statements:** intervals under the assumption that the condition holds.
- **Function calls:** either the return value (for standalone calls) or the assigned interval (if used on the left-hand side).
- **Loops:** intervals for variables updated within the loop body after fixpoint convergence.

All outputs are mapped to source line numbers and displayed directly in the Solidity editor, providing immediate, deployment-free feedback to developers.

3.2 Running Example

To make the architecture concrete, we walk through a small running example that exercises the main components of SolQDebug. The system operates as follows. First, each incremental source fragment is interpreted under abstract semantics to compute interval for the variables it touches. Second, the corresponding expression is stored in a CFG node that is inserted at a semantically valid point, determined from the edit's context and the existing control-flow. Third, when batch annotations are present, the entire function is reinterpreted using the pre-built CFG with the updated abstract state. We illustrate both modes—incremental source edits and batch annotations—using the burn function from Listing 1, and then refer back to the detailed mechanisms in §§3.3–3.5.

3.2.1 Source Code Analysis Example

Table 1 lists the incremental fragments a developer types for the function burn in Listing 1. SolQDebug accepts two kinds of fragments: (i) block fragments such as a function header or an if/else block, and (ii) single statements that end with a semicolon. Most editors auto-insert a closing brace when "{" is typed, so a block fragment arrives as two lines at once (e.g., function ... { and the matching }). As the body is filled, the auto-inserted closing brace is pushed downward. The line numbers shown in the table refer to the listing 1; intermediate edits may temporarily place the closing brace earlier.

Table 1 Incremental inputs for the running example

Step	Lines of Input Fragment	Fragment
1	11--12	`function burn(uint256 amount) public onlyOwner {` `}`
2	12	`uint256 bal = balances[msg.sender];`
3	13	`uint256 delta;`
4	14--15	`if (bal >= amount) {` `}`
5	15	`balances[msg.sender] = bal - amount;`
6	16	`delta = amount;`
7	18--19	`else {` `}`
8	19	`delta = 0;`
9	21	`totalSupply -= delta;`

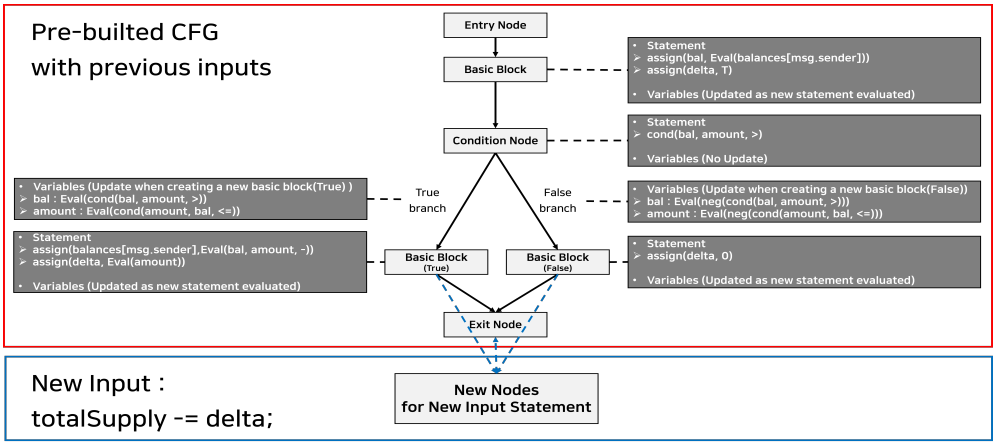

Fig. 3 Pre-built CFG (after Steps 1–8) and dynamic insertion of Step 9

Figure 3 visualizes the state of the CFG after Steps 1–8 have already been integrated and shows how the new input in Step 9 (totalSupply -= delta;) is analyzed and inserted. For clarity, we focus on the function body in this running example and ignore the modifier referenced in the header; modifiers and their placement are treated in §3.5.

SolQDebug uses the following node semantics and bookkeeping rules:

- **Basic blocks.** A basic block contains a straight-line sequence of statements. As statements are evaluated in order, the block maintains the abstract environment at the end of the block—i.e., the interval for each variable. For later re-analysis (e.g., under batch annotations), the block also records its statement list.
- **Condition nodes.** A condition node stores only the predicate (e.g., `bal >= amount`) and does not update the environment at that point.

- **Branch refinement.** When the true/false successors of a condition are created, the incoming environment is pruned along each edge: the true successor refines intervals under the predicate, while the false successor refines them under its negation. Subsequent statements are analyzed under these pruned environments.

With these rules in place, the arrival of Step 9 proceeds as follows:

(1) Parsing and semantics. The interactive parser recognizes `totalSupply -= delta;` as a single assignment and constructs its abstract transfer function (§3.3).

(2) Finding the insertion point. Using the current edit context (line number and the stack of enclosing constructs) together with the existing CFG, SolQDebug locates the semantically valid insertion point. In this case the context indicates the join after the if/else, not merely the preceding line (§3.5).

(3) Join environment (and localized fixpoint if needed). If the insertion point has multiple predecessors, SolQDebug computes the least upper bound ($\sqcup$, JoinBranches) of their environments to build a join block and inserts the new statement there. If a loop header is on the path, a localized fixpoint is computed at the header before joining (§3.5).

(4) Evaluation and rewiring. The assignment is evaluated once in the join environment to produce the new interval for `totalSupply`. The outgoing edges from the two branch leaves are rewired to flow through the join block and then to the exit.

This design allows SolQDebug to reuse all path-local computations accumulated up to the leaves while maintaining semantic correctness at the insertion site. Although the narrative assumes sequential input, the same procedure applies to out-of-order edits (e.g., adding an else later). The insertion-point search, leaf collection, join/fixpoint handling, and edge rewiring remain unchanged and safely update the existing CFG.

3.2.2 Batch Annotation Analysis Example

Listing 2 Burn function with batch annotations

```
1  function burn(uint256 amount) public onlyOwner {
2      // @Debugging BEGIN
3      // @StateVar balances[msg.sender] = [100,200]
4      // @LocalVar amount = [50,150]
5      // @Debugging END
6      uint256 bal = balances[msg.sender];
7      uint256 delta;
8      if (bal >= amount) {
9          balances[msg.sender] = bal - amount;
10         delta = amount;
11     }
12     else {
13         delta = 0;
14     }
15     totalSupply -= delta;
16 }
```

Batch annotations provide a declarative way for developers to specify initial state and parameters as symbolic (interval) values and to obtain line-level results by reinterpreting the program in a single pass over the already built CFG. In this work, "debugging" refers to the interactive exploration during pre-deployment editing in which the developer varies inputs (and state) to observe branch reachability, guard validity, and value bounds. Consequently, variables that are not given an initial range via annotations remain at the conservative $\top$, which can make results vacuous; this underscores the need for meaningful initialization in debugging. Batch annotations supply this initialization in a consistent and reproducible form.

In the running example, we augment the function `burn` in Listing 1 with the following lines: //@StateVar `balances[msg.sender]` $= [100, 200]$ and //@LocalVar `amount` $= [50, 150]$. We set the initial total supply to $totalSupply = 1000$. This choice makes the condition $bal \geq amount$ partially true, so both the then and else branches are reachable; it thereby exposes how pruning at branches and joining after the conditional affect the resulting bounds.

An annotation block is written between `//@Debugging BEGIN` and `//@Debugging END`, with one directive per line of the form "target L-value $\leftarrow$ abstract value (interval or symbolic)." Targets may be global, state, or local variables, and nested L-values (e.g., `a[i].x`, `balances[addr]`) are allowed. Integers are normalized to intervals respecting their declared bit width; addresses are interpreted as 160-bit unsigned intervals; booleans as $\{0, 1\}$.

Given a batch block, SolQDebug executes a lightweight pipeline: (i) parse each line, resolve symbols, and type-check; (ii) snapshot the current abstract memory and overlay the initial environment with the annotated values; (iii) traverse the existing CFG once from the function entry and perform abstract interpretation; condition nodes record only the predicate and do not immediately change the environment, while the true/false successor blocks refine (prune) their incoming environments under the predicate and its negation; if loops are present, a localized fixpoint is computed at the loop header; and (iv) restore the snapshot to guarantee isolation across runs. This process is coordinated by the Debugging Isolation Manager.

Importantly, batch annotations do not alter the CFG structure. No new nodes are inserted; only the initial environment changes, and the same CFG is reused. Each basic block retains its statement list and the abstract environment at the end of the block (interval per variable), enabling fast reevaluation. The rules for branch pruning, least upper bound (LUB) at joins (JoinBranches), and loop fixpoints are identical to those in the source-code example (§3.2.1). On the pre-built CFG in Figure 3, a batch run proceeds "entry $\rightarrow$ branch pruning $\rightarrow$ join $\rightarrow$ exit" in a single pass.

The concrete effect of the above annotations is as follows. From the statement $bal =$ `balances[msg.sender]` we obtain

$$bal \in [100, 200], \qquad amount \in [50, 150].$$

The guard $bal \geq amount$ is only partially true, thus both branches are reachable. After pruning, along the true branch the constraint $bal \geq amount$ raises the lower bound of

$bal - amount$ to 0, yielding

$$\texttt{balances[msg.sender]} := bal - amount \;\Rightarrow\; [0,\, 200-50] = [0, 150], \qquad \delta := amount \;\Rightarrow\; [50, 150].$$

Along the false branch we only set $\delta := 0$, and `balances[msg.sender]` remains at its annotated initial range $[100, 200]$. At the join we compute

$$\delta \in [50, 150] \sqcup [0, 0] = [0, 150], \qquad \texttt{balances[msg.sender]} \in [0, 150] \sqcup [100, 200] = [0, 200].$$

We then evaluate the assignment to the total supply once in the join environment. With $totalSupply = [1000, 1000]$ initially,

$$totalSupply -= \delta \;\Rightarrow\; [1000, 1000] - [0, 150] = [850, 1000].$$

Thus, by combining branch-specific pruning with an LUB at the join, even a simple interval domain avoids unnecessary blow-up (e.g., the negative region of $bal - amount$ is eliminated on the true path) while conservatively aggregating the effects of both paths.

Containers (arrays, mappings, structs) are kept at $\top$ by default and are concretized on access or when a specific key/field is annotated. In our example, the mapping entry `balances[msg. sender]` is concretized by the annotation, and its effects propagate through the branch body and the join. State variables with explicit initial values, such as *totalSupply*, start from a fixed interval, so a single assignment yields directly interpretable bounds.

In summary, batch annotations standardize the essential debugging act of initial state specification via a simple comment syntax, enabling the developer to explore an intended input range in one shot. SolQDebug reuses the CFG and performs a single-pass reinterpretation, delivering lightweight yet semantically sound results. Formal details and algorithms appear in §3.3 and §3.5.

3.3 Interactive Parser

The standard Solidity parser accepts only whole files (`sourceUnit` $\rightarrow$ `EOF`) and thus rejects partial fragments produced during editing. SolQDebug's interactive parser chooses fragment-specific entry rules based on the current editing context and parses each fragment into a syntactically well-formed subtree suitable for incremental analysis. The entry rules fall into two groups: (A) rules for Solidity program fragments (functions, blocks, and other constructs) and (B) rules for debugging-annotation blocks (`debugUnit`). We illustrate both groups with representative inputs.

3.3.1 Entry rules for Solidity program fragments

1) `interactiveSourceUnit` — top-level declaration fragments

- *Purpose.* Accepts top-level snippets such as function headers with empty bodies, contracts, interfaces, libraries, pragmas, imports, and state variables. Editors typically auto-insert a closing brace when "{" is typed, so a "skeleton" declaration arrives as two lines.

- *Example (cf. Table 1, Step 1).*

```
function burn(uint256 amount) public onlyOwner {
}
```

Selected entry: `interactiveSourceUnit`. *Internal match:* `interactiveFunctionElement` → `functionDefinition`.

- *Other top-level examples.*

```
contract Example {}
uint256 public totalSupply = 1000;
```

Selected entry: `interactiveSourceUnit` (matching `contractDefinition` / `stateVariableDeclaration`).

2) `interactiveEnumUnit` **— enum *member* lists added incrementally**

- *Two-phase input.*

1. First, the empty enum *shell* at top level:

```
enum Status {}
```

Selected entry: `interactiveSourceUnit`. *Internal match:* `interactiveStateVariableElement` → `interactiveEnumDefinition`.

2. Then, members are supplied in subsequent fragments:

```
Pending, Shipped

Delivered
```

Selected entry: `interactiveEnumUnit`. *Internal match:* `interactiveEnumItems`.

- *Rationale.* The enum *definition shell* and *member items* are parsed by different entry rules, allowing members to be typed incrementally after the shell is present.

3) `interactiveStructUnit` **— struct *member* declarations added incrementally**

- *Two-phase input.*

1. First, the empty struct *shell*:

```
struct A {}
```

Selected entry: `interactiveSourceUnit`. *Internal match:* `interactiveStateVariableElement` → `interactiveStructDefinition`.

2. Then, members are added one line at a time:

```
uint a;
address owner;
```

Selected entry: `interactiveStructUnit`. *Internal match:* `structMember`.

4) `interactiveBlockUnit` **— block-local statements and skeleton control flow**

- *Purpose.* Accepts semicolon-terminated statements and fully-braced control-flow skeletons typed inside a block or function body.

- *Examples (cf. Table 1, Steps 2,3,5,6,8,9).*

```
uint256 bal = balances[msg.sender];
uint256 delta;
balances[msg.sender] = bal - amount;
delta = amount;
delta = 0;
totalSupply -= delta;
```

Selected entry: `interactiveBlockUnit`. *Internal match:* `interactiveBlockItem` → `interactiveStatement` → `interactiveSimpleStatement`.

- *If-skeleton (cf. Table 1, Step 4).*

```
if (bal >= amount) {
}
```

Selected entry: `interactiveBlockUnit`. *Internal match:* `interactiveBlockItem` → `interactiveIfStatement` .

5) `interactiveDoWhileUnit` — **the *while-tail* of a** `do{...}` **loop**

- *Two-phase input.*

1. First, the `do` body skeleton:

```
do {
}
```

Selected entry: `interactiveBlockUnit`. *Internal match:* `interactiveBlockItem` → `interactiveDoWhileDoStatement`.

2. Then, the `while` tail is typed later:

```
while (i < n);
```

Selected entry: `interactiveDoWhileUnit`. *Internal match:* `interactiveDoWhileWhileStatement`.

6) `interactiveIfElseUnit` — `else` / `else if` **tails**

- *Two-phase input.*

1. First, the `if` skeleton (as above, parsed by `interactiveBlockUnit`).

```
if (cond) {
}
```

2. Then, the tail is added:

```
else {
}
```

or

```
else if (guard) {
}
```

Selected entry: `interactiveIfElseUnit`. *Internal match:* `interactiveElseStatement`.

- *Context handling.* The parser uses the construct stack to attach the tail to the closest unmatched `if`, not merely the preceding line.

7) `interactiveCatchClauseUnit` — `catch` clauses following a `try`

- *Two-phase input.*

  1. First, the `try` skeleton:

     ```
     try doSomething() {
     }
     ```

     Selected entry: `interactiveBlockUnit`. *Internal match:* `interactiveTryStatement`.
  2. Then, one or more `catch` clauses:

     ```
     catch {
     }
     ```

     or

     ```
     catch Error(string memory) {
     }
     ```

     Selected entry: `interactiveCatchClauseUnit`. *Internal match:* `interactiveCatchClause`.

Context-aware selection. A lightweight *construct stack* maintained by the context analyzer (contract/function/block nesting; unmatched `if`/`do`/`try`) guides the choice among these entry rules, ensuring that each edit produces a valid subtree and keeping the global AST/CFG consistent during live editing.

3.3.2 Entry rules for debugging-annotation fragments

`debugUnit` — batch-annotation lines inside `//@Debugging` blocks

- *Purpose.* Parses one or more annotation lines that assign abstract values (intervals or symbolic tags) to designated variables, independent of Solidity AST completeness.
- *Examples.*

  ```
  // @StateVar balances[msg.sender] = [100,200]
  // @LocalVar amount = [50,150]
  ```

  Selected entry: `debugUnit`. *Internal match:* `debugStateVar` / `debugLocalVar`.
- *Usage.* The parsed annotations are consumed by the analysis pipeline to overlay the initial abstract environment and re-interpret the function on the pre-built CFG without recompilation or deployment.

3.4 Dynamic CFG Construction

This section explains how we dynamically extend the control-flow graph while a user edits code. We proceed in three steps. First, we construct and splice a CFG fragment for each statement form and rewire only the neighborhood of the current node. Second,

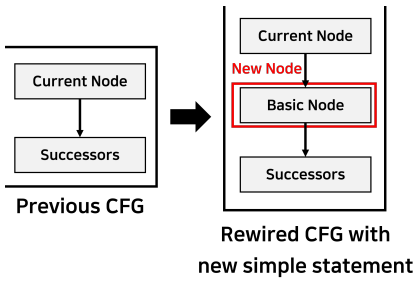

Fig. 4 new simple statement

we locate the insertion site (the current block) using a successor-first, line-aware selection strategy. Third, we re-interpret only the affected region after splicing to update abstract environments.

3.4.1 Statement-Local, Incremental Construction

We introduce the node kinds and then summarize how each major statement is translated into a small CFG fragment and spliced locally.

Node kinds.

- BASIC NODE: holds exactly one statement (e.g., a variable declaration, an assignment, or a function call).
- CONDITION NODE: represents branching constructs such as `if`, `else if`, `while`, `require`/`assert`, and `try`.
- RETURN NODE: a statement node whose outgoing edge is immediately rewired to the function's unique RETURN EXIT.
- ERROR NODE: the function's unique ERROR EXIT (targets the exceptional path).
- FIXPOINT EVALUATION NODE (ϕ): the loop join used for widening and narrowing.
- LOOP EXIT NODE: the false branch that leaves a loop.

Every edit operates at an INSERTION SITE—the block immediately preceding the new fragment—without restructuring the rest of the graph.

Figure 4 shows a simple statement. The builder creates one BASIC NODE and splices it between the current node and the original successors. Incoming environment is copied from the current node; all outgoing edges of the current node are reattached to the new basic node. We deliberately store *exactly one* statement per basic node so that mid-line insertions become $O(1)$ splices via the editor-to-CFG line map, without scanning or splitting multi-statement blocks.

Figure 5 shows an `if`. The builder inserts a CONDITION NODE for the guard, two BASIC NODES for the true/false arms, and an IF JOIN. Edges: current $\rightarrow$ condition; condition $\rightarrow$ true basic (true edge) and $\rightarrow$ false basic (false edge); both basics $\rightarrow$ if join; the if join reconnects to the original successors. Environments on the two edges are refined by the truth value of the guard.

Figure 6 shows an `else if`. The builder removes the previously created false arm of the nearest preceding `if`/`else if` at the same nesting depth and splices a fragment

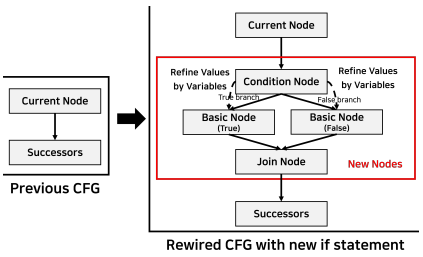

Fig. 5 new if

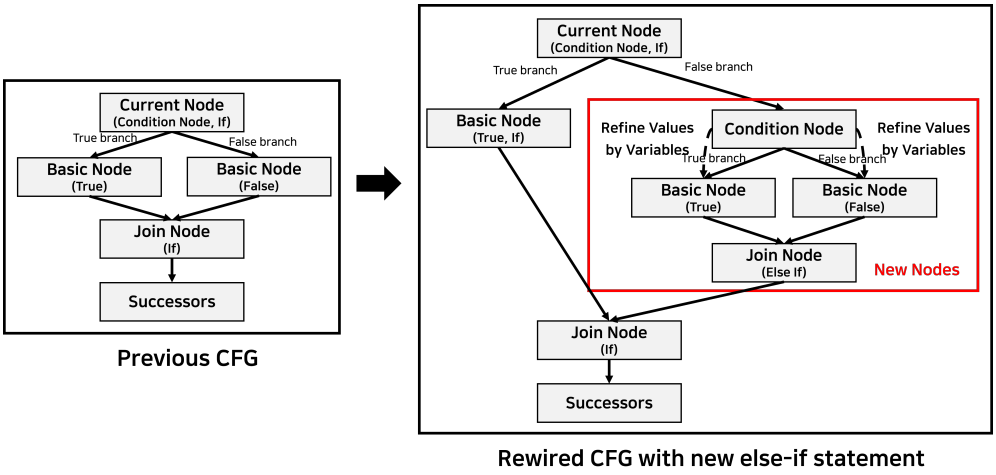

Fig. 6 new else if

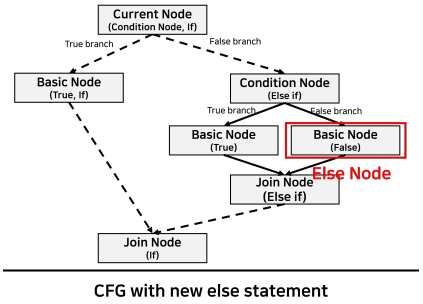

Fig. 7 new else

consisting of a new CONDITION NODE, two BASIC NODES, and an ELSE-IF JOIN. The else-if join is connected to the existing IF JOIN so the overall shape remains a single diamond toward the if join.

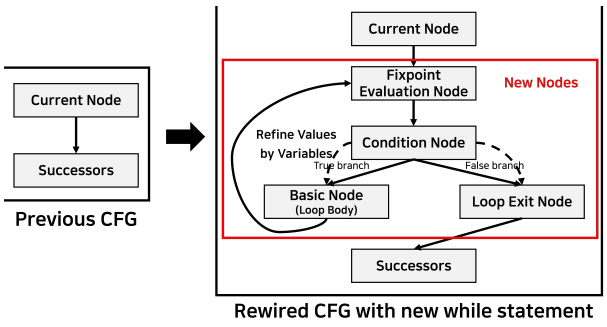

Fig. 8 new while

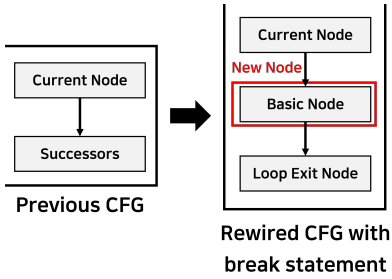

Fig. 9 new break

Figure 7 shows an `else`. No new condition is created; the builder attaches a BASIC NODE to the false branch of the corresponding `if/else if` and connects it to the same IF JOIN as the true branch. The figure assumes a canonical `if/else if/else` chain. For nested patterns (e.g., `if { if {} else {} }`), the else attaches to the false arm of its matching guard according to standard block matching.

Figure 8 shows a `while`. The builder creates a FIXPOINT EVALUATION NODE ϕ, a CONDITION NODE, a true-arm BASIC NODE as the loop-body entry, and a LOOP EXIT NODE (false arm). Rewiring: current $\rightarrow \phi \rightarrow$ condition; condition(true) $\rightarrow$ body; body $\rightarrow \phi$ (back edge); condition(false) $\rightarrow$ loop exit; the loop exit reconnects to the original successors. The ϕ node stores both the pre-loop baseline and the running snapshot for widening/narrowing.

Figure 9 shows a `break`. The statement becomes a BASIC NODE whose outgoing edge is redirected to the LOOP EXIT NODE. The loop exit's environment is conservatively joined with the environment at the break site.

Figure 10 shows a `continue`. The statement becomes a BASIC NODE whose outgoing edge is redirected to the loop's FIXPOINT EVALUATION NODE ϕ. Operationally, this keeps the back-edge shape and joins the current environment into the loop's join state.

Figure 11 shows a `return`. The statement becomes a RETURN NODE and is immediately rewired to the function's unique RETURN EXIT; the return value is recorded there and the original successors of the current node are detached.

Figure 12 shows `require/assert`. The builder inserts a CONDITION NODE for the predicate, makes the true edge point to a BASIC NODE, and connects the false edge

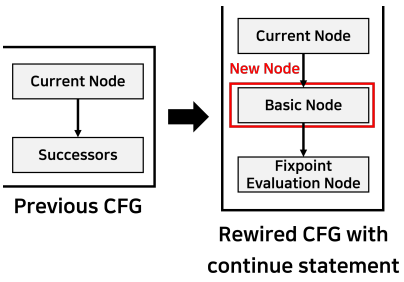

Fig. 10 new continue

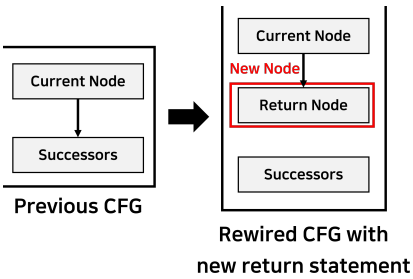

Fig. 11 new return

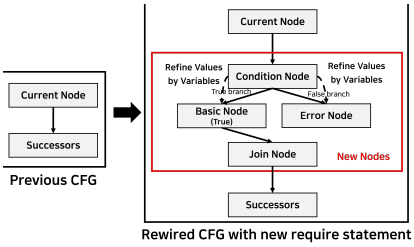

Fig. 12 new require

directly to the function's ERROR EXIT. The true basic then reconnects to the original successors, forming a one-sided diamond.

Other constructs.

`for` loops are handled as a ϕ node and condition like `while`, optionally preceded by an initialization node and followed by an increment node on the back edge. `do{}` `while` is built in two steps: a body pair is created and closed; later the trailing `while` line attaches a ϕ, condition, and loop exit and wires the back edge to the existing body. `try{ } catch{ }` is represented by a CONDITION NODE tagged as `try` whose true edge goes to the success block and whose false edge is replaced by a catch entry/end pair when a matching `catch` appears.

3.4.2 Line-Aware Successor-First Insertion-Site Selection

Line–to–Node Index

We keep a lightweight line–to–node index to make insertion local. Each newly created node is attached to one or two source lines depending on whether the construct is single-line (terminated by ";") or brace-delimited.

Single-line statements (e.g., variable declaration, assignment, function call, require/assert): we index the newly created basic or condition node at the statement's source line.

`if` / `else if`: we index the condition node at the guard's line and the join node at the closing line of the selection.

`else`: we index the basic node of the else arm at the else line and reuse the same join line as the preceding guard.

`while`: we index the condition node at the guard line and the loop-exit node at the closing line of the loop body.

This index is used only to locate the insertion site; the algorithm below does not mutate the graph.

Overview.

The procedure takes the edit span ending at line L and returns the INSERTION SITE—the node immediately preceding the fragment to be spliced. It never mutates the graph. The key idea is SUCCESSOR-FIRST: we first locate the earliest CFG node that appears after L in the source and then decide the most local predecessor that should dominate the new fragment.

Locating the successor (lines 1–2).

FIRSTNODEAFTER(L) scans lines strictly larger than L in the line–to–node index and returns the first node; if none is found, we use the function's unique `EXIT`. We also read its source line ℓ for proximity decisions.

Case 1 — loop exit (lines 4–7).

If s is a loop–exit node, we jump to the header that owns that exit: LOOPHEADERFOREXIT (s) returns the unique loop condition whose false edge points to s. The insertion site is the true branch of this header, i.e., the loop body entry, because newly added statements at this point belong to the body before the loop closes.

Case 2 — join (lines 9–21).

If s is a join, we prefer predecessors that are already joins to preserve nesting. Among those, NEARESTBYLINE picks the one whose source line is closest to ℓ. If no predecessor join exists, we try to recover the immediate guard. GUARDONSAMELINEAS(s) returns the last node indexed at ℓ; if it is not a condition, CONDOFJOIN(s) follows "join ← branch ← cond" to find the guard. BRANCHHINTFROMCONTEXT (ℓ) decides whether we are closing the true or false arm: by default true for `if` /`else-if` and false for `else`. We then return that branch block if available. As a final choice inside this case, we return NEARESTBYLINE among all predecessors of s.

Algorithm 1 Line-Aware Successor-First Insertion-Site Selection (GetInsertionSite)

Require: CFG $G = (V, E)$, edit span ending at line L
Ensure: Insertion-site node $A \in V$ (no graph mutation here)

$s \leftarrow$ FirstNodeAfter(L) ▷ scan lines $> L$ until the first indexed node
if $s = \bot$ **then**
  $s \leftarrow$ EXIT
end if
$\ell \leftarrow$ LineOf(s)
if isLoopExit(s) **then** ▷ closing a loop
  $c \leftarrow$ LoopHeaderForExit(s) ▷ the unique loop condition whose false-edge targets s
  if $c \neq \bot$ **then**
    return BranchBlock(c, true) ▷ the loop body entry (true branch of c)
  end if
else if isJoin(s) **then** ▷ closing a selection or loop body
  $Pred \leftarrow$ Predecessors(s)
  if $\exists p \in Pred$: isJoin(p) **then**
    return NearestByLine($\{p \in Pred \mid \text{isJoin}(p)\}$, ℓ)
  else
    $c \leftarrow$ GuardOnSameLineAs(s) ▷ last node indexed at ℓ
    if $\neg$isCond(c) **then**
      $c \leftarrow$ CondOfJoin(s) ▷ walk join ← branch ← cond
    end if
    if $c \neq \bot$ **then**
      $t \leftarrow$ BranchHintFromContext(ℓ) **default** true ▷ true for `if`/`else if`, false for `else`
      $b \leftarrow$ BranchBlock(c, t)
      if $b \neq \bot$ **then**
        return b
      end if
    end if
    return NearestByLine($Pred$, ℓ)
  end if
else ▷ basic successor
  $Pred \leftarrow$ Predecessors(s)
  if $|Pred| = 1$ **then**
    return the unique element of $Pred$
  else
    return NearestByLine($Pred$, ℓ) ▷ never create a join here
  end if
end if
return ENTRY ▷ defensive fallback (should be unreachable)

Case 3 — basic successor (lines 23–27).

If s is a regular basic node, the insertion site is normally its unique predecessor. When multiple predecessors exist (rare in our skeleton because explicit joins are created earlier), we again choose the predecessor whose source line is closest to ℓ. We never synthesize a join here.

Defensive fallback (line 29).

Although unreachable under well–formed skeletons, we return `ENTRY` as a safety net.

Helper semantics.

PREDECESSORS(s) is the predecessor set; NEARESTBYLINE(X, ℓ) returns $\arg\min_{x \in X} |\text{LINEOF}(x) - \ell|$; GUARDONSAMELINEAS($s$) returns the last node attached to $\ell = \text{LINEOF}(s)$; CONDOFJOIN(s) walks two steps up from the join via the branch to the guard; BRANCHBLOCK(c, t) returns the successor of condition c along truth value t; BRANCHHINTFROMCONTEXT(ℓ) reads the edit context to decide whether the closed branch is true or false.

3.4.3 Change-Driven Reinterpretation

Seed selection per statement.

The builder returns a `single` seed node per edit (never a sink). We use the following mapping:

- `simple` (variable declaration, assignment, function call): the newly inserted basic node.
- `if`: the outer join of the selection.
- `else if`: the outer join of the enclosing guard (fallback to the local join if the outer join is not yet indexed).
- `else`: the same outer join as the true arm.
- `while`/`for`/`do{} while`: the loop exit node (the false branch).
- `continue`: the loop exit node (its environment is conservatively joined with the site).
- `break`: the loop exit node (likewise).
- `return`: a single non-sink dominator that represents the original successors before rewiring to the return exit.[1]
- `require`/`assert`: the true-branch successor along the normal path (if it would be `EXIT` only, reinterpretation is vacuous).

Overview.

Given a single seed node s provided by the builder, the procedure propagates abstract environments forward only along paths that can be affected by the change. It uses a worklist, an in-queue membership set, and a per-node snapshot map *Out* that stores the last observed output environment of each node during this reinterpretation.

[1] In the implementation, the builder may enqueue the original successors directly; conceptually this is equivalent to using one dominating seed.

Algorithm 2 Change-Driven Reinterpretation

Require: CFG $G = (V, E)$; a single seed node s returned by the builder
Ensure: Environments updated along forward-reachable paths from s

$WL \leftarrow \langle\rangle$; $inQ \leftarrow \emptyset$; $Out \leftarrow$ snapshot map ▷ initially current $(\cdot)$
enqueue s into WL; add s to inQ ▷ builder guarantees s is not a sink
while WL not empty **do** ▷ worklist exploration
  $n \leftarrow WL.\mathsf{pop}()$; $inQ \leftarrow inQ \setminus \{n\}$ ▷ FIFO/LIFO is immaterial
  $\hat{\sigma}_{in} \leftarrow \bot$ ▷ accumulator for incoming flow
  for all $p \in$ PREDECESSORS(n) **do** ▷ edge-level pruning for each incoming edge
    if $\mathsf{isCond}(p) \wedge \mathsf{hasTruthLabel}(p \to n)$ **then** ▷ guarded predecessor
      $t \leftarrow \mathsf{edgeLabel}(p \to n)$ ▷ True/False
      $\sigma \leftarrow$ PRUNE$((p), p.\mathsf{cond}, t)$ ▷ refine by guard and truth value
      if FEASIBLE$(\sigma, p.\mathsf{cond}, t)$ **then**
        $\hat{\sigma}_{in} \leftarrow \hat{\sigma}_{in} \sqcup \sigma$
      end if ▷ drop infeasible edges
    else
      $\hat{\sigma}_{in} \leftarrow \hat{\sigma}_{in} \sqcup (p)$ ▷ unconditional predecessor
    end if
  end for
  if $\mathsf{isLoopHeader}(n)$ **then** ▷ handle loops locally by a fresh fixpoint
    $F \leftarrow \{ f \in (n) \mid \mathsf{edgeLabel}(n \to f) = \mathsf{false} \}$ ▷ false (exit) successors
    if $|F| = 1 \wedge \mathsf{isSink}$(the sole f) **then** ▷ trivial fall-through to a sink
      for all $f \in F$ **do**
        if $f \notin inQ$ **then**
          enqueue f; add f to inQ
        end if
      end for
    else
      FIXPOINT(n) ▷ compute loop exit env; widening/narrowing at the header's join
      for all $u \in$ (LOOPEXIT(n)) **do** ▷ resume from the loop exit's successors
        if $u \notin inQ$ **then**
          enqueue u; add u to inQ
        end if
      end for
    end if
    continue ▷ skip the standard transfer at the header
  end if
  $\hat{\sigma}_{out} \leftarrow$ TRANSFER$(n, \hat{\sigma}_{in})$ ▷ identity on joins/conds; apply statements otherwise
  if $\hat{\sigma}_{out} \neq Out[n]$ **then** ▷ change guard: propagate only when output changed
    $(n) \leftarrow \hat{\sigma}_{out}$; $Out[n] \leftarrow \hat{\sigma}_{out}$ ▷ commit the new snapshot
    for all $u \in (n)$ **do**
      if $\neg\mathsf{isSink}(u) \wedge u \notin inQ$ **then**
        enqueue u; add u to inQ
      end if ▷ do not enqueue sinks
    end for
  end if
end while

Edge-level pruning.

For each predecessor p of the current node n, we compute the incoming contribution as follows. If p is a condition and the edge $p \to n$ carries a truth label, we refine (p) with respect to the guard and the edge's truth value, and include it only if the branch is feasible. Otherwise, we join (p) directly. The result $\hat{\sigma}_{in}$ is the join of all feasible incoming flows.

Loop headers.

When n is a loop header, we distinguish the trivial case where its false successor immediately sinks into the function (single false edge to `EXIT`/`RETURN`/`ERROR`) from the general case. In the general case we invoke FIXPOINT, which computes the fixpoint over the loop and updates the environment at the loop exit. We then continue propagation from the successors of that exit. This design ensures that changes inside the body trigger a fresh fixpoint once the header is reached; no special seed augmentation is required.

Change guard.

After applying the transfer function to n, we compare the new output $\hat{\sigma}_{out}$ with $Out[n]$. Only if they differ do we update (n), refresh $Out[n]$, and enqueue the non-sink successors. This guard guarantees termination and avoids needless work. Importantly, it does not miss downstream effects: if an upstream change alters the incoming environment of a node that does not write anything, its output equals the (changed) input, hence $\hat{\sigma}_{out}$ differs from the previous snapshot and successors are still enqueued.

Note on return/revert.

While the implementation may hand multiple old successors back to the engine for `return` /`revert`, the analysis is equivalent to using a single dominating non-sink seed; the algorithm above is presented in this single-seed form for clarity.

3.5 Design of the Abstract Interpretation Framework for Solidity

3.5.1 Program Syntax

Scope of the subset.

We focus on standard structured constructs (variable declarations, assignments, if, while, do–while, for, return, assert/require, delete, and calls as statements). Low-level features (e.g., inline assembly, unchecked arithmetic) are out of scope for this section; modifiers are assumed to be desugared into the control flow.

3.5.2 Concrete Semantics (Denotational)

Domains and helpers.

Let stores be $\sigma : \mathsf{Var} \rightharpoonup \mathsf{CVal}$. L-value resolution $\mathrm{loc}_\sigma(lv) = \ell$ and write $\mathrm{write}(\sigma, \ell, v)$ update the store (arrays/mappings lazily materialize missing cells). Expressions are pure: $e_\sigma \in \mathsf{Val}$.

Table 2 Abstract syntax (subset of Solidity) used by our analysis — meta, literals, and types

Symbol	Set	Definition / Forms
Meta and identifiers		
N	BitW	Integer bit width, $N \in \{8, 16, \ldots, 256\}$.
K	BLen	Fixed byte width, $K \in \{1, \ldots, 32\}$.
x	Var	Program variables (state or local).
f	Field	Struct fields.
C	Struct	Struct identifiers.
E	Enum	Enum identifiers.
Literals		
n	$\mathbb{Z}_{2^N}$	Integer numeral typed by N (signed via unary $-$ when needed).
b	$\mathbb{B}$	$\{\mathsf{true}, \mathsf{false}\}$.
$addr$	$\mathbb{A}$	Address literal $(0 \ldots 2^{160}-1)$.
by	$\mathbb{BY}_K$	Fixed bytes literal (length K); opaque to arithmetic.
Types		
τ_b	ValType	$\mathsf{uint}N \mid \mathsf{int}N \mid \mathsf{bool} \mid \mathsf{address} \mid \mathsf{bytes}K$.
κ	Key	$\mathsf{uint}N \mid \mathsf{int}N \mid \mathsf{address} \mid \mathsf{bytes}K \mid \mathsf{enum}\ E$.
μ	CType	$\mathsf{mapping}(\kappa \Rightarrow \tau) \mid \tau[] \mid \mathsf{struct}\ C \mid \mathsf{enum}\ E$, where $\tau ::= \tau_b \mid \mu$.

We model control effects through an *outcome* domain

$$\mathsf{Res} \ ::= \ (\sigma) \ \mid \ (v, \sigma) \ \mid$$

with a sequencing (Kleisli) operator

$$\begin{aligned} (\sigma) \ &\triangleright \ K := K(\sigma), \\ (v, \sigma) \ &\triangleright \ K := (v, \sigma), \\ &\triangleright \ K := . \end{aligned}$$

We write $s : \sigma \mapsto \mathsf{Res}$ for the denotation of statements.

Statement-by-statement meaning.

Materialization of arrays/mappings.

$\mathrm{loc}_\sigma(a[i])$ for dynamic arrays extends a up to i with default cells if needed; $\mathrm{loc}_\sigma(m[k])$ creates $m[k]$ lazily if absent. These conventions are used by both reads and writes.

Notes. for/do-while are standardly desugared to while. We identify require with assert at the level of control effects (both abort on failure).

Table 3 Abstract syntax (subset) — l-values, expressions, statements, programs

Symbol	Set	Definition / Forms
L-values and expressions		
lv	LVal	$x \mid lv.f \mid lv[idx] \mid lv[key]$.
idx	IdxExp	Numeric index expression (int/uint).
key	KeyExp	Mapping key expression of type κ.
a	AExp	$n \mid addr \mid lv \mid -a \mid \sim a \mid a \oplus a$, $\oplus \in \{+, -, *, /, \%, \ll, \gg, \&, \mid, \wedge\}$.
p	BExp	$b \mid a \bowtie a \mid \neg p \mid p \wedge p \mid p \vee p$, $\bowtie \in \{=, \neq, <, \leq, >, \geq\}$.
Statements		
s	Stmt	skip |s; s|{$\overline{s}$}
		$\tau_b\ x$; |$\tau_b\ x = a$; |lv := a| delete lv
		if p then s else s
		while p do s | do s while p
		for(s_0; p?; u?) s (u is an update as an assignment/compound)
		return a? | assert(p) | require(p)
		call($\overline{a}$) (*external or unknown call, as a statement*)
Programs		
f	Fun	function $id(\overline{x : \tau_b})\ s$ (*single-contract, single-tx*; modifiers desugared).

3.5.3 Abstract Domain

Atomic abstract values.

$$\widehat{\mathbb{U}}_N := \{[\ell, u] \mid 0 \leq \ell \leq u \leq 2^N{-}1\} \ \cup\ \{\bot, \top_N\},$$
$$\widehat{\mathbb{Z}}_N := \{[\ell, u] \mid -2^{N-1} \leq \ell \leq u \leq 2^{N-1}{-}1\} \ \cup\ \{\bot, \top_N^{\pm}\},$$
$$\widehat{\mathbb{B}} := \{\bot, \widehat{\mathsf{false}}, \widehat{\mathsf{true}}, \top\}, \qquad \widehat{\mathbb{A}} := \widehat{\mathbb{U}}_{160},$$
$$\widehat{\mathbb{BY}}_K := \{\bot, \top_K\}, \qquad \widehat{\mathsf{Enum}}(E) := \{[\ell, u] \mid 0 \leq \ell \leq u \leq |E| - 1\} \cup \{\bot, [0, |E| - 1]\}.$$

Order/Join/Meet. For intervals,

$$[\ell_1, u_1] \sqsubseteq [\ell_2, u_2] \iff \ell_2 \leq \ell_1 \wedge u_1 \leq u_2, \quad [\ell_1, u_1] \sqcup [\ell_2, u_2] = [\min(\ell_1, \ell_2), \max(u_1, u_2)],$$

$$[\ell_1, u_1] \sqcap [\ell_2, u_2] = \begin{cases} [\max(\ell_1, \ell_2), \min(u_1, u_2)] & \text{if } \max(\ell_1, \ell_2) \leq \min(u_1, u_2), \\ \bot & \text{otherwise.} \end{cases}$$

Widening ∇ is the standard interval widening (per bit width); narrowing Δ follows the dual pattern.

Composite values.

- *Structs*: $\widehat{\mathsf{Struct}}(C) = \prod_{f \in \mathsf{fields}(C)} \widehat{\mathsf{Val}}_f$ (pointwise order).

Table 4 Concrete denotational semantics (statements)

Statement	Meaning
skip	$\mathsf{skip}(\sigma) = (\sigma)$.
$s_1; s_2$	$s_1; s_2(\sigma) = \big(s_1(\sigma)\big) \triangleright (\lambda\sigma'.\, s_2(\sigma'))$.
$\{\overline{s}\}$	Right-associative fold of sequencing over $\overline{s}$.
$\tau\ x;$	$\tau\ x; (\sigma) = \big(\sigma[x \mapsto \mathrm{zero}_\tau]\big)$.
$\tau\ x = e;$	$\tau\ x = e; (\sigma) = \big(\sigma[x \mapsto e_\sigma]\big)$.
$lv := e$	$lv := e(\sigma) = \big(\mathrm{write}(\sigma,\, \mathrm{loc}_\sigma(lv),\, e_\sigma)\big)$.
delete lv	$\mathsf{delete}\ lv(\sigma) = \big(\mathrm{write}(\sigma,\, \mathrm{loc}_\sigma(lv),\, \mathrm{zero}_{\tau(lv)})\big)$.
if p then s_t else s_f	$\cdot(\sigma) = \begin{cases} s_t(\sigma) & \text{if } p_\sigma = \mathsf{true}, \\ s_f(\sigma) & \text{if } p_\sigma = \mathsf{false}. \end{cases}$
while p do s	Let $F(H)(\sigma) = \begin{cases} \big(s(\sigma)\big) \triangleright H & \text{if } p_\sigma = \mathsf{true}, \\ (\sigma) & \text{if } p_\sigma = \mathsf{false}. \end{cases}$ Then $\mathsf{while}\ p\ \mathsf{do}\ s = \mathrm{lfp}(F)$.
return e	$\mathsf{return}\ e(\sigma) = (e_\sigma,\, \sigma)$.
assert(p), require(p)	$\cdot(\sigma) = \begin{cases} (\sigma) & \text{if } p_\sigma = \mathsf{true}, \\ & \text{if } p_\sigma = \mathsf{false}. \end{cases}$
revert($\cdots$)	$\mathsf{revert}(\cdots)(\sigma) =$.
call($\overline{e}$)	Internal calls evaluate the callee's body with parameter binding; external/unknown calls are left unspecified here (treated in the abstract setting by a conservative effect).

- *Arrays (on-access materialization)*: $\widehat{\mathsf{Arr}}(\tau) = (\hat{\ell}, \hat{d}, M)$ where $\hat{\ell} \in \widehat{\mathbb{U}}_{256}$ is a length summary, $\hat{d} \in \widehat{\tau}$ a default element, and $M : \mathbb{N}_{\mathrm{fin}} \rightharpoonup \widehat{\tau}$ a finite map for observed indices (strong updates).
- *Mappings (on-access materialization)*: $\widehat{\mathsf{Map}}(\kappa \Rightarrow \tau) = (\hat{d}, M)$ with default $\hat{d} \in \widehat{\tau}$ and finite $M : \widehat{\kappa}_{\mathrm{fin}} \rightharpoonup \widehat{\tau}$.
- *Enums*: bounded unsigned intervals over $[0, |E| - 1]$.

Value domains and store.

$$\widehat{\mathsf{Val}} ::= \bigcup_N (\widehat{\mathbb{U}}_N \cup \widehat{\mathbb{Z}}_N) \ \cup\ \widehat{\mathbb{B}} \ \cup\ \widehat{\mathbb{A}} \ \cup\ \bigcup_{K=1}^{32} \widehat{\mathbb{BY}}_K \ \cup\ \widehat{\mathsf{Enum}}(E),$$

$\widehat{\mathsf{CVal}} ::= \widehat{\mathsf{Val}} \mid \widehat{\mathsf{Struct}}(C) \mid \widehat{\mathsf{Arr}}(\tau) \mid \widehat{\mathsf{Map}}(\kappa \Rightarrow \tau)$, $\quad \hat{\sigma} : \mathsf{Var} \rightharpoonup \widehat{\mathsf{CVal}}$ (pointwise order/join).

Design notes.

Addresses are modeled as 160-bit unsigned intervals; strings and dynamic bytes are treated as opaque (symbolic) values when they arise in programs. Arrays and mappings use finite observed maps with defaults, ensuring soundness under unknown indices/keys while enabling strong updates for observed ones.

Table 5 Type → abstract domain mapping (summary)

Solidity type	Abstract domain
$\mathsf{uint}N$	$\widehat{\mathbb{U}}_N \quad (N \in \{8, \ldots, 256\})$
$\mathsf{int}N$	$\widehat{\mathbb{Z}}_N$
bool	$\widehat{\mathbb{B}}$
address	$\widehat{\mathbb{U}}_{160}$
$\mathsf{bytes}K$	$\widehat{\mathbb{BY}}_K$ (opaque)
enum E	$\widehat{\mathsf{Enum}}(E)$
$\tau[]$	$\widehat{\mathsf{Arr}}(\tau)$
$\mathsf{mapping}(\kappa \Rightarrow \tau)$	$\widehat{\mathsf{Map}}(\kappa \Rightarrow \tau)$
struct C	$\widehat{\mathsf{Struct}}(C)$

3.5.4 Abstract Semantics (Denotational)

Abstract domains and results.

Let $\hat{\sigma} : \mathsf{Var} \rightharpoonup \widehat{\mathsf{CVal}}$ be the abstract store (cf. Abstract Domain). Expressions evaluate to abstract values $e^{\sharp}_{\hat{\sigma}} \in \widehat{\mathsf{Val}}$ with bit-width–aware interval arithmetic, booleans, addresses, and composites.

Abstract outcomes:

$$\widehat{\mathsf{Res}} ::= \curlywedge(\hat{\sigma}) \mid \curlywedge(\hat{v}, \hat{\sigma}) \mid \widehat{\curlywedge},$$

ordered componentwise, with sequencing

$$\begin{aligned} \curlywedge(\hat{\sigma}) \ \triangleright^{\sharp} \ K &:= K(\hat{\sigma}), \\ \curlywedge(\hat{v}, \hat{\sigma}) \ \triangleright^{\sharp} \ K &:= \curlywedge(\hat{v}, \hat{\sigma}), \\ \widehat{\curlywedge} \ \triangleright^{\sharp} \ K &:= \widehat{\curlywedge}. \end{aligned}$$

Branch refinement $\mathrm{refine}(\hat{\sigma}, p, b)$ narrows operands of p by interval meets; abstract write $\widehat{\mathrm{write}}(\hat{\sigma}, lv, \hat{v})$ is *strong* if the target cell is unique (e.g., singleton index/key), otherwise *weak* (join with the old value).

For joining branch outcomes we use

$$\mathrm{joinRes}(r_1, r_2) = \begin{cases} \curlywedge(\hat{\sigma}_1 \sqcup \hat{\sigma}_2) & r_i = \curlywedge(\hat{\sigma}_i), \\ \curlywedge(\hat{v}_1 \sqcup \hat{v}_2, \ \hat{\sigma}_1 \sqcup \hat{\sigma}_2) & r_i = \curlywedge(\hat{v}_i, \hat{\sigma}_i), \\ \text{the obvious mixed cases: componentwise join and/or carry } \widehat{\curlywedge}. \end{cases}$$

Statement-by-statement meaning.

Arrays/mappings (reads/writes).

Reading with a singleton index/key returns the cell; with a range/non-singleton key, return the join of materialized cells (or the element type top if none). Dynamic array

Table 6 Abstract denotational semantics (statements)

Statement	Meaning
skip	$\mathsf{skip}^\sharp(\hat{\sigma}) = \widehat{\gamma}(\hat{\sigma})$.
$s_1; s_2$	$s_1; s_2^\sharp(\hat{\sigma}) = \big(s_1^\sharp(\hat{\sigma})\big) \ \triangleright^\sharp \ (\lambda\hat{\sigma}'.\, s_2^\sharp(\hat{\sigma}'))$.
$\tau\ x;$	$\tau\ x;^\sharp(\hat{\sigma}) = \widehat{\gamma}\big(\hat{\sigma}[x \mapsto \hat{\mathrm{init}}(\tau)]\big)$, where $\hat{\mathrm{init}}$ sets $\mathsf{int/uint/bool} \mapsto \bot$, $\mathsf{address} \mapsto \top_{160}$, composites to empty summaries.
$\tau\ x = e;$	$\tau\ x = e;^\sharp(\hat{\sigma}) = \widehat{\gamma}\big(\hat{\sigma}[x \mapsto \alpha_\tau(e^\sharp_{\hat{\sigma}})]\big)$.
$lv := e$	$lv := e^\sharp(\hat{\sigma}) = \widehat{\gamma}\big(\widehat{\mathrm{write}}(\hat{\sigma},\, lv,\, e^\sharp_{\hat{\sigma}})\big)$; non-singleton index/key $\Rightarrow$ weak update (join).
delete lv	delete $lv^\sharp(\hat{\sigma}) = \widehat{\gamma}\big(\widehat{\mathrm{write}}(\hat{\sigma},\, lv,\, \hat{\mathrm{zero}}_{\tau(lv)})\big)$; arrays/maps/structs wiped recursively.
if p then s_t else s_f	Let $\hat{\sigma}_t = \mathrm{refine}(\hat{\sigma}, p, \mathsf{true})$ and $\hat{\sigma}_f = \mathrm{refine}(\hat{\sigma}, p, \mathsf{false})$. Then $\cdot^\sharp(\hat{\sigma}) = \mathrm{joinRes}\big(s_t^\sharp(\hat{\sigma}_t),\ s_f^\sharp(\hat{\sigma}_f)\big)$.
while p do s	Define $G^\sharp(H)(\hat{\sigma}) = \mathrm{joinRes}\big(s^\sharp(\mathrm{refine}(\hat{\sigma}, p, \mathsf{true})) \ \triangleright^\sharp\ H, \widehat{\gamma}(\mathrm{refine}(\hat{\sigma}, p, \mathsf{false}))\big)$. Then $\mathsf{while}\ p\ \mathsf{do}\ s^\sharp = \underbrace{\mathrm{lfp}^\nabla(G^\sharp)}_{\text{widening pass}} \ \triangle \ \underbrace{\mathrm{narrow}^k}_{\textbf{mandatory},\ k\geq 1}$, i.e., compute the widening-based post-fixpoint and then apply at least one narrowing round to regain precision.
return e	return $e^\sharp(\hat{\sigma}) = \widehat{\gamma}(e^\sharp_{\hat{\sigma}},\, \hat{\sigma})$.
assert(p), require(p)	As guards: if p must-hold $\Rightarrow$ $\widehat{\gamma}(\mathrm{refine}(\hat{\sigma}, p, \mathsf{true}))$; if p must-fail $\Rightarrow$ $\widehat{;}$ otherwise both may happen and joinRes carries the possibilities.
revert($\cdots$)	revert$(\cdots)^\sharp(\hat{\sigma}) = \widehat{\cdot}$
call($\overline{e}$)	Internal calls analyze the callee body under parameter binding (same abstract machinery); external/unknown calls conservatively havoc their footprint or are modeled by $\widehat{\cdot}$ per analysis policy.

`length` is a singleton when observed; otherwise conservatively $\top_{\mathsf{uint256}}$. Writes follow the same singleton/non-singleton criterion for strong/weak updates.

4 Evaluation

To evaluate how SolQDebug performs in practical debugging scenarios, we organize our study around three research questions:

- **RQ1 – Responsiveness**: How much edit–to–inspect latency does SolQDebug eliminate compared to Remix?
- **RQ2 – Precision Sensitivity to Annotation Structure**: In a common Solidity pattern where inputs are normalized by division, how does the structure of operand intervals—overlapping vs. distinct—impact interval growth?
- **RQ3 – Loops**: Which loop structures lead to loss of precision, and how do symbolic inputs influence the stability of analysis?

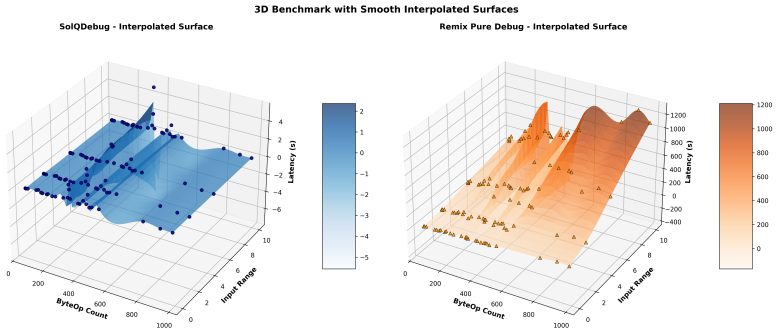

Fig. 13 Edit-to-inspect latency comparison between Remix and SolQDebug across varying test-case widths and execution passes. The x-axis represents the cost estimate, y-axis shows TestCase width (Δ), and z-axis displays latency in seconds. While Remix maintains constant high latency regardless of iteration, SolQDebug demonstrates significantly lower latency that quickly reaches a floor after the initial pass.

4.1 Experimental Setup

We evaluate SolQDebug on a controlled local setup with the following hardware and software configuration:

- **CPU**: 11th Gen Intel® Core™ i7-11390H @ 3.40GHz
- **RAM**: 16.0 GB
- **Operating System**: Windows 10 (64-bit)
- **Implementation Language**: Python

The dataset is derived from DAppSCAN (40), a large-scale real-world benchmark for smart contract analysis. From 3,345 Solidity files using `>=0.8.0`, we sample 128 contracts across three size brackets (1–10 KB, 11–20 KB, and over 20 KB). After filtering out logic-free functions (e. g., those containing only assignments or return statements), we retain 242 single-transaction handlers. From these, we select 13 representative examples covering key Solidity idioms, including structs, mappings, dynamic arrays, control flow, and arithmetic logic.

Although SolQDebug is designed for interactive use within a Solidity editor, all experiments simulate this behavior in a controlled scripting environment. For each function, we reconstruct a sequence of incremental edits and annotations that mimic realistic developer activity. These fragments are streamed into the interpreter to measure latency and interval growth under reproducible conditions.

4.2 RQ1 - Responsiveness

To evaluate responsiveness, we measure edit to inspect latency—defined as the time from a code change to the appearance of updated variable information—under a single contract, single transaction scenario..

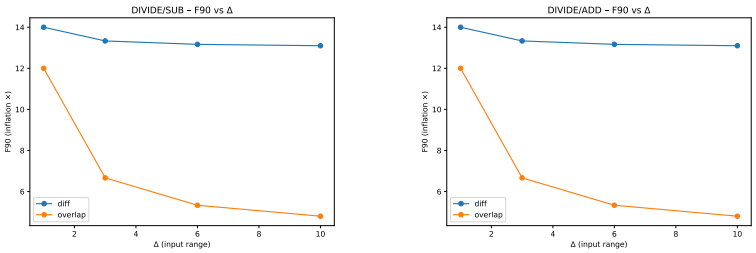

Fig. 14 Interval growth after normalization in `pending` function from `Lock.sol`. Left: original version with subtraction; right: modified version where subtraction is replaced with addition

For Remix, this delay includes a full compile–deploy–execute cycle and is modeled as

$$T_{\text{Remix}} = \underbrace{3}_{\text{compile+deploy}} + \underbrace{2N_s}_{\text{state setup}} + \underbrace{2}_{\text{parameter entry}} + \underbrace{\frac{\text{ops}}{6}}_{\text{EVM trace}} \quad \text{s}$$

where N_s is the number of storage variables that must be manually initialized, and *ops* is the number of executed bytecode operations. Since Remix replays the same steps on every run, the second pass offers no gain: $T^{(2)}_{\text{Remix}} = T^{(1)}_{\text{Remix}}$.

In contrast, SolQDebug updates only the affected region and developer-supplied annotations. Its latency is modeled as

$$T^{(1)}_{\text{SolQ}} = 10v + c_f, \qquad T^{(2)}_{\text{SolQ}} = 5v + c_f \quad \text{s},$$

where v is the number of annotated variables and c_f is a fixed per-function overhead. The coefficients reflect annotation time: first-pass latency includes fresh annotation for each variable (approx. 10 ms per variable), while the second pass assumes partial reuse, as previously annotated variables may be modified rather than newly added. This reflects a realistic debugging workflow in which only a subset of inputs change between passes. Empirically, c_f remains under 0.18 seconds across all tested functions (median: 21 ms, max: 170 ms), confirming that annotation width dominates latency.

We evaluated 13 functions (see Table 1) across 4 test-case widths $\Delta \in \{0, 2, 5, 10\}$ and two passes. Remix required 17.4–41.3 seconds per run (median: 26.2 s), independent of iteration. SolQDebug, by contrast, completed its first pass in 8–62 ms (median: 14 ms), and the second in 5–35 ms. Fig. 6 visualizes this gap in 3D space using a model-derived x-axis (cost estimate), y-axis (TestCase width), and z-axis (latency). While Remix latency scales linearly with test-case complexity, SolQDebug quickly reaches a latency floor after the initial pass.

Since batch execution contributes at most 100 ms, future optimizations should prioritize reducing v—e.g., via annotation templates—rather than optimizing the interpreter core.

4.3 RQ2 - Precision Sensitivity to Annotation Structure

Smart contracts often normalize raw inputs via division—e.g., converting timestamps to time units—before combining the results using addition or subtraction. To isolate the impact of the final arithmetic operator from the shared division step, we analyze two variants of the same control-flow structure: one using addition, the other using subtraction.

Each variant is tested under two annotation styles. In the DIFF style, each operand is assigned a distinct input interval (e.g., $[10, 20]$ and $[30, 40]$). In the OVERLAP style, the intervals are partially aligned (e.g., $[10, 20]$ and $[15, 25]$), such that they share a subrange but are not fully identical. For each combination, we sweep the annotation width $\Delta \in \{1, 3, 6, 10\}$ and report F_{90}, the 90th percentile of the inflation factor $F = \textit{exit\_width}/\textit{input\_width}$.

Results in Fig.7 show that interval growth is more sensitive to the structure of input ranges than to the arithmetic operator. DIFF inputs consistently trigger early widening as Δ increases, while OVERLAP inputs maintain tighter bounds even under addition, which typically increases output range.

This suggests that in division-normalized logic, the alignment of operand intervals—whether disjoint or overlapping—has a stronger influence on interval growth than the choice between addition and subtraction. Overlapping inputs consistently result in smaller output ranges, reducing the degree of over-approximation as input width increases.

4.4 RQ3 - Loops

Two loop patterns emerge from the benchmarks. In the first, the loop condition itself bounds the updated variable, and the loop body performs only direct assignments. In such cases, abstract interpretation converges naturally. For example, `updateUserInfo` in AOC_BEP iterates from 1 to 4, and the interval for `level` stabilizes at $[1, 4]$.

In contrast, loops that interact with external or conditionally populated state tend to diverge under widening. For instance, `revokeStableMaster` in CORE terminates immediately under the contract's default state, but diverges once annotations populate the relevant lists. These lists trigger cascading updates, and their interactions produce imprecision even though the loop count is implicit.

In short, loop precision tends to hold when updates are tightly coupled to bounded loop indices, but approximation still arises due to joins at merge points. Precision degrades further when variables are updated independently of the loop condition. This includes dormant paths activated by symbolic input, or loops that iterate over data-driven structures such as mappings or dynamic lists.

5 Discussion

5.1 Why use Abstract Interpretation for Debugging

In this work, we use debugging to mean a developer-led, interactive exploration activity that happens before deployment during code authoring: the developer varies symbolic (interval) inputs and immediately observes branch reachability, guard validity, and

value bounds at the source level. This edit-time feedback loop calls for a technique that (1) terminates quickly, (2) explains results in a way developers can inspect, and (3) scales to near-keystroke responsiveness.

We chose abstract interpretation (AI) over symbolic execution and proof-based verification for three reasons:

- **Termination.** AI enforces convergence via widening at loops and joins at merges, avoiding the path explosion common in symbolic execution.
- **Explainability.** Each result is an abstract value in a well-defined lattice. With interval domains, the mapping from inputs to outputs is explicit as ranges, which makes dataflow effects easy to trace and debug at the line level.
- **Responsiveness.** Interval transfer functions are lightweight, enabling millisecond-scale updates that fit the edit cycle. Symbolic engines routinely explore many paths even for small edits, which can break interactivity.

Formal verification provides stronger guarantees, but requires fully specified properties and invariants, which are costly to author during early iterations. SolQDebug is designed to bridge the gap between writing code and running tests or verification—offering immediate, sound, conservative feedback with low annotation overhead.

Why the interval domain?

For debugging, intervals strike a practical balance between precision and speed. They (i) align with developers' mental model of "possible ranges," (ii) expose boundary effects (e.g., overflow thresholds, guard satisfaction regions) without committing to a single concrete input, and (iii) compose predictably through joins and widenings. In our setting, intervals are also a natural surface for annotations: developers can *shape* symbolic inputs (e.g., make them overlapping or disjoint) and directly see how that affects control flow and computed ranges.

Managing the accuracy–latency trade-off.

AI's precision is conservative by design; edit-time usability depends on giving developers simple levers to steer precision without sacrificing responsiveness. We expose three such levers that proved effective in our study:

- **Annotation structure.** Overlapping operand intervals often bound output ranges more tightly than disjoint ones in division-normalized arithmetic (cf. RQ2). This reduces false alarms with no runtime cost.
- **Annotation width.** Narrower inputs shrink joins and delay widening; developers can start narrow and broaden gradually ("zoom out") to probe stability.
- **Guard-guided narrowing.** Making explicit the intended `require`/`if` guards in annotations tightens feasible states early and improves precision along the taken branch at negligible cost.

Where stricter precision is essential (e.g., inside data-driven loops), the workflow can temporarily fall back to concrete inputs for local inspection, then return to intervals

for broader exploration. This "concrete when needed, symbolic by default" rhythm preserves interactivity while keeping results actionable.

5.2 Evaluation Implication

RQ1: Edit-time responsiveness, not just "a few seconds faster."

Traditional debuggers (e.g., Remix, Hardhat Debug) require compile–deploy–execute per iteration, typically taking tens of seconds. In contrast, our interpreter updates in milliseconds (median ∼14 ms on the first pass and 5–35 ms on the second), yielding *orders-of-magnitude* lower edit-to-inspect latency. This difference is qualitative: it enables near-keystroke feedback, which changes how developers explore code. Because results are symbolic, a single pass summarizes many concrete executions; developers can see when guards always hold/fail for an interval, when a branch becomes unreachable, or when a value may cross a critical threshold—all without leaving the editor. In short, SolQDebug complements runtime debuggers by moving fast, informative checks *into* the authoring loop.

RQ2: Annotation design as a precision knob.

RQ2 shows that, in division-normalized patterns common in Solidity, *how* intervals are shaped can matter more than *which* arithmetic operator is used. Overlapping inputs systematically produced smaller output ranges than disjoint inputs, delaying or avoiding early widening. Practical takeaway: when investigating arithmetic joins, start with partially overlapping intervals and widen only as needed; keep operands aligned where normalization is present.

RQ3: When loops converge—and when they don't.

Widening can degrade precision in loops, but RQ3 highlights that not all loop updates lead to divergence. Loops whose updated variables are bounded by the loop index (or by monotone guards) often converge to tight ranges quickly; data-driven loops over symbolic containers tend to widen early. Practical takeaway: (i) prefer index-bounded annotations (e.g., bound the iteration count or accessed keys) for loop-local exploration, (ii) materialize only the keys or indices the loop actually touches, and (iii) where necessary, switch to concrete inputs for a small slice of the loop to confirm behavior, then return to symbolic exploration.

Overall, these findings suggest a debugging workflow that starts symbolic and broad, then *shapes* annotations to tighten precision where it matters (overlap, narrow, guard-guided), and finally uses concrete spot checks only for stubborn hot spots (e.g., deeply data-dependent loops).

5.3 Limitation

Our current scope and measurements introduce several limitations.

Scope (external validity).

We focus on single-contract, single-transaction functions. Inter-contract calls, multi-transaction workflows, proxies, and inheritance hierarchies are out of scope in the

present implementation. As a result, we have not yet conducted a developer study in larger project settings; the usability and interpretability of edit-time feedback across multi-contract workflows remain unvalidated.

Measurement (internal validity).

Latency numbers combine interpreter execution time (timed in Python) with an estimate for annotation effort per variable (manual input). This procedure ignores UI-event latency and cursor dynamics, and it assumes a consistent operator for annotation entry. Likewise, our precision metric (F_{90}: 90th percentile of exit-/input-width inflation) captures a salient aspect of interval growth but does not reflect all developer notions of "useful precision." These choices provide a consistent basis for tool-level comparison but may under- or over-estimate end-to-end IDE latency or perceived precision.

Mitigations and future work.

We plan to (i) extend the analysis to inter-contract calls and multi-transaction scenarios, (ii) instrument editor events to directly measure human-in-the-loop latency and refine the annotation cost model, and (iii) run a controlled developer study once multi-contract support stabilizes. On the analysis side, loop summarization and selective use of lightweight relational domains (e. g., applied on demand to hot spots) are promising avenues to improve precision while preserving interactivity.

6 Related Works

6.1 Solidity IDEs and Debuggers

Modern Solidity development environments either embed a debugger or integrate external debugging plug-ins. Remix IDE (23) is the most widely used web IDE; it supports syntax highlighting, one-click compilation, and a bytecode-level debugger that lets users step through EVM instructions and inspect stack, memory, and storage. Hardhat (12) is a Node. js–based framework that couples the Solidity compiler with an Ethereum runtime; its Hardhat Debug plug-in attaches a Remix-style debugger to locally broadcast transactions inside Visual Studio Code. Foundry Forge (8) is a command-line toolchain oriented toward fast, reproducible unit testing; the command `forge test` spins up an ephemeral fork, deploys contracts, executes annotated test functions, and enables replay through Forge Debug. Solidity Debugger Pro (30) is a Visual Studio Code extension that performs runtime debugging over concrete transactions and integrates with Hardhat; in practice, many workflows create a small auxiliary contract that calls the target functions so that state changes can be observed step by step.

In short, these debuggers operate on compiled artifacts or post-deployment traces and rely on transaction replay and EVM-level stepping. They do not accept partial, in-flight source fragments nor provide symbolic (interval) input modeling or millisecond edit-time feedback. By contrast, SolQDebug targets pre-deployment authoring, accepts partial fragments and symbolic annotations, and reports line-level effects via abstract interpretation during editing.

6.2 Solidity Vulnerability Detection and Verification

A rich body of work analyzes smart contracts for security issues using four main families of techniques. Static analysis tools reason over source or bytecode without running the contract. Representative systems include rule- or pattern–based analyzers such as Securify and Slither (35; 36), symbolic-execution–assisted detectors like Mythril (38), knowledge-graph–based reasoning such as Solidet (13), and bytecode CFG refinement as in Ethersolve (22). Dynamic testing and fuzzing exercise deployed or locally simulated contracts to uncover faults and security issues: ContractFuzzer mutates ABI-level inputs (15), Echidna brings property-based fuzzing into developer workflows (9), sFuzz adapts scheduling for higher coverage (21), TransRacer finds transaction-ordering races (18), and Ityfuzz leverages snapshotting to decouple executions from chain nondeterminism (26). Formal verification aims to prove safety properties or refute counterexamples at compile time; examples include ZEUS, VeriSmart, and SmartPulse (16; 27; 34). Finally, AI-based approaches train models to predict vulnerabilities or triage candidates, e. g., via data-flow–aware pretraining, IoT-oriented classifiers, or prompt-tuning for detector adaptation (37; 39; 41).

These approaches have substantially advanced vulnerability detection and property checking for fully written contracts. However, they are not designed to provide interactive, edit-time feedback to developers while code is still under construction. They typically analyze post-compilation artifacts or deployed bytecode and expect complete program units. SolQDebug complements this line of work by focusing on pre-deployment authoring: it accepts partial fragments and symbolic (interval) inputs and produces line-by-line feedback inside the editor.

6.3 Solidity-Specific Abstract Interpretation Frameworks

Abstract interpretation is a well-established framework for static analysis and has been adapted to many programming languages. Two recent studies apply it to Solidity (10; 11). The first uses the Pos domain to construct a theoretical model for taint (information-flow) analysis Halder et al. (10), while the second employs the Difference-Bound Matrix (DBM) domain to generate state invariants and detect re-entrancy vulnerabilities, including the DAO attack (11; 19). However, both approaches operate on fully written contracts and provide no support for line-by-line interpretation or developer interaction within an IDE.

SolQDebug adapts abstract interpretation for an interactive setting. It incrementally updates both the control-flow graph and the abstract state in response to each edit. Developer-supplied annotations serve as a first-class input mechanism, reflecting how debugging often involves varying symbolic inputs. These annotations are internally represented as linear-inequality constraints, and form an integral part of interactive debugging by enabling symbolic reasoning over developer-specified inputs. This design improves interpretability and control within the interval domain by leveraging symbolic constraints, while maintaining keystroke-level responsiveness. As a result, SolQDebug updates variable ranges directly in the Solidity editor, allowing developers to observe how values evolve in response to each edit.

6.4 Interactive Abstract Interpretation for Traditional Languages

In recent years, traditional languages have seen a surge of interest in making abstract interpretation interactive, integrating it directly into IDEs to provide live analysis feedback during editing (4; 7; 24; 32; 33). Stein et al. (32) proposed demanded abstract interpretation, which incrementally rebuilds only the analysis nodes touched by an edit. A follow-up Stein et al. (33) generalized this to procedure summaries, enabling inter-procedural reuse. Erhard et al. (7) extended Goblint with incremental support for multithreaded C, selectively recomputing only genuinely affected facts and maintaining IDE-level responsiveness. Riouak et al. (24) introduced IntraJ, an LSP-integrated analyzer for Java 11 that computes only the AST and data-flow facts needed for the current view, keeping feedback under 100 ms. Chimdyalwar (4) achieved fast yet precise interval analysis on call graphs via one top-down and multiple bottom-up passes, and later introduced an incremental variant that revisits only the impacted functions.

Unlike these frameworks for C or Java, SolQDebug is designed specifically for Solidity. It supports in-flight code fragments and range annotations as first-class input. It incrementally updates only the current basic block in the CFG while reusing previously computed abstract states. Finally, it combines these with an interval domain guided by developer-supplied annotations, which act as input to represent the exploratory nature of debugging. This architecture enables keystroke-level feedback without requiring recompilation, redeployment, or transaction execution. It bridges the gap between Solidity development and the interactive tooling common in traditional programming environments.

7 Conclusion

We introduced SolQDebug, a source-level interactive debugger for Solidity that provides millisecond feedback without requiring compilation, deployment, or transaction replay. By combining interactive parsing, dynamic control-flow graph updates, and interval domain based abstract interpretation seeded by annotations, SolQDebug enables responsive, line-by-line inspection directly within the Solidity editor. Our evaluation shows that it reduces debugging latency compared to Remix, while enabling actionable feedback in response to symbolic inputs. These results demonstrate that SolQDebug's design effectively bridges the interactivity gap in Solidity debugging and brings the development experience closer to that of modern debugging workflows.

Future work includes extending SolQDebug to inter-contract and multi-transaction contexts, incorporating loop summarization for higher precision, and conducting user studies to assess its practical adoption and usability. We also plan to apply analysis based on the EVM Object Format (EOF) to support inter-contract debugging when source code is unavailable, as Ethereum moves toward structured bytecode formats in upcoming hard forks.

References

ANTLR: https://www.antlr.org/ (2025). Accessed September 2025

ChatGPT: https://chatgpt.com/ (2025). Accessed September 2025

Chen, X., et al.: Characterizing smart contract evolution. ACM Transactions on Software Engineering and Methodology (2025)

Chimdyalwar, B.: Fast and precise interval analysis on industry code. In: 2024 IEEE 35th International Symposium on Software Reliability Engineering Workshops (ISSREW) (2024)

ConsenSys Diligence: Python Solidity Parser. https://github.com/ConsenSysDiligence/python-solidity-parser (2025). Accessed September 2025

Cousot, P., Cousot, R.: Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints. In: Proceedings of the 4th ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages (POPL) (1977)

Erhard, J., et al.: Interactive abstract interpretation: reanalyzing multithreaded C programs for cheap. International Journal on Software Tools for Technology Transfer (2024)

Foundry Forge: https://book.getfoundry.sh/reference/forge/forge/ (2025). Accessed September 2025

Grieco, G., et al.: Echidna: effective, usable, and fast fuzzing for smart contracts. In: Proceedings of the 29th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA), pp. 557–560 (2020)

Halder, R., et al.: Analyzing information flow in Solidity smart contracts. In: Distributed Computing to Blockchain, pp. 105–123. Academic Press (2023)

Halder, R.: State-based invariant property generation of Solidity smart contracts using abstract interpretation. In: 2024 IEEE International Conference on Blockchain (2024)

Hardhat: https://hardhat.org/ (2025). Accessed September 2025

Hu, T., et al.: Detect defects of Solidity smart contract based on the knowledge graph. IEEE Transactions on Reliability 73(1), 186–202 (2023)

JetBrains: PyCharm. https://www.jetbrains.com/pycharm/ (2025). Accessed September 2025

Jiang, B., Liu, Y., Chan, W.K.: ContractFuzzer: fuzzing smart contracts for vulnerability detection. In: Proceedings of the 33rd ACM/IEEE International Conference on Automated Software Engineering (ASE) , pp. 259–269 (2018)

Kalra, S., Goel, S., Dhawan, M., Sharma, S.: ZEUS: analyzing safety of smart contracts. In: Proceedings of the 2018 Network and Distributed System Security Symposium (NDSS) (2018)

Llama: https://www.llama.com/ (2025). Accessed September 2025

Ma, C., Song, W., Huang, J.: TransRacer: function dependence-guided transaction race detection for smart contracts. In: Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software

Engineering (ESEC/FSE), pp. 947–959 (2023)

Mehar, M.I., et al.: Understanding a revolutionary and flawed grand experiment in blockchain: the DAO attack. Journal of Cases on Information Technology (2019)

Microsoft Visual Studio: https://visualstudio.microsoft.com/ko/ (2025). Accessed September 2025

Nguyen, T.D., et al.: sFuzz: an efficient adaptive fuzzer for Solidity smart contracts. In: Proceedings of the 42nd ACM/IEEE International Conference on Software Engineering (ICSE), pp. 778–788 (2020)

Pasqua, M., et al.: Enhancing Ethereum smart-contracts static analysis by computing a precise control-flow graph of Ethereum bytecode. Journal of Systems and Software 200, 111653 (2023)

Remix IDE: https://remix.ethereum.org/ (2025). Accessed September 2025

Riouak, I., et al.: IntraJ: an on-demand framework for intraprocedural Java code analysis. International Journal on Software Tools for Technology Transfer (2024)

Rival, X., Yi, K.: Introduction to Static Analysis: an Abstract Interpretation Perspective (2020)

Shou, C., Tan, S., Sen, K.: Ityfuzz: snapshot-based fuzzer for smart contract. In: Proceedings of the 32nd ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA), pp. 322–333 (2023)

So, S., et al.: Verismart: a highly precise safety verifier for Ethereum smart contracts. In: 2020 IEEE Symposium on Security and Privacy (SP), pp. 1678–1694 (2020)

Solidity Compiler in Python (solcx): https://solcx.readthedocs.io/en/latest/ (2025). Accessed September 2025

Solidity documentation: https://docs.soliditylang.org/en/v0.8.29/ (2025). Accessed September 2025

Solidity Debugger Pro: https://www.soliditydbg.org/ (2025). Accessed September 2025

Solidity Language Grammar Rule of SolQDebug : https://github.com/iwwyou/SolDebug/blob/main/Parser/Solidity.g4 . Accessed September 2025

Stein, B., Chang, B.-Y.E., Sridharan, M.: Demanded abstract interpretation. In: Proceedings of the 42nd ACM SIGPLAN International Conference on Programming Language Design and Implementation (PLDI) (2021)

Stein, B., Chang, B.-Y.E., Sridharan, M.: Interactive abstract interpretation with demanded summarization. ACM Transactions on Programming Languages and Systems (2024)

Stephens, J., et al.: SmartPulse: automated checking of temporal properties in smart contracts. In: 2021 IEEE Symposium on Security and Privacy (SP), pp. 555–571 (2021)

Tsankov, P., et al.: Securify: practical security analysis of smart contracts. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS), pp. 67–82 (2018)

Tsankov, P., et al.: Slither: a static analysis framework for smart contracts. In: 2019 IEEE/ACM 2nd International Workshop on Emerging Trends in Software Engineering for Blockchain (WETSEB), pp. 8–15 (2019)

Wu, H., et al.: Peculiar: smart contract vulnerability detection based on crucial data-flow graph and pre-training techniques. In: 2021 IEEE 32nd International Symposium on Software Reliability Engineering (ISSRE), pp. 378–389 (2021)

Yao, Y., et al.: An improved vulnerability detection system of smart contracts based on symbolic execution. In: 2022 IEEE International Conference on Big Data (Big Data), pp. 3225–3234 (2022)

Yu, L., et al.: PSCVFinder: a prompt-tuning based framework for smart contract vulnerability detection. In: 2023 IEEE 34th International Symposium on Software Reliability Engineering (ISSRE), pp. 556–567 (2023)

Zheng, Z., et al.: Dappscan: building large-scale datasets for smart contract weaknesses in dApp projects. IEEE Transactions on Software Engineering (2024)

Zhou, Q., et al.: Vulnerability analysis of smart contract for blockchain-based IoT applications: a machine learning approach. IEEE Internet of Things Journal 9(24), 24695–24707 (2022)

Zou, W., et al.: Smart contract development: challenges and opportunities. IEEE Transactions on Software Engineering (2019)

Algorithm 3 Loop Fixpoint at Header

Require: loop header (condition) node h
Ensure: Converged abstract environments at the loop exit and inside the loop

$L \leftarrow$ TraverseLoopNodes(h) ▷ nodes dominated by h and on some back-edge to h
$vis[\cdot] \leftarrow 0$; $In[\cdot], Out[\cdot] \leftarrow \bot$
$Start \leftarrow \bigsqcup\{(p) \mid p \in (h) \setminus L\}$ ▷ pre-loop env (exclude back-edges)
$In[h] \leftarrow Start$
$\tau \leftarrow$ EstimateIterations$(h, Start)$ ▷ visit threshold for widening

// Widening phase (ascending)
$WL \leftarrow \langle h \rangle$
while $WL \neq \langle\rangle$ **do**
  $n \leftarrow WL.\mathsf{pop}()$; $vis[n] \leftarrow vis[n] + 1$
  $\hat{o} \leftarrow$ Transfer$(n, In[n])$
  if IsJoin$(n) \wedge vis[n] > \tau$ **then**
    $\hat{o} \leftarrow$ Widen$(Out[n], \hat{o})$
  else
    $\hat{o} \leftarrow Out[n] \sqcup \hat{o}$
  end if
  if IsJoin$(n) \wedge$ CondConverged(n) **then** ▷ optional early stop
    $Out[n] \leftarrow \hat{o}$; **break**
  end if
  if $\hat{o} \neq Out[n]$ **then**
    $Out[n] \leftarrow \hat{o}$
    for all $s \in (n) \cap L$ **do**
      $In[s] \leftarrow \bigsqcup \{$ Flow$(p \to s) \mid p \in (s) \cap L \}$ ▷ edge-pruned join
      $WL.\mathsf{push}(s)$
    end for
  end if
end while

// Narrowing phase (descending)
$WL \leftarrow$ any worklist ordering over L
while $WL \neq \langle\rangle$ **do**
  $n \leftarrow WL.\mathsf{pop}()$
  $\hat{o} \leftarrow$ Transfer$(n, In[n])$
  if IsJoin(n) **then**
    $\hat{o} \leftarrow$ Narrow$(Out[n], \hat{o})$ ▷ at least one round; cap by $k_{\max}$
  end if
  if $\hat{o} \neq Out[n]$ **then**
    $Out[n] \leftarrow \hat{o}$
    for all $s \in (n) \cap L$ **do**
      $WL.\mathsf{push}(s)$
    end for
  end if
end while
return Out ▷ in particular (LoopExit(h)) is now converged