Modular Static Analysis of Actor Programs
Stiévenart, Quentin; Nicolay, Jens; De Roover, Coen; De Meuter, Wolfgang

Publication date:
2017

Document Version
Peer reviewed version

Citation for published version (APA):
Modular Static Analysis of Actor Programs

Quentin Stiévenart, Jens Nicolay, Wolfgang De Meuter, Coen De Roover
{gstieven,jnicolay,wdeMEUTER,cderoove}@vub.ac.be
Vrije Universiteit Brussel, Belgium

Abstract—Existing static analyses for actor programs explicitly model all possible execution interleavings. Because the number of interleavings increases exponentially with the number of actors and messages exchanged, such analyses scale poorly. We describe the first modular analysis for actor programs, that analyzes each actor in separation of each other. This analysis over-approximates over the different interleavings without explicitly modeling them, rendering it scalable. We demonstrate its enhanced scalability by comparing the analysis of the Savina benchmark suite with a non-modular analysis and our new modular analysis. Our technique succeeds in analyzing all of the Savina benchmarks in a matter of seconds, while the non-modular analysis times out on more than half of these benchmarks. Moreover, we show that the precision of our modular remains on par with the precision of the non-modular analysis.

I. INTRODUCTION

Actor programs consist of isolated processes, called actors, that run concurrently to each other and communicate through messages. At run time, actors can change their behavior, create an unbounded number of other actors, and communicate actor references through messages. This results in a possibly highly dynamic topology for actor programs.

Existing static analyses for actor programs explicitly represent every possible interleaving of concurrent actor execution, and are therefore limited in scalability. Modular analyses allow for a scalable analysis, and have been explored in the context of shared-memory concurrent programs, limited to systems where the number of processes is fixed and known in advance. We explore modular analysis for actor programs. Compared to existing work for shared-memory concurrent programs, we need to account for dynamic creation of new actors.

The core insight behind our modular analysis is that the dynamic behavior of an actor is entirely defined by its code and the set of messages it receives. While the code is known beforehand, the set of messages received must be inferred by analyzing other actors. The goal of the analysis at the level of an actor is to infer the set of created actors and messages sent to other actors for every possible message the actor understands. We refer to this analysis as the intra-actor analysis, an analysis that analyzes one actor in isolation and infers the interferences generated by this actor. The intra-actor analysis enables a modular analysis, at the level of the entire program, to infer the set of running actors and the set of messages received by each running actor. We refer to this analysis as the inter-actor analysis, an analysis that uses information from the intra-actor analysis of a set of actors to drive the next set of intra-actor analyses on a more complete set of actors. Repeated iterations of the inter-actor analysis will achieve a sound over-approximation of the behavior of all actors in the actor program in a scalable manner.

Our approach infers properties about concurrent actor programs that are required for tool support for pressing problems in software engineering such as program comprehension and bug detection. The inferred properties concern the actors created, the messages exchanged, and the behaviors changed at run time—in addition to the traditional data and control flow properties known from analyses for sequential programs. We demonstrate that our approach scales beyond the current state-of-the-art, as the first analysis to analyze the entire Savina benchmark suite [3] in a matter of seconds, with the same precision.

II. DYNAMIC BEHAVIOR OF ACTOR PROGRAMS

Consider the program in Listing 1, implemented in a Scheme dialect that supports concurrent actors [4]. The program computes and displays the factorial of a number.

After the definition of three behaviors through the actor construct (lines 2, 11, and 16), two actors are spawned through the create construct: actor fact with behavior fact-actor (line 20), and actor displayer that displays the messages it receives (line 21). On line 22 the number returned by read-integer (assumed non-negative) is sent to actor fact along with actor displayer, the customer that will receive the answer that fact computes.

When actor fact receives the compute message with \( n = 0 \), it sends 1 as answer to the customer (line 5). Otherwise \( n \neq 0 \), and fact spawns an actor with behavior customer-actor (line 6), passing \( n \) and fact’s customer as arguments. Actor fact then sends itself a compute message to compute \((n-1)!\), passing along the newly created customer (line 8).

Actors with the customer-actor behavior multiply the number they receive by the number given at their creation, and send the result to the customer given at their creation (line 13). Putting everything together, when fact receives the initial compute message with number \( n \) and customer displayer, it creates a customer \( c \) with \( n \) and displayer as arguments, and computes \((n-1)!\) by sending a compute message to itself with \( n-1 \) and \( c \) as arguments. When the computation of \((n-1)!\) is completed, the result is sent to customer \( c \), which computes \( n\cdot(n-1)! = n! \), and sends the result to displayer, which displays the result on the screen.

Figure 1 depicts the evolution of the actor topology throughout the execution of this program. We see that the fact actor creates a chain of customers that propagate and multiply a value until the computed factorial reaches the displayer actor.
At the start of the first iteration no actor has been created yet, so only the main process is analyzed. Because the main process does not receive any messages, it can be fully analyzed in this first iteration. The intra-actor analysis deduces that the main process performs the following actions:
- it creates two actors: one with process identifier $p_{fact}$ and behavior fact-actor (line 20), and one with process identifier $p_{disp}$ and behavior displayer-actor (line 21), and
- it sends a message with tag compute and two arguments (an integer and $p_{disp}$) to $p_{fact}$ (line 22).

**Iteration 2.**

The main process does not have to be analyzed in this iteration, because it was already fully analyzed and cannot receive messages. As a message was sent to a newly created actor $p_{fact}$ in the previous iteration, this iteration now analyzes the code of this actor with behavior fact-actor and a mailbox containing the message compute($Int, p_{disp}$). The body of the processor for message compute is analyzed with its parameters bound to the given values, and the intra-actor analysis of actor $p_{fact}$ infers the following actions:
- send the message result($Int$) to customer (bound to $p_{disp}$).
- create a new actor with process identifier $p_{cust}$ and behavior customer-actor, and
- send the message compute($Int, p_{cust}$) to itself.

This second iteration also has to analyze $p_{disp}$ created in the previous iteration. Because its mailbox is empty, no code can be executed in this actor and the analysis trivially finds no new effects. This trivial analysis is omitted from Table I.

**Iteration 3.**

The third iteration re-analyzes actor $p_{fact}$ because this actor
Analysis result.

After reaching a fixed point, the analysis has discovered every
of actually running a λα program.

1https://github.com/acieroid/scala-am/tree/modularactors

received a new message \(\text{compute}(\text{Int}, p_{\text{cust}})\). The behavior
inferred by the intra-actor analysis is now more complete than
in the previous iteration as the following additional behavior is
discovered:

- send the message \(\text{result}(\text{Int})\) to \(p_{\text{cust}}\),
- create a new actor with process identifier \(p_{\text{cust}}\) and
  behavior customer-actor(\(\text{Int}, p_{\text{cust}}\)).

The third iteration also has to analyze \(p_{\text{cust}}\) with an empty
mailbox, and trivially finds no new effects.

**Iteration 4.**

In the fourth and final iteration, \(p_{\text{fact}}\) is not re-analyzed as no
new effects affecting this actor have been discovered. Therefore,
the analysis has reached a fixed point for this actor.

Actor \(p_{\text{cust}}\) does have to be re-analyzed with a
non-empty mailbox containing \(\text{result}(\text{Int})\), and with
two possible behaviors: customer-actor(\(\text{Int}, p_{\text{cust}}\)) and
customer-actor(\(\text{Int}, p_{\text{disp}}\)). As two distinct behaviors can
process the message, the intra-actor analysis also infers two
possible message sends.

The newly sent message \(\text{result}(\text{Int})\) has already been
taken into account for \(p_{\text{cust}}\) and \(p_{\text{disp}}\) in the previous iteration,
hence the analysis has reached a fixed point and no additional
intra-actor analyses will be triggered.

**Analysis result.**

After reaching a fixed point, the analysis has discovered every
actor that is created and every message that is sent to or received
by each actor, with information about the values contained in
the messages. Figure 2 depicts the corresponding actor topology
that has been inferred, with the abstract actor states and abstract
message values as expected. Note that the analysis has merged
the unbounded chain of customers into a single abstract actor.

### IV. Experimental Evaluation

To evaluate our approach, we implemented the analysis in
Scala on top of the SCALA-AM static analysis framework [5].
Our implementation is available online\(^1\). We also implemented
a concrete interpreter for the input language in Racket. This
enables comparing the results of the analysis with the results
of actually running a λα program.

![Fig. 2: An over-approximation of the actor topology's evolution, as inferred by the analysis. Actors denoted by disp and cust respectively denote actors with behaviors displayer-actor and customer-actor.](image_url)

For our experiments, we translated all of the Savina benchmark
programs into our Scheme-based input language. The original Savina benchmark programs, written in Scala (with
fragments in Java), range from 102 to 616 lines of code. After
translation into the input language, they range from 17 to 293
lines of code. Benchmarks of this size are beyond what is
currently supported by existing static analyses. To support that
claim, we compare our modular analysis with a non-modular
analysis on the same set of benchmark programs and report
on the results.

We ran each benchmark 1000 times with our concrete inter-
preter, and recorded every actor that is created, every become
statement that is executed, and every message that is received.
We then compared these observations with the information
computed by the analysis. This enables the soundness and
precision of our implementation to be respectively verified and
measured empirically.

**A. Performance**

To assess performance of our modular analysis, we analyzed
each Savina benchmark program two different analyses.

1) Our implementation of the modular analysis presented
here.

2) Our previous non-modular analysis MBA [4], which pro-
vides several mailbox abstractions that vary in precision.

<table>
<thead>
<tr>
<th>#</th>
<th>Pid</th>
<th>Behavior</th>
<th>Message received</th>
<th>Messages sent</th>
<th>Actors created</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(p_{\text{main}})</td>
<td>main</td>
<td>–</td>
<td>({p_{\text{fact}}, \text{compute}(\text{Int}, p_{\text{disp}})}) ({p_{\text{disp}}, \text{disp}(\text{Int})})</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>(p_{\text{fact}})</td>
<td>fact()</td>
<td>compute((\text{Int}, p_{\text{disp}}))</td>
<td>({p_{\text{disp}}, \text{result}(\text{Int})}) ({p_{\text{fact}}, \text{compute}(\text{Int}, p_{\text{cust}})})</td>
<td>(p_{\text{cust}}, \text{cust}(\text{Int}, p_{\text{disp}}))</td>
</tr>
<tr>
<td>3</td>
<td>(p_{\text{disp}})</td>
<td>disp()</td>
<td>compute((\text{Int}, p_{\text{cust}}))</td>
<td>result((\text{Int}))</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>(p_{\text{cust}})</td>
<td>cust((\text{Int}, p_{\text{cust}}))</td>
<td>result((\text{Int}))</td>
<td>({p_{\text{cust}}, \text{result}(\text{Int})}) ({p_{\text{disp}}, \text{result}(\text{Int})})</td>
<td>–</td>
</tr>
</tbody>
</table>

Table I: Results of analyzing the program from Listing 1. The # column denote the inter-actor analysis iteration. Columns *Pid* and *Behavior* denote the process identifier under analysis with the given behavior. Column *Message received* denotes the message for which the actor is analyzed, and columns *Messages sent* and *Actors created* give the effects inferred by the intra-actor analysis.

---

\(^{1}\)https://github.com/acieroid/scala-am/tree/modularactors
We used MBA\textsubscript{Set}, that is MBA with the least precise, set-based, mailbox abstraction, to match the mailbox abstraction used by our modular approach.

Both analyses were executed under Scala 2.12.2 using Java 1.8.0_102 on a Mid-2014 MacBook Pro with a 2.8 GHz Intel Core i7 and 16 GB of RAM, and we report on the average timing of 20 runs after 10 warmup runs. We employed a timeout of two minutes, after which we denote an analysis time as infinite (\(\infty\)). If the analysis completes under 0.01 second, we denote the analysis time as \(\epsilon\).

Table II lists the results of our experiments, which give a general idea of the scalability and precision of each analysis. Although the analyses differ slightly on what is computed, it is clear that our modular technique scales to the entirety of the Savina benchmark suite, analyzing it in a matter of seconds.

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>LOC</th>
<th>Observed</th>
<th>t</th>
<th>Spurious</th>
<th>t</th>
<th>Spurious</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP</td>
<td>27</td>
<td>9</td>
<td>0.03</td>
<td>0</td>
<td>0.11</td>
<td>0</td>
</tr>
<tr>
<td>COUNT</td>
<td>29</td>
<td>8</td>
<td>0.02</td>
<td>0</td>
<td>0.09</td>
<td>0</td>
</tr>
<tr>
<td>FJT</td>
<td>38</td>
<td>3</td>
<td>0.02</td>
<td>0</td>
<td>0.03</td>
<td>0</td>
</tr>
<tr>
<td>FJC</td>
<td>17</td>
<td>2</td>
<td>(\epsilon)</td>
<td>0</td>
<td>(\epsilon)</td>
<td>0</td>
</tr>
<tr>
<td>THR</td>
<td>43</td>
<td>5</td>
<td>0.02</td>
<td>2</td>
<td>(\infty)</td>
<td>–</td>
</tr>
<tr>
<td>CHAM</td>
<td>81</td>
<td>10</td>
<td>0.05</td>
<td>0</td>
<td>(\infty)</td>
<td>–</td>
</tr>
<tr>
<td>BIG</td>
<td>52</td>
<td>10</td>
<td>0.06</td>
<td>2</td>
<td>(\infty)</td>
<td>–</td>
</tr>
<tr>
<td>CDICT</td>
<td>67</td>
<td>13</td>
<td>0.09</td>
<td>1</td>
<td>21.25</td>
<td>1</td>
</tr>
<tr>
<td>CSLI</td>
<td>61</td>
<td>16</td>
<td>0.08</td>
<td>1</td>
<td>(\epsilon)</td>
<td>1</td>
</tr>
<tr>
<td>PCB</td>
<td>98</td>
<td>13</td>
<td>0.66</td>
<td>0</td>
<td>102.99</td>
<td>0</td>
</tr>
<tr>
<td>PHIL</td>
<td>58</td>
<td>13</td>
<td>0.05</td>
<td>0</td>
<td>(\infty)</td>
<td>–</td>
</tr>
<tr>
<td>SBAR</td>
<td>77</td>
<td>19</td>
<td>0.08</td>
<td>1</td>
<td>(\epsilon)</td>
<td>–</td>
</tr>
<tr>
<td>CIG</td>
<td>49</td>
<td>9</td>
<td>0.05</td>
<td>1</td>
<td>1.40</td>
<td>1</td>
</tr>
<tr>
<td>LOGM</td>
<td>106</td>
<td>15</td>
<td>0.11</td>
<td>7</td>
<td>(\infty)</td>
<td>–</td>
</tr>
<tr>
<td>BTX</td>
<td>61</td>
<td>9</td>
<td>0.08</td>
<td>0</td>
<td>2.50</td>
<td>0</td>
</tr>
<tr>
<td>RSORT</td>
<td>60</td>
<td>10</td>
<td>0.03</td>
<td>0</td>
<td>12.98</td>
<td>0</td>
</tr>
<tr>
<td>FBANK</td>
<td>143</td>
<td>38</td>
<td>0.15</td>
<td>0</td>
<td>(\infty)</td>
<td>–</td>
</tr>
<tr>
<td>SIEVE</td>
<td>37</td>
<td>8</td>
<td>0.03</td>
<td>0</td>
<td>0.09</td>
<td>0</td>
</tr>
<tr>
<td>UCT</td>
<td>145</td>
<td>20</td>
<td>0.85</td>
<td>11</td>
<td>(\epsilon)</td>
<td>–</td>
</tr>
<tr>
<td>OPL</td>
<td>293</td>
<td>13</td>
<td>5.75</td>
<td>0</td>
<td>(\epsilon)</td>
<td>–</td>
</tr>
<tr>
<td>TRAPR</td>
<td>72</td>
<td>7</td>
<td>0.08</td>
<td>0</td>
<td>10.24</td>
<td>0</td>
</tr>
<tr>
<td>PIPREC</td>
<td>74</td>
<td>8</td>
<td>0.05</td>
<td>0</td>
<td>0.38</td>
<td>0</td>
</tr>
<tr>
<td>RMMI</td>
<td>113</td>
<td>14</td>
<td>0.45</td>
<td>0</td>
<td>(\epsilon)</td>
<td>–</td>
</tr>
<tr>
<td>QSORT</td>
<td>69</td>
<td>6</td>
<td>1.41</td>
<td>0</td>
<td>(\epsilon)</td>
<td>–</td>
</tr>
<tr>
<td>APSP</td>
<td>188</td>
<td>5</td>
<td>1.06</td>
<td>1</td>
<td>(\epsilon)</td>
<td>–</td>
</tr>
<tr>
<td>SOR</td>
<td>201</td>
<td>12</td>
<td>2.39</td>
<td>31</td>
<td>(\epsilon)</td>
<td>–</td>
</tr>
<tr>
<td>ASTAR</td>
<td>92</td>
<td>11</td>
<td>0.26</td>
<td>0</td>
<td>(\epsilon)</td>
<td>–</td>
</tr>
<tr>
<td>NQN</td>
<td>106</td>
<td>11</td>
<td>0.55</td>
<td>0</td>
<td>(\epsilon)</td>
<td>–</td>
</tr>
</tbody>
</table>

Fully analyzed: 28/28 12/28

Table II: Scalability of our modular analysis and a non-modular analysis. The LOC column indicate the length of the benchmarks a Scheme-based input language. Column Observed reports the number of different elements observed at run-time (messages received, become statements, actors created) in 1000 runs of a benchmark program. For each analysis, we list the time taken to analyze each benchmark; and the number of spurious elements that have been reported by the analysis. Timings are in seconds where \(\epsilon\) represents a time strictly smaller than 0.01s and \(\infty\) a time out, which was set to two minutes (120s). The last row shows how many benchmarks could be analyzed by each analysis within the given time limit.

B. Soundness testing

We provide empirical evidence for the soundness of our implementation through soundness testing [2]. To this end, we verified that all information recorded during the concrete runs of each benchmark program is indeed over-approximated by the analysis of the same program. No unsound results were reported for any of the benchmarks, i.e., the analysis implementation over-approximated every value that was observed during the concrete executions.

C. Precision

Similar to the empirical verification of soundness, we measured the precision of the analysis. We compared the maximum-precision abstraction of the observed values in concrete runs of the program with the abstract values computed by the analysis.

Table II reports on the number of spurious values computed by the analysis. These are values that do not correspond to any value observed during any concrete execution. We also produced the same evaluation for our previous non-modular technique [4], and report on the spurious elements computed in benchmark programs for which the analysis does not time out. We observe that on these benchmarks, the precision of both techniques is identical: the same number of spurious elements appears in all benchmarks supported by the non-modular technique. Overall, the precision of our modular analysis on the Savina benchmark suite is 84% (317 true positives and 61 false positives).

The reported numbers are merely an upper-bound on the number of potential false positives, as they correspond to elements that have not been observed in any concrete executions but that may be present in non-explored executions.

V. Conclusion

This paper presents a modular static analysis for concurrent actor programs that scales to all programs in the Savina benchmark suite. Relying on a sequential analysis to perform intra-actor analysis, we explain how a complementary inter-actor analysis uses the effects inferred by the intra-actor analysis to run the intra-actor analysis on other actors. Eventually, each actor is analyzed with an over-approximation of its mailbox, and the whole inter-actor analysis terminates. This yields a sound analysis that does not explicitly model every possible interleaving and, unlike existing techniques, scales well.

We evaluated our modular analysis through its implementation on top of an existing static analysis framework and show that, unlike non-modular approaches, it supports the entire Savina benchmark suite. The modular analysis generates few spurious elements and achieves a precision of 84% with respect to detecting messages sends, actor creations, and executions of become statements.

In summary, this paper demonstrates that it is possible to construct a scalable, sound, and precise analysis for actor programs. This is achieved by taking advantage of the fact that actor programs are made of actors that only interfere through messages. By relying on a sequential analysis for single actors, we are able to build a modular, scalable analysis for actor programs.
REFERENCES


