



The Probabilistic Termination Tool Amber

Marcel Moosbrugger¹(✉), Ezio Bartocci¹, Joost-Pieter Katoen²,
and Laura Kovács¹

¹ TU Wien, Vienna, Austria

marcel.moosbrugger@tuwien.ac.at

² RWTH Aachen University, Aachen, Germany

Abstract. We describe the AMBER tool for proving and refuting the termination of a class of probabilistic while-programs with polynomial arithmetic, in a fully automated manner. AMBER combines martingale theory with properties of asymptotic bounding functions and implements relaxed versions of existing probabilistic termination proof rules to prove/disprove (positive) almost sure termination of probabilistic loops. AMBER supports programs parameterized by symbolic constants and drawing from common probability distributions. Our experimental comparisons give practical evidence of AMBER outperforming existing state-of-the-art tools.

Keywords: Almost sure termination · Martingales · Asymptotic bounds

1 Introduction

Probabilistic programming obviates the need to manually provide inference methods and enables rapid prototyping [13]. Automated formal verification of probabilistic programs, however, is still in its infancy. Our tool AMBER provides a step towards solving this problem when it comes to automating the termination analysis of probabilistic programs, which is an active research topic [1, 6, 7, 9–12, 14, 16]. Probabilistic programs are almost-surely terminating (AST) if they terminate with probability 1 on all inputs. They are positively AST (PAST) if their expected runtime is finite [5]. We describe AMBER, a fully automated software artifact to prove/disprove (P)AST. Proving (P)AST is a notoriously difficult problem; in fact it is harder than proving traditional program termination [15]. AMBER supports the analysis of a class of polynomial probabilistic programs. Programs in the supported class consist of single loops whose body is a sequence of random assignments with acyclic variable dependencies. Moreover, AMBER’s

This research was supported by the WWTF ICT19-018 grant ProbInG, the ERC Starting Grant SYMCAR 639270, ERC Consolidator Grant ARTIST 101002685, the ERC AdG Grant FRAPPANT 787914, the Austrian FWF project W1255-N23, and the SecInt Doctoral College funded by TU Wien.

$bop \in \{+, -, *, **, /\}$, $cop \in \{>, <\}$
 $dist \in \{\text{uniform, gauss, laplace, bernoulli, binomial, geometric, hypergeometric, exponential, beta, chi-squared, rayleigh}\}$

$$\begin{aligned}
\langle program \rangle &::= \langle i_assign \rangle^* \text{ while } \langle poly \rangle \langle cop \rangle \langle poly \rangle : \langle rv_assign \rangle^+ \langle v_assign \rangle^+ \\
\langle i_assign \rangle &::= \langle var \rangle = \langle const \rangle \mid \langle var \rangle = \langle rv_expr \rangle \quad \langle rv_assign \rangle ::= \langle var \rangle = \langle rv_expr \rangle \\
\langle v_assign \rangle &::= \langle var \rangle = \langle branches \rangle \quad \langle rv_expr \rangle ::= RV(\langle dist \rangle [, \langle const \rangle]^*) \\
\langle branches \rangle &::= \langle poly \rangle \mid \langle poly \rangle @ \langle const \rangle ; \langle branches \rangle \\
\langle poly \rangle &::= p \in C[V] \quad \langle sym \rangle ::= [a-zA-Z][a-zA-Z0-9]^* \quad \langle var \rangle \vee ::= [a-zA-Z][a-zA-Z0-9]^* \\
\langle const \rangle C &::= n \in \mathbb{N} \mid \langle sym \rangle \mid - \langle const \rangle \mid \langle const \rangle \langle bop \rangle \langle const \rangle
\end{aligned}$$

Fig. 1. The AMBER input syntax. $C[V]$ denotes the set of polynomials in V (program variables) with coefficients from C (constants). The power operator is ‘**’.

programming model supports programs parametrized by symbolic constants and drawing from common probability distributions. To automate termination analysis, AMBER automates relaxations of various existing martingale-based proof rules ensuring (non-)(P)AST [8] and combines symbolic computation with asymptotic bounding functions. AMBER certifies (non-)(P)AST without relying on user-provided templates/bounds over termination conditions. Our experiments demonstrate AMBER outperforming the state-of-the-art in the automated termination analysis of probabilistic programs (Sect. 3).

Related Work. The tools MGen [6] and LexRSM [1] use linear programming techniques to certify PAST and AST, respectively. The recent tools Absynth [20], KoAT2 [18] and ecoimp [2] can establish upper bounds on expected costs, therefore also on expected runtimes, and thus certify PAST. While powerful on respective AST/PAST domains, we note that none of the aforementioned tools support both proving and disproving (P)AST. AMBER is the first tool able to prove and disprove (P)AST. Our recent work introduces relaxations of existing proof rules for probabilistic (non-)termination together with automation techniques based on *asymptotic bounding functions* [19]. We utilize these proof rule relaxations in AMBER and extend the technique of asymptotic bounding functions to programs drawing from common probability distributions and including symbolic constants.

Contributions. This tool demonstration paper describes *what* AMBER can do and *how* it can be used for certifying (non-)(P)AST.

- We present AMBER, a fully automatic open-source software artifact¹ for certifying probabilistic (non-)termination (Sect. 2).
- We exhaustively compare AMBER to related tools and report on our experimental findings (Sect. 3).
- We provide a benchmark suite of 50 probabilistic programs as a publicly available repository of probabilistic program examples (Sect. 3).

¹ <https://github.com/probing-lab/amber>.

<pre> 1 x = RV(<i>gauss</i>, 0, 1) 2 y = RV(<i>gauss</i>, 0, 1) 3 while x**2+y**2 < c: 4 s = RV(<i>uniform</i>, 1, 2) 5 t = RV(<i>gauss</i>, 0, 1) 6 x = x+s @1/2; x+2*s 7 y = y+x+t**2 @1/2; y-x-t**2 </pre> <p style="text-align: center;">(a)</p>	<pre> 1 x = x0 2 while x > 0: 3 x = x+c @1/2; x-c </pre> <p style="text-align: center;">(b)</p> <pre> 1 x = x0 2 while x > 0: 3 x = x+c @1/2+e; x-c </pre> <p style="text-align: center;">(c)</p>
--	---

Fig. 2. Two programs supported by AMBER, with symbolic constants $c, x_0, e \in \mathbb{R}^+$; Program 2a is PAST, program 2b is AST but not PAST and program 2c is not AST.

2 Usage and Components

Programming Model. AMBER supports analyzing the probabilistic termination behavior of a class of probabilistic programs involving polynomial arithmetic and drawing from common probability distributions, parameterized by symbolic constants which represent arbitrary real numbers. All symbolic constants are assumed to be positive. Negative constants can be modeled with the explicit use of “-”. The grammar in Fig. 1 defines the input programs to AMBER. Inputs consist of an initialization part and a while-loop, whose guard is a polynomial inequality over program variables. The initialization part is a sequence of assignments either assigning (symbolic) constants or values drawn from probability distributions. Within the loop body, program variables are updated with either (i) a value drawn from a distribution or (ii) one of multiple polynomials over program variables with some probability. Additional to the structure imposed by the grammar in Fig. 1, input programs are required to satisfy the following *structural constraint*: each variable updated in the loop body depends at most linearly on itself and at most polynomially on variables preceding. On a high-level, this constraint enables the use of algebraic recurrence techniques for probabilistic termination analysis [19]. Despite the syntactical restrictions, most existing benchmarks on automated probabilistic termination analysis [19] and dynamic Bayesian networks [3] can be encoded in our programming language. Figure 2 shows three example programs for which AMBER is able to automatically infer the respective termination behavior.

Implementation and Usage. AMBER is implemented in `python3` and relies on the `lark-parser`² package to parse its input programs. Further, AMBER uses the `diofant`³ package as its computer-algebra system. To compute closed-form expressions for statistical moments of monomials over program variables only depending on the loop counter, AMBER uses the tool `Mora` [4]. However, for efficient integration within AMBER, we reimplemented and adapted the `Mora` functionalities exploited by AMBER (`Mora v2`), in particular by employing dynamic

² <https://github.com/lark-parser/lark>.

³ <https://github.com/diofant/diofant>.

programming to avoid redundant computations. Altogether, AMBER consists of ~ 2000 lines of code. Figure 3 shows AMBER’s output when run on the program from Fig. 2a. AMBER can be used through a Docker container [17] or installed locally. Detailed installation and usage instructions are available at <https://github.com/probing-lab/amber>.

Run with Docker. AMBER can be used through a Docker container [17] by running: `$ docker run -ti marcelmoosbrugger/amber`
 AMBER can be run on our `2d_bounded_random_walk` benchmark with:
`$./amber benchmarks/past/2d_bounded_random_walk`

```

AMBER

Martingale expression: -x**2 - 11*x - 115/6
Computing bounds for x**2
Computing bounds for 1
Computing bounds for x

PAST: Yes
AST: Yes

Witness for PAST
-----
Ranking SM          c - x**2 - y**2
SM expression       -x**2 - 11*x - 115/6
SM expression bound -c*20*n**2

Explanation:
Eventually, 'c - x**2 - y**2' is a ranking supermartingale. That's because eventually
the bound of the supermartingale expression is '-c*20*n**2'.

Computation time: 3.6706s
(.venv) → amber git:(master) █
[13] 0:zsh*
```

Fig. 3. The output of AMBER when run on the program from Fig. 2a.

Local Installation. First, clone the repository by running the following command in your terminal: `$ git clone git@github.com:probing-lab/amber.git`
 Change directories to AMBER’s root folder and make sure `python3.8` and the package manager `pip` are installed on your system. All required python packages can be installed by running `$ pip install -r requirements.txt`
 Create an input program (see Sect. 2) and save it in the `benchmarks` folder for example with the file name `my-benchmark`. AMBER can now be run with respect to the input program `benchmarks/my-benchmark` with the following command:
`$ python ./amber.py --benchmarks benchmarks/my-benchmark.`

Components. Figure 4 illustrates AMBER’s main components. AMBER uses four existing probabilistic termination proof rules [6, 9, 12, 16] and their relaxations [19]. Additionally, AMBER extends the algorithms for these relaxations to further support drawing from common probability distributions and symbolic

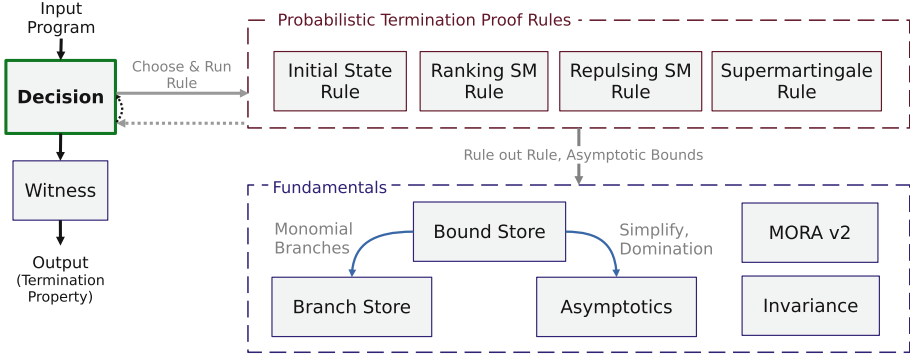


Fig. 4. Main components of AMBER and interactions between them.

constants (cf. Fig. 1). After parsing the input program, AMBER initializes the four proof rule relaxations and determines their applicability [19]. AMBER then executes applicable proof rules consecutively and reports the analysis result containing potential witnesses for (non-)P)AST. The proof rule algorithms require the computation of asymptotic bounding functions which is implemented in the *Bound Store* component.

3 Evaluation

Experimental Setup. AMBER and our benchmarks, are publicly available at <https://github.com/probing-lab/amber>. The output of AMBER is an answer (“Yes”, “No” or “Maybe”) to PAST and AST, together with a potential witness. We took all 39 benchmarks from [19] and extended them by 11 new programs to test AMBER’s capability to handle symbolic constants and drawing from probability distributions. The 11 new benchmarks are constructed from the 39 original programs, by adding noise drawn from common probability distributions and replacing concrete constants with symbolic ones. As such, we conduct experiments using a total of 50 challenging benchmarks, involving polynomial arithmetic, probability distributions and symbolic constants. Further, we compare AMBER not only against **Absynth** and **MGen** (as in [19]), but also evaluate AMBER in comparison to the recent tools **LexRSM** [1], **KoAT2** [18] and **ecoimp** [2]. Note that **MGen** can only certify PAST and **LexRSM** only AST. Moreover, the tools **Absynth**, **KoAT2** and **ecoimp** mainly aim to find upper bounds on expected costs. Tables 1, 2 and 3 summarize our experimental results, with benchmarks separated into *PAST* (Table 1), *AST but not PAST* (Table 2), and *not AST* (Table 3). Benchmarks marked with * are part of our 11 new examples. In every table, ✓ (✗) marks a tool (not) being able to certify the respective termination property. Moreover, **NA** symbolizes that a benchmark is out-of-scope for a tool, for instance, due to not supporting some distributions or polynomial arithmetic. All benchmarks have been run on a machine with a 2.6 GHz Intel i7 (Gen 10)

Table 1. 27 programs which are PAST.

Program	AMBER	Absynth	MGen	LexRSM	KoAT2	ecosimp	Program	AMBER	Absynth	MGen	LexRSM	KoAT2	ecosimp
2d_bounded_random_walk	✓	✗	NA	NA	✗	✗	linear_past_1	✓	✗	✗	✗	✗	✗
biased_random_walk_const	✓	✓	✓	✓	✓	✓	linear_past_2	✓	✗	NA	✗	✗	✗
biased_random_walk_exp	✓	✗	✓	✗	✗	✗	nested_loops	NA	✓	✗	✓	✓	✓
biased_random_walk_poly	✓	✗	✗	NA	✗	✗	polynomial_past_1	✓	✗	NA	NA	✗	✗
binomial_past	✓	✓	✓	✓	✓	✓	polynomial_past_2	✓	✗	NA	NA	✗	✗
complex_past	✓	✗	NA	NA	✗	✗	sequential_loops	NA	✓	✗	✓	✓	✓
consecutive_bernoulli_trails	✓	✓	✓	✓	✓	✓	tortoise_hare_race	✓	✓	✓	✓	✓	✓
coupon_collector_4	✓	✗	✓	✓	✓	✓	dependent_dist*	NA	NA	NA	NA	✗	✓
coupon_collector_5	✓	✗	✓	✓	✓	✓	exp_rw_gauss_noise*	✓	NA	NA	NA	NA	NA
dueling_cowboys	✓	✓	✓	✓	✓	✓	gemoetric_gaussian*	✓	NA	NA	NA	NA	NA
exponential_past_1	✓	NA	NA	NA	✗	NA	race_uniform_noise*	✓	✗	✓	✓	✗	✓
exponential_past_2	✓	NA	NA	NA	✗	NA	symb_2d_rw*	✓	✗	NA	NA	✗	✗
geometric	✓	✓	✓	✓	✓	✓	uniform_rw_walk*	✓	✓	✓	✓	✓	✓
geometric_exp	✗	✗	✗	✗	✗	✗	Total ✓	23	9	11	12	11	13

processor and 32 GB of RAM and finished within a timeout of 50 s, where most experiments terminated within a few seconds.

Experimental Analysis. AMBER successfully certifies 23 out of the 27 PAST benchmarks (Table 1). Although Absynth, KoAT2 and ecosimp can find expected cost upper bounds for large programs [2, 18, 20], they struggle on small programs whose termination is not known a priori. For instance, they struggle when a benchmark probabilistically “chooses” between two polynomials working against each other (one moving the program state away from a termination criterion and one towards it). Our experiments show that AMBER handles such cases successfully. MGen supports the continuous uniform distribution and KoAT2 the geometric distribution whose support is infinite. With these two exceptions, AMBER is the only tool supporting continuous distributions and distributions with infinite support. To the best of our knowledge, AMBER is the first tool certifying PAST supporting both discrete and continuous distributions as well as distributions with finite and infinite support. AMBER successfully certifies 12 benchmarks to be AST which are not PAST (Table 2). Whereas the LexRSM tool can certify non-PAST programs to be AST, such programs need to contain subprograms which are PAST [1]. The well-known example of `symmetric_1D_random_walk`, contained in our benchmarks, does not have a PAST subprogram. Therefore, the LexRSM tool cannot establish AST for it. In contrast, AMBER using the *Supermartingale Rule* can handle these programs. To the best of our knowledge, AMBER is the first tool capable of certifying non-AST for polynomial probabilistic programs involving drawing from distributions and symbolic constants. AMBER is also the first tool automating (non-)AST and (non-)PAST analysis in a unifying manner.

Table 2. 14 programs which are AST and not necessarily PAST.

Program	AMBER	LexRSM
fair_in_limit_random_walk	NA	NA
gambling	✓	✗
symmetric_2d_random_walk	✗	NA
symmetric_random_walk_constant_1	✓	✗
symmetric_random_walk_constant_2	✓	✗
symmetric_random_walk_exp_1	✓	✗
symmetric_random_walk_exp_2	✓	NA
symmetric_random_walk_linear_1	✓	✗
symmetric_random_walk_linear_2	✓	✗
symmetric_random_walk_poly_1	✓	NA
symmetric_random_walk_poly_2	✓	NA
gaussian_rw_walk*	✓	NA
laplacian_noise*	✓	NA
symb_1d_rw*	✓	NA
Total ✓	12	0

Table 3. 9 programs which are not AST.

Program	AMBER
biased_random_walk_nast_1	✓
biased_random_walk_nast_2	✓
biased_random_walk_nast_3	✓
biased_random_walk_nast_4	✓
binomial_nast	✓
polynomial_nast	✗
binomial_nast_noise*	✓
symb_nast_1d_rw*	✓
hypergeo_nast*	✓
Total ✓	8

Experimental Summary. Tables 1, 2 and 3 demonstrate that (i) AMBER outperforms the state-of-the-art in certifying (P)AST, and (ii) AMBER determines (non-)(P)AST for programs with various distributions and symbolic constants.

4 Conclusion

We described AMBER, an open-source tool for analyzing the termination behavior for polynomial probabilistic programs, in a fully automatic way. AMBER computes asymptotic bounding functions and martingale expressions and is the first tool to prove and disprove (P)AST in a unifying manner. AMBER can analyze continuous, discrete, finitely- and infinitely supported distributions in polynomial probabilistic programs parameterized by symbolic constants. Our experimental comparisons give practical evidence that AMBER can (dis)prove (P)AST for a substantially larger class of programs than state-of-the-art tools.

References

1. Agrawal, S., Chatterjee, K., Novotný, P.: Lexicographic ranking supermartingales: an efficient approach to termination of probabilistic programs. Proc. of POPL 2, 1–32 (2017). <https://doi.org/10.1145/3158122>

2. Avanzini, M., Moser, G., Schaper, M.: A modular cost analysis for probabilistic programs. *Proc. of OOPSLA* **4**, 1–30 (2020). <https://doi.org/10.1145/3428240>
3. Bartocci, E., Kovács, L., Stanković, M.: Analysis of bayesian networks via probabilistic loops. In: Pun, V.K.I., Stolz, V., Simao, A. (eds.) *ICTAC 2020*. LNCS, vol. 12545, pp. 221–241. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-64276-1_12
4. Bartocci, E., Kovács, L., Stankovic, M.: Mora - automatic generation of moment-based invariants. In: *Proceedings of TACAS (2020)*. <https://doi.org/10.1007/978-3-030-45190-5>
5. Bournez, O., Garnier, F.: Proving positive almost-sure termination. In: Giesl, J. (ed.) *RTA 2005*. LNCS, vol. 3467, pp. 323–337. Springer, Heidelberg (2005). https://doi.org/10.1007/978-3-540-32033-3_24
6. Chakarov, A., Sankaranarayanan, S.: Probabilistic program analysis with martingales. In: Sharygina, N., Veith, H. (eds.) *CAV 2013*. LNCS, vol. 8044, pp. 511–526. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39799-8_34
7. Chatterjee, K., Fu, H., Goharshady, A.K.: Termination analysis of probabilistic programs through positivstellensatz's. In: Chaudhuri, S., Farzan, A. (eds.) *CAV 2016*. LNCS, vol. 9779, pp. 3–22. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41528-4_1
8. Chatterjee, K., Fu, H., Novotný, P.: Termination Analysis of Probabilistic Programs with Martingales. *Foundations of Probabilistic Programming*, pp. 221–258 (2020). <https://doi.org/10.1017/9781108770750.008>
9. Chatterjee, K., Novotný, P., Zikelic, D.: Stochastic invariants for probabilistic termination. In: *Proceedings of POPL (2017)*. <https://doi.org/10.1145/3009837.3009873>
10. Chen, J., He, F.: Proving almost-sure termination by omega-regular decomposition. In: *Proceedings of PLDI (2020)*. <https://doi.org/10.1145/3385412.3386002>
11. Esparza, J., Gaiser, A., Kiefer, S.: Proving termination of probabilistic programs using patterns. In: Madhusudan, P., Seshia, S.A. (eds.) *CAV 2012*. LNCS, vol. 7358, pp. 123–138. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31424-7_14
12. Ferrer Fioriti, L.L.M., Hermanns, H.: Probabilistic termination: soundness, completeness, and compositionality. In: *Proceedings of POPL (2015)*. <https://doi.org/10.1145/2676726.2677001>
13. Ghahramani, Z.: Probabilistic machine learning and artificial intelligence. *Nature* **521**, 452–459 (2015). <https://doi.org/10.1038/nature14541>
14. Huang, M., Fu, H., Chatterjee, K.: New approaches for almost-sure termination of probabilistic programs. In: Ryu, S. (ed.) *APLAS 2018*. LNCS, vol. 11275, pp. 181–201. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-02768-1_11
15. Kaminski, B.L., Katoen, J.P.: On the hardness of almost-sure termination. In: Italiano, G.F., Pighizzini, G., Sannella, D.T. (eds.) *MFCFS 2015*. LNCS, vol. 9234, pp. 307–318. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-48057-1_24
16. McIver, A., Morgan, C., Kaminski, B.L., Katoen, J.P.: A new proof rule for almost-sure termination. *Proc. of ACM Program. Lang.* (2018). <https://doi.org/10.1145/3158121>
17. Merkel, D.: Docker: lightweight linux containers for consistent development and deployment. *Linux J.* **239**, 2 (2014)
18. Meyer, F., Hark, M., Giesl, J.: Inferring expected runtimes of probabilistic integer programs using expected sizes. In: *TACAS 2021*. LNCS, vol. 12651, pp. 250–269. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72016-2_14

19. Moosbrugger, M., Bartocci, E., Katoen, J.P., Kovács, L.: Automated termination analysis of polynomial probabilistic programs. In: ESOP 2021. LNCS, vol. 12648, pp. 491–518. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72019-3_18
20. Ngo, V.C., Carbonneaux, Q., Hoffmann, J.: Bounded expectations: resource analysis for probabilistic programs. In: Proceedings of PLDI (2018). <https://doi.org/10.1145/3192366.3192394>