Probabilistic Program Analysis with Martingales

Aleksandar Chakarov and Sriram Sankaranarayanan
CAV'13

Report date: March 17, 2021

Introduction

- Probablistic programs: Standard imperative program + *random value generators*
  - Branching.
  - Assignment.
- Problem: Invariant synthesis and termination checking in probablistic settings.

Contributions

- Extend *quantitative invariants*, using Azuma-Hoeffding theorem to generate probabilistic assertions.
- Define *super martigale ranking functions* (SMRF) to prove almost sure termination ($\texttt{Pr}(terminates) = 1$) of probablistic programs.
- A constraint-based algorithm for supermartingale expression generation.

Restrictions

- Only applies to stochastic programs. (Not a demonic program with non-determinism)
- Restricted on linear expressions and systems.
- A.s. termination proving is sound but incomplete.

Motivating Examples

Example (Hare and Tortoise)

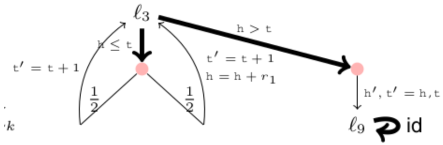
```
real h, t;
// h is hare and t is tortoise
h = 0; t = 30;
while ( h <= t ){
  if (flip (0.5) )
        h = h + unifRand(0,10);
  t = t +1;
} // almost sure terminate?
```

- Worst case: non-terminating.
- The expression $t - h$ decrease by 1.5 in expectation each iteration.

Probablistic Transition System

Definition 1 (Probabilistic Transition System). *A Probabilistic Transition System (PTS) Π is defined by a tuple $\langle X, R, L, \mathcal{T}, \ell_0, \boldsymbol{x}_0 \rangle$ such that*

1. *X, R represent the program and random variables, respectively.*
2. *L represents a finite set of* locations. *$\ell_0 \in L$ represents the* initial location, *and $\boldsymbol{x}_0$ represents the* initial values *for the program variables.*
3. *$\mathcal{T} = \{\tau_1, \ldots, \tau_p\}$ represents a finite set of* transitions. *Each transition $\tau_j \in \mathcal{T}$ is a tuple $\langle \ell, \varphi, f_1, \ldots, f_k \rangle$ consisting of (see Fig 2):*
   (a) *Source location $\ell \in L$, and guard assertion φ over X,*
   (b) Forks *$\{f_1, \ldots, f_k\}$, where each fork $f_j : (p_j, F_j, m_j)$ is defined by a fork probability $p_j \in (0, 1]$, a (continuous) update function $F_j(X, R)$ and a destination $m_j \in L$. The sum of the fork probabilities is $\sum_{j=1}^{k} p_j = 1$.*



No Demonic: mutually exclusive and mutually exhaustive.

State and Post-Distribution

A *state* of PTS is a tuple $(l, \mathbf{x})$ where $l \in L$ and $\mathbf{x}$ is a valuation of X.

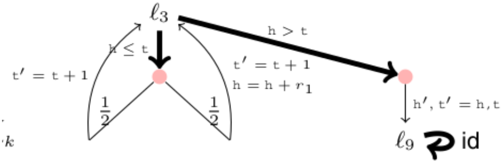
Given a transition $\tau = \langle l, \phi, f_1, \ldots, f_k \rangle$, if $\mathbf{x} \models \phi$ then the result of executing τ is a *probability distribution* over post states, obtained by:

1. Choose fork f_j with probability p_j, and a vector of random variables $\mathbf{r} : (r_1, \ldots, r_m)$ is drawn according to the joint distribution.
2. Update the states by computing the function $\mathbf{x}' = F_j(\mathbf{x}, \mathbf{r})$ and update l to m_j.

$$\textsc{Post-Distrib}(s, \tau), \textsc{Post-Distrib}(s)$$

Operationally, PTS is a Markov chain.

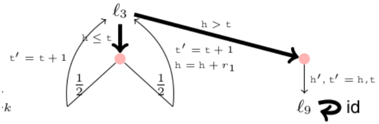
State and Post-Distribution

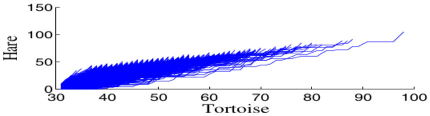


Sampled Executions

Definition 2 (Sample Executions). *Let Π be a transition system. A sample execution σ of Π is a countably infinite sequence of states $\sigma : (\ell_0, \boldsymbol{x}_0) \xrightarrow{\tau_1} (\ell_1, \boldsymbol{x}_1) \xrightarrow{\tau_2} \cdots \xrightarrow{\tau_n} (\ell_n, \boldsymbol{x}_n) \cdots$, such that (a) $(\ell_0, \boldsymbol{x}_0)$ is the unique initial state. (b) The state $s_j : (\ell_j, \boldsymbol{x}_j)$ for $j \geq 0$ satisfies the guard for the transition τ_{j+1}. Note that by the no demonic restriction, τ_{j+1} is uniquely defined for each s_j. (c) Each state $s_{j+1} : (\ell_{j+1}, \boldsymbol{x}_{j+1})$ is a sample from* POST-DISTRIB(s_j).

Example





Almost Sure Termination

Definition (Termination)

Let Π be a PTS with a special *final location* l_F. l_F has only one outgoing transition `id`. A sampled execution σ of Π *terminates* if it eventually reaches a state $(l_F, \mathbf{x})$.

Probability of terminating paths:

- For a finite syntactic path $\pi : l_0 \xrightarrow{\tau_1} l_1 \xrightarrow{\tau_2} l_2 \ldots l_F$, there is a well-defined probability $\mu(\pi) \in [0, 1]$ that characterizes the probaility of the path going through the locations.
- The overall probability of termination can be obtained as the sum of probability of all such finite syntactic paths.

Almost Sure Termination

The main idea to show μ on the infinite space is well-defined:
Let $\Omega = \Pi_{j=1}^{\infty}\Omega_j$ and $\mathcal{F} = \Pi_{j=1}^{\infty}\mathcal{F}_j$.

- At each location l_i, there is a probability space $(\Omega_i, \mathcal{F}_i, \mu_i)$.
- For a given n, a measurable cylinder can be constructed by $B_n = \{\omega \in \Omega \mid (\omega_1, \ldots, \omega_n) \in B^n\}$, where $B^n = \Pi_{j=1}^{n} A_j, A_j \in \mathcal{F}_j$. Assume the measure is P_n.
- Theorem of cylinder construction, a probability measure space $(\Omega, \mathcal{F}, P)$ such that $P\{\omega \in \Omega \mid (\omega_1, \ldots, \omega_n) \in B^n\} = P_n(B^n)$.

Almost Sure Termination

Definition (a.s. Termination)

A PTS is said to be almost sure terminating iff the sum of probabilities of all terminating syntactic paths is 1.

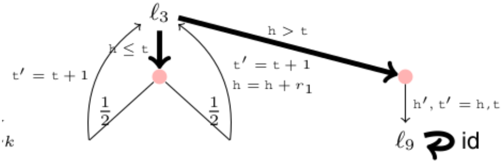
Pre-Expectation

Definition (Pre-expectation of an expression)

Let state $s : (l, \mathbf{x})$ be a state and τ be the enabled transition on s. The pre-expectation $\mathbb{E}(\mathbf{e}'|s)$ is defined as the conditional expected value of $\mathbf{e}'$ over $\textsc{Post-Distrib}(s)$ as

$$\mathbb{E}_\tau(\mathbf{e}'|s) = \Sigma_{j=1}^{k} p_j \mathbb{E}_R(\mathtt{pre}(\mathbf{e}', F_j))$$

Pre-Expectation



Example

$$\mathbb{E}(5t - 2h|(l_3, h, t))$$

Martingale and Supermartingale Expression

A discrete-time stochastic process $\{M_n\}$ is a countable sequence of random variables $M_0, M_1, M_2, \ldots$ where M_n is distributed based on the samples drawn from $M_0, \ldots, M_{n-1}$. By convention, M_n denotes the random variable and m_n its sample.

Definition 4 (Martingales and Super Martingales). *A process* $\{M_n\}$ *is a* martingale *iff for each* $n > 0$, $\mathbb{E}(M_n | m_{n-1}, \ldots, m_0) = m_{n-1}$. *In other words, at each step the expected value at the next step is equal to the current value. Likewise* $\{M_n\}$ *is a* super-martingale *iff for each* $n > 0$, $\mathbb{E}(M_n | m_{n-1}, \ldots, m_0) \leq m_{n-1}$.

Adapting the original definition to PTS:

Definition 5 (Martingale Expressions). *An expression* $\mathbf{e}[X]$ *over program variables* X *is a* martingale *for the PTS* Π *iff for every transition* $\tau : (\ell, \varphi, f_1, \ldots, f_k)$ *in* Π *and for every state* $s : (\ell, \boldsymbol{x})$ *for which* τ *is enabled, the pre-expectation of* $\mathbf{e}$ *equals its current state value:* $\forall\ \boldsymbol{x}.\ \varphi[\boldsymbol{x}] \ \Rightarrow\ \mathbb{E}_\tau(\mathbf{e}' | \ell, \boldsymbol{x}) = \mathbf{e}$. *Likewise, an expression is a* super-martingale *iff for each transition* τ, $\forall\ \boldsymbol{x}.\ \varphi[\boldsymbol{x}] \ \Rightarrow\ \mathbb{E}_\tau(\mathbf{e}' | \ell, \boldsymbol{x}) \leq \mathbf{e}$.

Almost-Sure Termination

Definition (Ranking Super Martingale(RSM))

A supermartingale $\{M_n\}$ is ranking iff

- There exists $\epsilon > 0$ s.t. for all sampled paths, $\mathbb{E}(M_{n+1}|m_n) \leq m_n - \epsilon$.
- For all $n \geq 0$, $M_n \geq -K$ for some $K > 0$. (Equiv. Def. For all $T(\omega) > j$, $M_j \geq 0$).

How a.s. Termination is Proved?

Ranking function: will finally become negative.
Ranking supermartingale: will almost surely become negative.

How a.s. Termination is Proved?

Ranking function: will finally become negative.
Ranking supermartingale: will almost surely become negative.

Theorem (Main result)

A ranking supermartingale with a positive initial value will a.s. become negative.

How a.s. Termination is Proved?

Ranking function: will finally become negative.
Ranking supermartingale: will almost surely become negative.

Theorem (Main result)

A ranking supermartingale with a positive initial value will a.s. become negative.

Proof.

Stopping time: $t = \inf_{n \geq 0} m_n \leq 0$. Let the r.v. be T. M_n^T.
$Y_n = M_n^T + \epsilon \min(n, T)$. $n < t, n \geq t$. □

How a.s. Termination is Proved?

Ranking function: will finally become negative.
Ranking supermartingale: will almost surely become negative.

Theorem (Main result)

A ranking supermartingale with a positive initial value will a.s. become negative.

Proof.
Stopping time: $t = \inf_{n \geq 0} m_n \leq 0$. Let the r.v. be T. M_n^T.
$Y_n = M_n^T + \epsilon \min(n, T)$. $n < t, n \geq t$. □

Lemma
$\{Y_n\}$ *is a supermartingale and* $Y_n \geq -K$

Proof.
Hint: discuss $n + 1 \geq t$ and $n + 1 < t$ □

How a.s. Termination is Proved?

Theorem (Super Martingale Convergence Theorem)

A lower-bounded supermartingale converges almost surely.

Lemma

For any convergent sample path $y_0, \ldots, y_n, \ldots$, the corresponding $\{M_n\}$ sample path $m_0, \ldots, m_n, \ldots$ eventually becomes negative eventually.

Proof.

Hint: assume $\{M_n\}$ sample path has stopping time $t = \infty$. By the definition of convergence and the definition of t. □

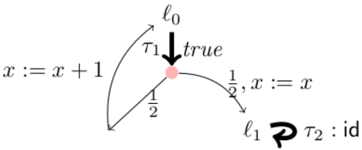
Expression Map

Definition 7 (Martingales and Super Martingale Expression Maps). *An expression map η is a martingale for a PTS Π iff for every transition $\tau : (\ell, \varphi, f_1, \ldots, f_k)$, we have $\forall\, \boldsymbol{x}.\ \varphi[\boldsymbol{x}] \ \Rightarrow\ \mathbb{E}_\tau(\eta'|\ell, \boldsymbol{x}) = \eta(\ell)[\boldsymbol{x}]$.*

Likewise, the map is a super-martingale iff for every transition τ, $\forall\, \boldsymbol{x}.\ \varphi[\boldsymbol{x}] \ \Rightarrow\ \mathbb{E}_\tau(\eta'|\ell, \boldsymbol{x}) \leq \eta(\ell)[\boldsymbol{x}]$.

Example

```
int x := 0;
while (flip (0.5))
    x ++;
// end
```



Super Martingale Ranking Function

Definition 9 (Super Martingale Ranking Function). *A super martingale ranking function (SMRF) η is a s.m. expression map that satisfies the following:*

- $\eta(\ell) \geq 0$ *for all* $\ell \neq \ell_F$, *and* $\eta(\ell_F) \in [-K, 0)$ *for some lower bound* K.
- *There exists a constant* $\epsilon > 0$ *s.t. for each transition* τ *(other than the self-loop* id *around* ℓ_F*) with guard* φ, $(\forall\, \boldsymbol{x})\ \varphi[\boldsymbol{x}] \ \Rightarrow\ \mathbb{E}_\tau(\eta'|\ell, \boldsymbol{x}) \leq \eta(\ell)[\boldsymbol{x}] - \epsilon$.

How SMRF works?

How to Synthesize SMRF

Affine linear templates.

$$(\forall\, \boldsymbol{x})\ (\varphi[\boldsymbol{x}])\ \Rightarrow\ \underbrace{\mathbb{E}_\tau(\eta'|\ell, \boldsymbol{x})}_{\text{template expression}}\ \leq\ \underbrace{\eta(\ell)[\boldsymbol{x}]}_{\text{template expression}}$$

By Farkas Lemma convert the system to the system of parameters. [CAV'06]

Example

Template: $f_{l_3}(h,t) = c_1 h + c_2 t$

- $\tau = (l_3, (h \leq t), f_1, f_2)$
- $f_1 : (\frac{1}{2}, (\lambda(h,t).h, t+1), l_3)$.
- $f_2 : (\frac{1}{2}, (\lambda(h,t).h + r_1, t+1))$