Analyzing Probabilistic Programs with Dynamic Logic

Tong Cheng *
School of Computer Science and Technology
Huazhong University of Science and Technology
Wuhan, Hubei 430074, China
Email: tcheng@hust.edu.cn

Abstract—We present a unified framework for the analysis of probabilistic programs based on Dynamic Logic in which both iterative and recursive programs can be analyzed. We augment the traditional rules of probabilistic Dynamic Logic with rules for assignment based on schematic Kleene algebra, as well as proof rules for recursive programs with call-by-value parameters. We show soundness of our system with respect to the standard Markov kernel semantics. We give an example of their use in the analysis of the Coupon Collector's Problem.

1. Introduction

The verification of probabilistic programs has been an active topic of research for almost four decades [1]. Recent applications in artificial intelligence and machine learning have generated renewed interest, but much recent research has ignored the early classical results in the field. In this note, we show how these classical systems can be easily adapted to provide tools that match the deductive power of recently proposed systems. In particular, we show that there is no need for specialized logical syntax for reasoning about expected behavior as advocated in [2], [3].

Our approach is based on Probabilistic Propositional Dynamic Logic (PPDL), a system for reasoning about partial correctness of probabilistic programs and deriving expected values of measurable function and other moments. We provide several enhancements to streamline the reasoning process; in particular, we illustrate the use of a new construct that treats an arithmetic expression as a proposition, thereby allowing expected values to be calculated by simple substitutions. In particular, this provides a substantially simpler method for reasoning about nested loops. We illustrate the construct by giving a relatively straightforward analysis of the Coupon Collector’s Problem, a problem that has been recognized as notoriously difficult to analyze formally [2].

We also augment the logic with syntactic constructs for representing general recursion, not just iteration or tail-recursion. We provide proof rules for recursion with call-by-value parameters and prove their soundness. This has been done by [4], but our point here is to show how it is a simple extension of much more classical work.

For parameter passing, we provide a mechanism to define and reason about local variables in the form of a let-in construct. We give a semantics based on Markov kernels [5] and give sound proof rules for this construct.

1.1. Related Work

Recent work on deductive systems for probabilistic programs has focused on deductive systems for iterative probabilistic programs [2] and recursive probabilistic programs [3]. In that work, calculi for iterative and recursive programs are defined separately, and specialized syntax for reasoning about expectations is used. Here we provide a unified system for termination, correctness and expected time analysis. In our approach, there is no need for specialized syntax for expectations; mechanisms already available from much earlier work not only subsume these results but are more flexible.

McIver and Morgan [6], [7] handle probabilistic programs with nondeterminism, but as nondeterminism presents certain complications, here we focus on probabilistic programs without nondeterminism.

Our semantics is based on the denotational semantics of probabilistic programs as presented in [1].

2. Syntax and Semantics of Probabilistic DL

In PPDL, terms are of two types: programs and measurable functions. Here we use assignments \( x \leftarrow e \) \((x \in \text{Var}, \ e \in \text{expr})\) as primitive programs and expressions \( e \) \((e \in \text{expr})\) as primitive measurable functions.

We assume a set of variables \( \text{Var} \), and an inductively defined set of expressions \( \text{expr} \). Elements in \( \text{expr} \) can be arithmetic expressions or Boolean expressions. In PPDL, Boolean expressions are called propositions, and usually used in tests. Here we generalize the concept of propositions to all measurable functions, allowing tests to hold with respect to some probability. This allows simpler evaluation for expected values, as shown in Section 4. We are changing our perspectives: propositions not only can represent some truth statement over a state, but also can represent a value over a state, due to the arithmetic nature of the probabilistic version of Dynamic Logic.
The grammar for propositions and programs are as follows.

\[
\begin{align*}
\text{prop} & ::= e \quad (e \in \text{expr}) \quad \text{(expression)} \\
& | k \cdot f \quad (k \in \mathbb{R}) \quad \text{(scalar)} \\
& | f_1 + f_2 \quad \text{(sum)} \\
& | f_1 \cdot f_2 \quad \text{(multiplication)} \\
& | \langle p \rangle f \quad \text{(eventuality)}
\end{align*}
\]

\[
\begin{align*}
\text{prog} & ::= x \leftarrow e \quad (e \in \text{expr}) \quad \text{(assignment)} \\
& | k \cdot p \quad (k \in \mathbb{R}) \quad \text{(scalar)} \\
& | p_1 + p_2 \quad \text{(sum)} \\
& | p_1; p_2 \quad \text{(composition)} \\
& | f? \quad \text{(test)} \\
& | p^n \quad \text{(power)} \\
& | p^* \quad \text{(Kleene star)}
\end{align*}
\]

Denote the set of possible states (variable valuation) as

\[
S = \{x_1, x_2, \ldots, x_n \mid x_i \in \mathbb{N} \text{ as the value of } i\text{-th variable}\}
\]

A state is a vector of values of variables [5]. Let \( \Sigma \) be a \( \sigma \)-algebra of measurable sets on \( S \).

Expressions are interpreted as measurable functions \( S \to [0, \infty] \), denoted by \( \mathcal{A}[e] \). Those are naturally given by the inductive definition on the structure of \( \text{expr} \).

Propositions are interpreted as measurable functions \( S \to [0, \infty] \):

\[
\begin{align*}
\mathcal{B}[e] &= \mathcal{A}[e] \quad (1) \\
\mathcal{B}[k \cdot f] &= k \cdot \mathcal{B}[f] \quad (2) \\
\mathcal{B}[f_1 + f_2] &= \mathcal{B}[f_1] + \mathcal{B}[f_2] \quad (3) \\
\mathcal{B}[f_1 \cdot f_2] &= \mathcal{B}[f_1] \cdot \mathcal{B}[f_2] \quad (4) \\
\mathcal{B}[\langle p \rangle f] &= \lambda s. \int_{t \in S} \mathcal{B}[f](t) \cdot \mathcal{B}[\langle p \rangle f](s, dt) \quad (5)
\end{align*}
\]

Programs are interpreted as Markov kernels \( S \to \Sigma \to [0, \infty] \):

\[
\begin{align*}
\mathcal{C}[x \leftarrow e] &= \lambda s. \Delta_s[x/\mathcal{A}[e](s)](A) \quad (6) \\
\mathcal{C}[k \cdot p] &= k \cdot \mathcal{C}[p] \quad (7) \\
\mathcal{C}[p_1 + p_2] &= \mathcal{C}[p_1] + \mathcal{C}[p_2] \quad (8) \\
\mathcal{C}[p_1; p_2] &= \lambda s. \int_{t \in S} \mathcal{C}[p_2](t, A) \cdot \mathcal{C}[p_1](s, dt) \quad (9) \\
\mathcal{C}[f?] &= \lambda s. A. \Delta_s(A) \cdot \mathcal{B}[f](s) \quad (10) \\
\mathcal{C}[p^n] &= \begin{cases} 
1 & \text{if } n = 0 \\
\mathcal{C}[p^{n-1}; p] & \text{if } n > 0
\end{cases} \quad (11) \\
\mathcal{C}[p^*] &= \sum_{n=0}^{\infty} \mathcal{C}[p^n] \quad (12)
\end{align*}
\]

where \( \Delta_s \) is the Dirac measure centered on state \( s \).

### 3. Proof Rules for First-order Terms

We present some proof rules for assignment in first-order terms. These rules are based on rules for Schematic Kleene algebra with tests (SKAT) as presented in [8].

\[
\begin{align*}
x \leftarrow e; y \leftarrow a &= y \leftarrow a[x/e]; x \leftarrow e, \quad (13) \\
x \leftarrow e; y \leftarrow a &= x \leftarrow e; y \leftarrow a[x/e], \quad (14) \\
x \leftarrow e; x \leftarrow a &= x \leftarrow a[x/e], \quad (15) \\
x \leftarrow e; f? &= f[x/e]?, x \leftarrow e, \quad (16) \\
\langle x \leftarrow e \rangle f &= f, \quad (17)
\end{align*}
\]

where equation (13) holds when \( x \neq y, y \in \text{FV}(e) \), equation (14) holds when \( x \neq y, x \notin \text{FV}(e) \), equation (17) holds when \( x \notin \text{FV}(f) \).

**Theorem 1.** Equations (13)–(17) are sound with respect to the Markov kernel semantics.

### 4. Well-Structured Iterative Programs

Our extension to admit arithmetic expressions as propositions allows a simpler representation for expected value evaluation, which is often used in analysis based on invariants. In this section, we will introduce this representation and present an example using invariant analysis.

In [5], a PPDL proposition used to compute the expected value of variable \( c \) in program \( p \) was proposed as follows:

\[
\langle (c \leftarrow c + 1)^n; c \geq 1 \rangle 1. \quad (18)
\]

This proposition is the stair function on \( c \)-axis, whose integration with program \( p \) will be the expectation of \( c \), if \( \mathcal{C}[p](s) \) is always a probability measure (which is always true for well-structured programs).

With our extension, we can directly represent the identity function on \( c \)-axis (as single letter \( c \)), whose integration with program \( p \) will also be the expectation of \( c \):

\[
\langle p \rangle c = \int_{s \in S} \mathcal{B}[c](t) \cdot \mathcal{C}[p](s, dt)
\]

Furthermore, we can use rules introduced in Section 3 to directly do transformations. For instance, we can transform \( (c \leftarrow c + 1) c \) to \( c + 1 \).

\[
\begin{align*}
(c \leftarrow c + 1)c &= (c \leftarrow c + 1) \cdot 1 \\
&= (c \leftarrow c + 1)\langle c \rangle 1 \\
&= (c \leftarrow c + 1; c) 1 \quad \text{by (16)} \\
&= (c + 1?; c \leftarrow c + 1) 1 \\
&= (c + 1)\langle c \leftarrow c + 1 \rangle 1 \\
&= (c + 1) \cdot (c \leftarrow c + 1) 1 \quad \text{by (17)} \\
&= (c + 1) \cdot 1 \\
&= c + 1.
\end{align*}
\]

PPDL has an induction rule

\[
f + \langle p \rangle g \leq g \rightarrow \langle p^* \rangle f \leq g
\]

that can be used to obtain the upper bound of expected value of a variable in a loop.
For example, for a loop with an uncommon form \((pbu)^*pbu\), we can instantiate an induction rule
\[
\langle pbu \rangle c + \langle pbu \rangle g \leq g \quad \rightarrow \quad \langle (pbu)^* \rangle \langle pbu \rangle c \leq g
\]
to prove \(g\) as the upper bound of expected value of variable \(c\).

An important instantiation of induction rule is the while rule
\[
\langle \bar{b} \rangle f + \langle bp \rangle g \leq g \quad \rightarrow \quad \langle (bp)^* \rangle \langle \bar{b} \rangle f \leq g,
\]
i.e.
\[
\langle \bar{b} \rangle f + \langle bp \rangle g \leq g \quad \rightarrow \quad \langle \text{while } b \text{ do } p \rangle f \leq g.
\]

In many circumstances, an upper bound is good enough. For example, when we analyze the expected running time of an algorithm and are going to represent the result in terms of Big-O notation, we just need an upper bound for the expected value. Nevertheless, we can get an lower bound by a strengthened version of induction rule
\[
g_0 = 0, \quad f + \langle p \rangle g_n \geq g_{n+1} \quad \rightarrow \quad \langle p^* \rangle f \geq \sup_n g_n
\],
\[
g_0 = 0, \quad f + \langle p \rangle g_n \leq g_{n+1} \quad \rightarrow \quad \langle p^* \rangle f \leq \sup_n g_n
\]

We can use PPDL to analyze the Coupon Collector’s Problem (Figure 1). Suppose we are collecting coupons of \(n\) sorts (assume \(n > 0\)). One coupon of a random sort can be acquired in each round (equally likely for all sorts). Our goal is to evaluate the expected number of coupons we need to get coupons of all sorts, which is also the expected running time for the collecting procedure.

Let
\[
\begin{align*}
\text{pin} & = \langle b_y \rangle \{ p_i; p_c \} \\
\text{pout} & = \langle b_y \rangle \{ p_i; p; p_c \} ; b_y
\end{align*}
\]

To analyze the inner loop, we use PPDL induction rule
\[
(f + \langle p \rangle g \leq g \quad \rightarrow \quad \langle p^* \rangle f \leq g)
\], which can be instantiated as:
\[
\langle b_y \rangle c + \langle b_y \rangle \{ p_i; p_c \} \leq I \quad \rightarrow \quad \langle (b_y \rangle \{ p_i; p_c \}^* \rangle \langle b_y \rangle c \leq I.
\]

We propose an invariant
\[
I = c + \langle b_y \rangle \frac{n}{n - k}, \quad \text{where } k = \sum_{i=0}^{n-1} \langle y[i] \rangle \neq 0
\]

which can also be verified in Section 7.1 to satisfy the left-hand condition in (21). Then we have
\[
\langle pin \rangle c \leq c + \langle b_y \rangle \frac{n}{n - k}.
\]

For the outer loop, we instantiate the induction rule as:
\[
\langle b_z \rangle c + \langle b_z \rangle \langle pin \rangle ; p_y ; p_x \langle b_z \rangle \leq J \quad \rightarrow \quad \langle (b_z \rangle \langle pin \rangle ; p_y ; p_x \rangle^* \rangle \langle b_z \rangle c \leq J.
\]

For \(x > 0\), we propose an invariant
\[
J = c + \langle b_y \rangle \frac{n}{n - k} + \sum_{j=n-k-x+1}^{n-1} \frac{n}{j},
\]

which can also be verified in Section 7.2. Then we have
\[
\langle p_{out} \rangle c \leq J.
\]

For the program overall,
\[
\langle c \leftarrow 0; y \leftarrow [0 \ldots 0]; p_i; p_c; x \leftarrow n \rangle \langle p_{out} \rangle c \leq J[\frac{c}{n}; \frac{c}{c+1}; i/\text{rand}(n), y[0]/0, \ldots, y[n-1]/0, c/0]
\]
\[
= 1 + \sum_{j=n-k-x+1}^{n-1} \frac{n}{j} \sum_{j=1}^{n} \frac{n}{j} = n \cdot H(n) = O(n \log n).
\]

5. Recursion

Our language can be easily extended to support recursion. We define a recursive procedure by writing down a program operator \(T : \text{prog} \rightarrow \text{prog}\) and use a new program construct \(\text{fix } T(\cdot)\) for the least fixpoint of the operator, i.e. the recursion procedure.

We define the semantics of the least fixpoint as follows:
\[
C[\text{fix } T(q)] = \sup_n C[T^n(0?)].
\]

For example, the loop while (true) do {skip} can be written as the least fixpoint of a program operator:
\[
\begin{align*}
p &= \text{fix } T(q), \\
T(q) &= \langle 0? + 1; 1?; q \rangle
\end{align*}
\]

Since \(T^n(0?) = 0?\) for all \(n\), we have \(p = \text{fix } T(q) = 0?\). We also present the induction rules for recursions:
\[
\langle p \rangle f \leq g \quad \rightarrow \quad \langle \text{fix } T(p) \rangle f \leq g
\]

(26)
\[ g_0 = 0, \quad (p)f \leq g_n \rightarrow (T(p))f \leq g_{n+1} \quad (27) \]
\[ g_0 = 0, \quad (p)f \leq g_n \rightarrow (T(p))f \geq g_{n+1} \quad (28) \]

where \( p = T^i(0?) \) for some \( i \) in each rule.

We can illustrate the usage of those rules with the following program:
\[
\begin{align*}
p &= fix T(q), \\
T(q) &= \frac{1}{2} \cdot \text{skip} + \frac{1}{2} (q; q),
\end{align*}
\]
whose upper bound for probability of termination can be calculated as follows:

Assume \( \langle q \rangle 1 = \phi = \frac{\sqrt{5} - 1}{2} \), we have
\[
\langle T(q) \rangle 1 = \left\langle \frac{1}{2} \cdot \text{skip} + \frac{1}{2} (q; q) \right\rangle 1
\]
\[
= \frac{1}{2} + \frac{1}{2} (q) \langle q \rangle 1
\]
\[
= \frac{1}{2} + \frac{1}{2} (q) \langle q \rangle \phi
\]
\[
= \frac{1}{2} + \frac{1}{2} \phi \langle q \rangle 1
\]
\[
= \frac{1}{2} + \frac{1}{2} \phi \langle q \rangle \phi
\]
\[
= \frac{1}{2} + \frac{1}{2} \phi \cdot \phi
\]
\[
= \phi.
\]

Then we can conclude that
\[
\langle p \rangle 1 = \langle fix T \rangle 1 \leq \phi.
\]

6. Local Variables

In this section we show how to add local variables to the language by providing semantics for a let-in construct. We add a new program term to the language:

\[
\text{let } x \leftarrow e \text{ in } p \text{ end}
\]
whose semantics is defined as
\[
C[\text{let } x \leftarrow e \text{ in } p \text{ end}] (s, A) = \int_{t \in S} \delta_t[x/A[x]][s][A][f][s][A][A][f][s][dt).
\]

Then we present some proof rules for let-expressions:
\[
\begin{align*}
\langle \text{let } x \leftarrow e \text{ in } p \text{ end} \rangle f &= \langle x \leftarrow e; p \rangle f, \quad (29) \\
\langle \text{let } x \leftarrow e \text{ in } p \text{ end} \rangle f &= f. \quad (30)
\end{align*}
\]

where equation (29) holds for \( x \notin FV(f) \), equation (30) holds for \( FV(f) = \{x\} \).

Theorem 2. Equations (29) and (30) are sound with respect to the Markov kernel semantics.

7. The Coupon Collector’s Problem

In this section we provide a complete proof of the expected running time of the Coupon Collector’s Problem.

7.1. Inner Loop Verification

\[
\begin{align*}
\langle \bar{b}_y? \rangle c + \langle b_y?; p_i; p_e \rangle I \\
= \langle \bar{b}_y? \rangle c + \langle b_y?; p_i; p_e \rangle c + \langle b_y?; p_i; p_e \rangle \langle b_y? \rangle \frac{n}{n-k} \\
= \langle \bar{b}_y? \rangle c + \langle b_y? \rangle (c + 1) + \langle b_y?; p_i; p_e \rangle \frac{n}{n-k} \\
= c + \langle b_y? \rangle \left(1 + \langle p_i; b_y? \rangle \frac{n}{n-k}\right) \\
= c + \langle b_y? \rangle \left(1 + \frac{1}{n} \sum_{j=0}^{n-1} (i \leftarrow j; b_y?) \frac{n}{n-k}\right) \\
= c + \langle b_y? \rangle \left(1 + \frac{1}{n} \sum_{j=0}^{n-1} (y[j] \neq 0?) \frac{n}{n-k}\right) \\
= c + \langle b_y? \rangle \left(1 + \frac{1}{n} \cdot \frac{n}{n-k} \sum_{j=0}^{n-1} (y[j] \neq 0?)\right) \\
= c + \langle b_y? \rangle \left(\frac{n}{n-k}\right) = I.
\end{align*}
\]

7.2. Outer Loop Invariant

\[
\begin{align*}
\langle \bar{b}_z? \rangle c + \langle b_z?; p_i; p_g; p_e \rangle J \\
= \langle p_i; p_g; p_e \rangle J \quad \text{(assumed initial state where } x > 0) \\
= \langle p_i; p_g; p_e \rangle c + \langle p_i; p_g; p_e \rangle \langle b_y? \rangle \frac{n}{n-k} \\
+ \langle p_i; p_g; p_e \rangle \sum_{j=n-k-2+1}^{n-1} \frac{n}{j} \\
= \langle p_i; p_g \rangle c + \langle p_i; p_g \rangle \langle b_y? \rangle \frac{n}{n-k} \\
+ \langle p_i; p_g \rangle \sum_{j=n-k-2+1}^{n-1} \frac{n}{j} \\
= \langle p_i; p_g \rangle c + \langle p_i; p_g \rangle \frac{n}{n-k} + \langle p_i; p_g \rangle \sum_{j=n-k-x+1}^{n-2} \frac{n}{j} \\
= \langle p_i; p_g \rangle c + \sum_{j=n-k-x+1}^{n-2} \frac{n}{j} \\
\text{(recall that } LValue(p_i) = \{i, c\}) \\
\leq c + \langle b_y? \rangle \frac{n}{n-k} + \sum_{j=n-k-x+1}^{n-1} \frac{n}{j} = J.
\end{align*}
\]

Some first-order formula used above can be verified as
\[
p_g; b_y? = 1 \neq 0?; p_y = p_g.
\]
\( \langle p_{in}; p_y \rangle \frac{n}{n-k} \)

\( = \langle p_{in}; \bar{b}_y; p_y \rangle \frac{n}{n-k} \)

\( = \langle p_{in}; y[i] = 0?; y[i] \leftarrow 1 \rangle \frac{n}{n - \left( \sum_{j} \langle y[j] \neq 0? \rangle 1 \right)} \)

\( = \langle p_{in}; y[i] = 0?; y[i] \leftarrow 1 \rangle \frac{n}{n - \left( \sum_{j \neq i} \langle y[j] \neq 0? \rangle 1 + \langle y[i] \neq 0? \rangle 1 \right)} \)

\( = \langle p_{in}; y[i] = 0? \rangle \frac{n}{n - \left( \sum_{j \neq i} \langle y[j] \neq 0? \rangle 1 + \langle 1 \neq 0? \rangle 1 \right)} \)

\( = \langle p_{in}; y[i] = 0? \rangle \frac{n}{n - \left( \sum_{j \neq i} \langle y[j] \neq 0? \rangle 1 + 1 \right)} \)

\( = \langle p_{in} \rangle \frac{n}{n - \left( \sum_{j \neq i} \langle y[j] \neq 0? \rangle 1 + \langle y[i] \neq 0? \rangle 1 \right)} \)

\( = \langle p_{in} \rangle \frac{n}{n - \left( \sum_{j \neq i} \langle y[j] \neq 0? \rangle 1 + 1 \right)} \)

\( = \langle p_{in} \rangle \frac{n}{n - \left( \sum_{j \neq i} \langle y[j] \neq 0? \rangle 1 + \langle y[i] \neq 0? \rangle 1 \right)} \)

\( = \langle p_{in} \rangle \frac{n}{n - (k+1)} \)

\( = \langle p_{in} \rangle \frac{n}{n - k - 1} \).

and similarily we have,

\( \langle p_{in}; p_y \rangle \sum_{j=n-k-x+2}^{n-k-1} \frac{n}{j} = \langle p_{in} \rangle \sum_{j=n-k-x+1}^{n-k-2} \frac{n}{j} \).

Disclaimer: This work has been submitted to MFCS 2017.

References


