A Kleene algebra with tests for union bound reasoning about probabilistic programs

Leandro Gomes, Patrick Baillot, Marco Gaboardi

A Kleene algebra with tests for union bound reasoning about probabilistic programs

Leandro Gomes[a1], Patrick Baillot[b1], and Marco Gaboardi[c2]

[1]Université de Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France
[2]Boston University, USA

[a]leandrogomes.moreiragomes@univ-lille.fr
[b]patrick.baillot@univ-lille.fr
[c]gaboardi@bu.edu

Abstract

Kleene Algebra with Tests (KAT) provides a framework for algebraic equational reasoning about imperative programs. The recent variant Guarded KAT (GKAT) allows to reason on non-probabilistic properties of probabilistic programs. Here we introduce an extension of this framework called approximate GKAT (aGKAT), which equips GKAT with a partially ordered monoid (real numbers) enabling to express satisfaction of (deterministic) properties *except* with a probability up to a certain bound. This allows to represent in equational reasoning 'à la KAT' proofs of probabilistic programs based on the union bound, a technique from basic probability theory. We show how a propositional variant of approximate Hoare Logic (aHL), a program logic for union bound, can be soundly encoded in our system aGKAT. We also show with an example that aGKAT is more general than aHL, in the sense that it can prove some probability bounds that aHL cannot.

1 Introduction

Kleene algebra with tests (KAT) has been introduced in [20] as an algebraic framework for program verification. A KAT is a two-sorted structure, consisting of a Kleene algebra and a Boolean algebra of tests: the Kleene algebra part accounts for programs, with sequential composition, branching and iteration; the Boolean algebra part accounts for the predicates used to build if-then-else instructions, while loops and assertions, as well as, being KAT able to subsume propositional Hoare logic [21], for the pre and post-conditions. This framework allowed to give algebraic proofs corresponding to several approaches in program verification, see e.g. [21, 1, 23], and has been implemented as a library for the Coq proof assistant [27]. It has also been followed by several variants, like NetKAT [2], which allows to reason about software defined networks, Concurrent NetKAT [30] for concurrent networks, CKAO [17] for concurrent programs and more recently TopKAT for reasoning about incorrectness [31].

Recently the variant Guarded KAT (GKAT) [29] has been proposed as a restriction of KAT where all sums and iterations are guarded by tests. It offers several advantages over KAT, including the fact that the complexity of its equational theory is lower (almost linear time, provided that the number of tests is fixed) and the existence of a probabilistic model. The latter paves the way for using GKAT for reasoning about probabilistic programs. However an important feature of this system is that the tests of GKAT remain the same as those of KAT, namely they express Boolean properties on states. Therefore the framework of GKAT allows to encode probabilistic programs, but the assertions about them are non-probabilistic.

In this paper our goal is to extend the GKAT approach to reason about probabilistic programs satisfying properties *with a given probability bound* β. The objective is not to design an expressive framework for advanced probabilistic proofs, but instead to allow for simple probabilistic reasoning with a low technical overhead.

Concretely we target proofs based on the *union bound principle*, a property from basic probability theory, which can be stated as follows: given some properties $A_1, \ldots, A_n$, one has $Pr[\cup_{i=1}^{n} A_i] \leq \sum_{i=1}^{n} Pr[A_i]$.

This principle is ubiquitous when reasoning about properties of randomized algorithms [26] and in their application in security, privacy [9], learning theory [18], etc.

A previous approach for reasoning about probabilistic imperative programs using the union bound principle had been provided by the *union bound program logic* aHL [5]. This is a Hoare logic for reasoning about probabilistic programs with non-probabilistic assertions but with judgements carrying a numeric index for tracking the failure probability. That is, judgments have the form $\vdash_\beta c : \phi \Rightarrow \psi$ where β is an upper bound on the probability that $\neg\psi$ is true after executing c starting from a memory satisfying ϕ. The authors illustrated how this logic could be used for the verification of accuracy of some algorithms, in particular in the setting of differential privacy. A relational variant of this logic is handled by the Easycrypt tool [8, 4] , which can be used for proving properties of cryptographic protocols as well as differential privacy properties of programs.

A natural idea is thus to adapt the union bound logic aHL to the GKAT framework. To do this and capture the union bound reasoning in an algebraic framework we extend GKAT with an additional relation, denoted $\lhd$, relating GKAT expressions with elements of a partially ordered monoid, typically real numbers on $[0, 1]$. We call this new system *approximate* GKAT (aGKAT). An important feature of this structure is that we want the new setting to subsume standard GKAT, requiring aGKAT to satisfy the theory of GKAT. A second feature is that we want the probabilistic model of sub-Markov kernels to be a model of our new structure, when we consider the monoid of real numbers. For this particular instantiation, the meaning of $\lhd$ will be that $c \lhd \beta$ holds if *the probability of successful execution of program c is bounded by β*. The theory of aGKAT extends the one of GKAT, by a small set of axioms characterizing the properties of the new relation $\lhd$. We illustrate how this theory allows for a concise form of equational reasoning for establishing probability bounds on some GKAT programs. Moreover, in order to demonstrate the expressivity of aGKAT, we encode aHL in it. This is inspired by the classical result of Kozen [21] showing that propositional Hoare logic can be encoded in KAT.

Outline. In Sect. 2 we will recall GKAT and its probabilistic model, and in Sect. 3 we will recall the Hoare logic aHL. Then in Sect. 4 we will define our system, aGKAT, its theory and its semantics. After that in Sect. 5 we will provide an encoding of the logic aHL in aGKAT and prove its soundness. In Sect. 6 we give an example that can be handled in aGKAT but not in aHL. Sect. 7 is devoted to another example, the analysis in aGKAT of the accuracy of the probabilistic algorithm Report-noisy-max. Finally, Sec. 8 overviews related work and Sec. 9 enumerates possible directions for future work.

2 Guarded Kleene algebra with tests

This section recalls the language and the semantics of *Guarded Kleene Algebra with Tests* (GKAT) [29], an abstraction of imperative programs where conditionals and loops are encoded as guarded sums $(c_1 +_b c_2)$ and guarded iterations $(c^{(b)})$, respectively, guarded by Boolean predicates b. The structure is a restriction of KAT in which we are not allowed to freely use operators $+$ and * to build terms. In other words, GKAT does not allow nondeterminism.

2.1 Syntax

The syntax of GKAT is defined with a set of *actions* Σ and a finite set of primitive tests $\mathtt{T}$, which are disjoint. We denote actions by a and primitive tests by p. The sets of Boolean expressions $\mathtt{BExp}$ (also called tests) and GKAT expressions $\mathtt{Exp}$ (also called programs) are then defined by the following grammars:

$b, b_1, b_2 \in \mathtt{BExp} ::=$	
0	**false**
1	**true**
$p \in \mathtt{T}$	p
$b_1 \cdot b_2$	b_1 **and** b_2
$b_1 + b_2$	b_1 **or** b_2
$\bar{b}$	**not** b

$c, c_1, c_2 \in \mathtt{Exp} ::=$	
$a \in \Sigma$	**do** a
$b \in \mathtt{BExp}$	**assert** b
$c_1 \cdot c_2$	c_1; c_2
$c_1 +_b c_2$	**if** b **then** c_1 **else** c_2
$c^{(b)}$	**while** b **do** c

where, for any $b, b_1, b_2 \in \mathtt{BExp}$, operators $\cdot$, $+$ and $\bar{}$ denote conjunction, disjunction and negation, respectively, and, for any $c, c_1, c_2 \in \mathtt{Exp}$, the operator $\cdot$ denotes sequential composition. The notations on the r.h.s. are given to help intuition and will sometimes be used when writing programs. We introduce command **skip** as a shorthand for **assert** 1, which is encoded by the Boolean expression 1.

The precedence of the operators is the usual one, i.e. the operator $\cdot$ has higher precedence than operator $+_b$, and $()^{(b)}$ has higher precedence that $\cdot$[1] To simplify the writing, we often omit the operator $\cdot$ by writing $c_1 c_2$ for the expression $c_1 \cdot c_2$, for any $c_1, c_2 \in \texttt{Exp}$.

We are interested in using GKAT for representing probabilistic programs. For that, let us first fix a few definitions. Given a set S, $\mathcal{D}(S)$ is the set of *probability sub-distributions*[2] over S with discrete support, i.e. the set of functions $f : S \to [0,1]$ such that $Supp(f) = \{x \in S \mid f(x) > 0\}$ is discrete and f sums up to at most 1, i.e. $\sum_{s \in S} f(s) \leq 1$. In particular, the *Dirac* distribution $\delta_s \in \mathcal{D}(S)$ is the map

$$w \to [w = s] = \begin{cases} 1, \text{ if } w = s \\ 0, \text{ otherwise} \end{cases}$$

Example 2.1 (Imperative programming language)**.** *Take a set* $\texttt{Var}$ *of variables and a set* $\texttt{Distr}$ *of sub-distributions over* $\mathbb{R}$ *with discrete support. Consider a simple imperative programming language defined by the following grammar:*

$$\begin{aligned} \textit{terms } t \in \texttt{Terms} &::= x \in \texttt{Var} \mid r \in \mathbb{R} \mid t_1 + t_2 \mid t_1 - t_2 \mid t_1 \times t_2 \\ \textit{distributions } d \in \texttt{Distr} & \\ \textit{tests } b \in \texttt{Tests} &::= \textbf{\textit{false}} \mid \textbf{\textit{true}} \mid t_1 < t_2 \mid t_1 = t_2 \mid \textbf{\textit{not}}\ b \mid b_1\ \textbf{\textit{and}}\ b_2 \mid b_1\ \textbf{\textit{or}}\ b_2 \\ \textit{commands } c \in \texttt{Comm} &::= \textbf{\textit{skip}} \mid x \leftarrow t \mid x \overset{\$}{\leftarrow} d \mid c_1;\ c_2 \mid \textbf{\textit{if}}\ b\ \textbf{\textit{then}}\ c_1\ \textbf{\textit{else}}\ c_2 \mid \textbf{\textit{while}}\ b\ \textbf{\textit{do}}\ c \end{aligned}$$

This language can be modeled in GKAT by taking as sets of actions and primitive tests respectively $\Sigma = \{x \leftarrow t,\ x \overset{\$}{\leftarrow} d \mid x \in \texttt{Var}, t \in \texttt{Terms}, d \in \texttt{Distr}\}$ *and* $T = \{t_1 < t_2,\ t_1 = t_2 \mid t_1, t_2 \in \texttt{Terms}\}$*. The first action evaluates term* t *and assigns the result to* x *and the second one samples from* d *and assigns the result to* x*. Note that technically speaking according to the definition of GKAT the set* T *should be chosen finite, which is not the case here, but as observed in [29] Sect. 2.3 Example 2.5 we can use a finite subset* T' *of* T *for reasoning on pairwise equivalence of programs which terminate.*

Observe that while programs c *may be probabilistic, due to the use of samplings, the tests* b *as for them are deterministic, i.e. they do not use any probabilistic primitives. In particular the conditional branching in programs is only done on deterministic tests.*

2.2 Semantics

We now present the semantic interpretation of GKAT that we will be using, the *Probabilistic model* [29] [3]. We first review some basic concepts needed for the semantics. The *Iverson bracket* $[b]$, for $b \in \texttt{BExp}$, is the function taking value 1 if b is true and 0 otherwise. Typical models of probabilistic imperative programming languages interpret programs as *Markov kernels* on a set S, i.e. maps from S to probability distributions. The semantic model defined below interprets programs as *sub-Markov* kernels, i.e. Markov kernels on sub-distributions.

Definition 2.1 (Probabilistic interpretation)**.** *Let* $i = (\texttt{State}, eval, sat)$ *be a triple where:*

- $\texttt{State}$ *is a set of states,*
- *for each action* $a \in \Sigma$*,* $eval(a) : \texttt{State} \to \mathcal{D}(\texttt{State})$ *is a sub-Markov kernel,*
- *for each primitive test* $p \in T$*,* $sat(p) \subseteq \texttt{State}$ *is a set of states.*

The *probabilistic interpretation* of an expression c with respect to i is the sub-Markov kernel $\mathcal{P}_i[\![c]\!] : \texttt{State} \to \mathcal{D}(\texttt{State})$ defined as follows:

1. $\mathcal{P}_i[\![a]\!] := eval(a)$
2. $\mathcal{P}_i[\![b]\!](\sigma) := [\sigma \in sat^\dagger(b)] \times \delta_\sigma$
3. $\mathcal{P}_i[\![c_1 \cdot c_2]\!](\sigma)(\sigma') := \sum_{\sigma''} \mathcal{P}_i[\![c_1]\!](\sigma)(\sigma'') \times \mathcal{P}_i[\![c_2]\!](\sigma'')(\sigma')$
4. $\mathcal{P}_i[\![c_1 +_b c_2]\!](\sigma) := [\sigma \in sat^\dagger(b)] \times \mathcal{P}_i[\![c_1]\!](\sigma) + [\sigma \in sat^\dagger(\bar{b})] \times \mathcal{P}_i[\![c_2]\!](\sigma)$

[1] For example the GKAT expression $c_1^{(b_1)} \cdot c_2 +_{b_2} c_3$ reads as $((c_1^{(b_1)}) \cdot c_2) +_{b_2} c_3$.

[2] Some examples of distributions are the tossing of a fair coin, with probability 0.5 for 0 and 1, and the (discrete version of the) Laplacian distribution $\mathcal{L}_p(a)$ centered in a with parameter p. The density function of $\mathcal{L}_p(a)$ is given by $\frac{1}{2p} \exp(\frac{|x-a|}{p})$.

[3] Note that more interpretations of GKAT are presented in [29], namely a relational model and a trace model.

5. $\mathcal{P}_i[\![c^{(b)}]\!](\sigma)(\sigma') := \lim_{n\to\infty} \mathcal{P}_i[\![(c +_b 1)^n \cdot \bar{b}]\!](\sigma)(\sigma')$

where $\mathrm{sat}^\dagger : \mathtt{BExp} \to 2^{\mathtt{State}}$ is the lifting of $\mathrm{sat} : \mathtt{T} \to 2^{\mathtt{State}}$ to arbitrary Boolean expressions over $\mathtt{BExp}$, and $\times$ denotes both multiplication on real numbers and the pointwise multiplication on sub-distributions. For instance the definition of $\mathcal{P}_i[\![b]\!](\sigma)$ means that it is either δ_σ if σ belongs to $sat^\dagger(b)$, or the constant sub-distribution equal to 0 otherwise. Intuitively $\mathcal{P}_i[\![c]\!](\sigma)(\sigma')$ is the probability that the execution of c on initial state σ terminates on state σ', and $\sum_{\sigma'} \mathcal{P}_i[\![c]\!](\sigma)(\sigma')$ is the probability that the execution of c on initial state σ terminates on a state (we then also say that it is a successfull execution). Observe thus that we really need to consider sub-distributions and not only distributions.

Remark 2.1 (Finite state case)**.** *In the case where* $\mathtt{State}$ *is a finite set of size n, say $\{s_1, \ldots, s_n\}$ then a sub-Markov kernel f can be represented as an $n \times n$ matrix $\mathcal{M} = (a_{i,j})_{i,j\in[1,n]}$. Each coefficient $a_{i,j}$ is defined as $a_{i,j} = f(s_i)(s_j)$. So in particular the sum over each line is inferior or equal to 1. We denote $f = \mathcal{M}$. For tests b, the matrix $\mathcal{P}_i[\![b]\!]$ has only diagonal coefficients, with value $a_{i,i} = 1$ if $s_i \in sat^\dagger(b)$, $a_{i,i} = 0$ if s_i not in $sat^\dagger(b)$. In the case of $c_1 \cdot c_2$, the matrix $\mathcal{P}_i[\![c_1 \cdot c_2]\!]$ is obtained by the matrix product of $\mathcal{P}_i[\![c_1]\!]$ and $\mathcal{P}_i[\![c_2]\!]$. See Appendix 14 for an example.*

In the following we will consider programs over a finite set of variables $\mathtt{Var}$ and the set of states will be the set of *memories*, that is to say functions in $\mathtt{Var} \to D$ where D is the domain of variables (we can take for instance $D = \mathbb{Q}$, the rational numbers). If $x \in \mathtt{Var}$ and σ is a memory, then $\sigma[x \leftarrow t]$ is the memory identical to σ except that it maps x to the evaluation of t in memory σ. The interpretation of actions $a \in \Sigma$ as sub-Markov Kernels is then given by $eval(x \leftarrow t)(\sigma) := \delta_{\sigma[x\leftarrow t]}$ and $eval(x \overset{\$}{\leftarrow} d)(\sigma) := \sum_{t\in Supp(d)} d(t) \cdot \delta_{\sigma[x\leftarrow t]}$.
In the sequel memories will often be denoted as m.

2.3 Axioms

The theory of GKAT introduced in [29] is given by the axioms from Fig. 1. Note in particular the

$$
\begin{array}{rcll}
c +_b c & = & c & (1)\\
c_1 +_b c_2 & = & c_2 +_{\bar{b}} c_1 & (2)\\
(c_1 +_{b_1} c_2) +_{b_2} c_3 & = & c_1 +_{b_1\cdot b_2} (c_2 +_{b_2} c_3) & (3)\\
c_1 +_b c_2 & = & b \cdot c_1 +_b c_2 & (4)\\
c_1 \cdot c_3 +_b c_2 \cdot c_3 & = & (c_1 +_b c_2) \cdot c_3 & (5)\\
(c_1 \cdot c_2) \cdot c_3 & = & c_1 \cdot (c_2 \cdot c_3) & (6)\\
0 \cdot c & = & 0 & (7)
\end{array}
\qquad
\begin{array}{rcll}
c \cdot 0 & = & 0 & (8)\\
1 \cdot c & = & c & (9)\\
c \cdot 1 & = & c & (10)\\
c^{(b)} & = & c \cdot c^{(b)} +_b 1 & (11)\\
(c +_{b_2} 1)^{(b_1)} & = & (b_2 \cdot c)^{(b_1)} & (12)\\
\dfrac{c_3 = c_1 \cdot c_3 +_b c_2}{c_3 = c_1^{(b)} \cdot c_2} & \text{if} & E(c_1) = 0 & (13)
\end{array}
$$

Figure 1: Axiomatisation of Guarded Kleene algebra with tests

fixpoint axiom (13). Intuitively, it says that if expression c_3 chooses (using guard b) between executing c_1 and looping again, and executing c_2, then c_3 is a b-guarded loop followed by c_2. However, the rule is not sound in general. In order to overcome this limitation, following [29] (Section 3.1, Definition 3.2), the side condition $E(c_1) = 0$ is introduced, ensuring that command c_1 is *productive*, i.e. that it performs some action. To this end, the function E is inductively defined as follows: $E(b) := b$, $E(a) := 0$, $E(c_1 +_b c_2) := b \cdot E(c_1) + \bar{b} \cdot E(c_2)$, $E(c_1 \cdot c_2) := E(c_1) \cdot E(c_2)$, $E(c^{(b)}) := \bar{b}$. We can see $E(c)$ as the weakest test that guarantees that command c terminates successfully but does not perform any action.

Moreover, note particularly the following observation: in KAT the encoding $c_1(bc_2 + \bar{b}c_3) = c_1bc_2 + c_1\bar{b}c_3$ is not an **if-then-else** statement; it is rather a nondeterministic choice between executing c_1, then testing b and executing c_2, and executing c_1, then testing $\bar{b}$ and executing c_3. The corresponding encoding in GKAT would be $c_1(c_2 +_b c_3) = c_1c_2 +_b c_1c_3$, an equality which is actually *not* valid in GKAT. Since GKAT is restricted to deterministic programs, there is no valid correspondence between the KAT encoding, which is not an **if-then-else** statement, and the hypothetical correspondent GKAT encoding, which is not valid. That is why left distributivity does not hold in GKAT for any $c \in \mathtt{Exp}$; it only holds for the particular case of $c_1 \in \mathtt{BExp}$, i.e. if c_1 is a test.

We define the relation $\leq$ on tests as: $b_1 \leq b_2$ iff $b_1 + b_2 = b_2$. Contrarily to KAT [20], the relation $\leq$ is not defined on an arbitrary GKAT expression, only on tests. In Appendix 10 we recall additional derivable equations in GKAT from [29].

Since any test is a program ($\texttt{BExp} \subseteq \texttt{Exp}$), the grammar also allows to write expressions as $b_1 +_b b_2$, for any $b \in \texttt{BExp}$. We thus establish the following proposition[4] (proof in Appendix 11) which expresses the guarded sum $+_b$, for any $b \in \texttt{BExp}$, in terms of the disjunction $+$ on tests.

Proposition 2.1. *For any tests b, b_1, b_2 one has: $b_1 +_b b_2 = bb_1 + \bar{b}b_2$.*

By Boolean reasoning, we can observe that $bb + \bar{b}\bar{b} = 1$. This observation will be useful later to prove the soundness of some aHL rules in aGKAT.

We also state the following proposition (see Appendix 12 for the proof):

Proposition 2.2. *For any tests b_1, b_2 one has: $b_1 + b_2 = b_1 +_{b_1} b_2$.*

3 Union bound logic - Approximate Hoare logic

In this section we recall *Approximate Hoare logic (aHL)* [5], a logic based on the union bound, a tool from probability theory for analyzing randomised algorithms. A judgment in aHL is of the form $\vdash_\beta c : \phi \Rightarrow \psi$ where: ϕ, ψ are first-order formulas representing non probabilistic pre- and post-conditions[5], respectively; β is a value in $[0, 1]$ and it is an upper bound on the probability that the post-condition ψ does *not* hold on the output distribution, assuming that ϕ holds on an initial memory m. We assume a probabilistic interpretation i and we will denote $m \models \phi$ if ϕ is valid in memory m. The validity of the judgement is thus stated by:

Definition 3.1 (Validity of aHL judgment)**.** *A judgment $\vdash_\beta c : \phi \Rightarrow \psi$ is valid if for every memory m such that $m \models \phi$, we have $\mathcal{P}_i[\![c]\!](m)[\bar{\psi}] \leq \beta$.*

Figure 2 presents the deduction rules of aHL. Let us comment on some of these rules. The rule (*Rand*) handles sampling from a distribution d; we can assume a postcondition ψ after the sampling, provided that under the assumption of precondition ϕ, the statement ψ fails with probability at most β.

The other rules are similar to standard Hoare logic rules annotated with suitable probability indexes β. The rule (*Seq*) says that when composing two programs c and c', the failure probabilities of the two programs with respect to their postconditions add together. The (*Cond*) rule states that if the two branches of the conditional have the same index β, then we can keep the index β for the conditional. In rule (*Weak*) the premise $\models \phi' \Rightarrow \phi$ means that, in any model, ϕ' implies ϕ. This (*Weak*) rule allows to strengthen the precondition, weaken the postcondition, and increase the index β (which means overapproximating the failure probability). The (*And*) rule can be seen as an application of the union bound principle. It enables to combine two postconditions by a conjunction, provided we add up the failure probabilities. As to the (*Or*) rule, it allows to take the disjunction of two preconditions, if they have the same failure probability, and keep this index for the disjunction. Note that thanks to the (*Weak*) rule we could also in (*Or*) consider two indexes β and β' in the premises, and their maximum in the conclusion (the same is also true for (*Cond*)). The rule (*False*) might first seem a bit strange as it allows to conclude false, but note that its index is 1, which means that false holds in the final memory with probability 0. Finally, considering the (*While*) rule, observe that it is slightly more restrictive than the corresponding classical one of Hoare logic. Its side conditions ensure that the loop terminates in at most k iterations except with probability $k\beta$. Its first side condition states that the variable b_v only takes non-negative integer values.

4 Approximate Guarded Kleene algebra with tests (aGKAT)

4.1 Definition and theory of aGKAT

Recalling that GKAT encodes only Boolean assertions on probabilistic programs, we want to extend this kind of reasoning in order to capture aHL properties. We want to define a structure which would allow to express the fact that a probabilistic program c satisfies a deterministic postcondition, *except* with a probability up to a certain bound. For that we will extend GKAT with a relation between a GKAT expression and a value β from a partially ordered set. Such a set is defined as follows:

Definition 4.1. *A preordered double monoid (pod-monoid) is a $\mathcal{M} = (M, \leq, \cdot, 1, +, 0)$ where:*

[4] We thank the anonymous reviewer of another paper for pointing out to us the fact that this property is derivable in GKAT.

[5] Note that ϕ and ψ are properties of memories rather that properties of distributions over memories.

- *Skip*:

$$\frac{}{\vdash_0 \mathtt{skip} : \phi \Rightarrow \phi}$$

- *Assn*:

$$\frac{}{\vdash_0 \mathbf{do}\ x \leftarrow t : \phi[t/x] \Rightarrow \phi}$$

- *Seq*:

$$\frac{\vdash_\beta c : \phi \Rightarrow \phi' \quad \vdash_{\beta'} c' : \phi' \Rightarrow \phi''}{\vdash_{\beta+\beta'} c;\ c' : \phi \Rightarrow \phi''}$$

- *Weak*:

$$\frac{\models \phi' \Rightarrow \phi \quad \vdash_\beta c : \phi \Rightarrow \psi \quad \models \psi \Rightarrow \psi' \quad \beta \leq \beta'}{\vdash_{\beta'} c : \phi' \Rightarrow \psi'}$$

- *Or*:

$$\frac{\vdash_\beta c : \phi \Rightarrow \psi \quad \vdash_\beta c : \phi' \Rightarrow \psi}{\vdash_\beta c : \phi \vee \phi' \Rightarrow \psi}$$

- *Rand*:

$$\frac{\forall\, m : m \models \phi \Rightarrow \mathcal{P}_i[\![x \xleftarrow{\$} d]\!](m)[\bar{\psi}] \leq \beta}{\vdash_\beta \mathbf{do}\ x \xleftarrow{\$} d : \phi \Rightarrow \psi}$$

- *Cond*:

$$\frac{\vdash_\beta c : \phi \wedge b \Rightarrow \psi \quad \vdash_\beta c' : \phi \wedge \bar{b} \Rightarrow \psi}{\vdash_\beta \mathbf{if}\ b\ \mathbf{then}\ c\ \mathbf{else}\ c' : \phi \Rightarrow \psi}$$

- *And*:

$$\frac{\vdash_\beta c : \phi \Rightarrow \psi \quad \vdash_{\beta'} c : \phi \Rightarrow \psi'}{\vdash_{\beta+\beta'} c : \phi \Rightarrow \psi \wedge \psi'}$$

- *False*:

$$\frac{}{\vdash_1 c : \phi \Rightarrow \bot}$$

- *While*:

$$\frac{b_v : \mathbb{N}, \qquad \models (\phi \wedge b_v \leq 0 \rightarrow \bar{b}),\ \vdash_\beta c : \phi \Rightarrow \phi,\ \forall_{\eta>0} : \vdash_0 c : \phi \wedge b \wedge (b_v = \eta) \Rightarrow (b_v < \eta)}{\vdash_{k\cdot\beta} \mathbf{while}\ b\ \mathbf{do}\ c : \phi \wedge (b_v \leq k) \Rightarrow \phi \wedge \bar{b}}$$

Figure 2: Approximate Hoare Logic rules (aHL)

- $\leq$ *is a preorder on* M,
- $(M, \cdot, 1)$ *and* $(M, +, 0)$ *are two monoid structures, whose operations* $\cdot$ *and* $+$ *are monotone w.r.t.* $\leq$.

Note that we do not include any axiom relating $\cdot$ and $+$. This structure is thus sufficient to model the probability bounds from aHL. In the sequel we will consider the pod-monoid consisting of the real unit interval $[0, 1]$ equipped with multiplication and addition truncated to 1, that is to say $min\,((\beta_1 + \beta_2), 1)$, where $+$ is the ordinary addition.

We then give the main definition of this section.

Definition 4.2. *Approximate GKAT, denoted as aGKAT, is an extension of GKAT with a pod-monoid* $\mathcal{M}$ *and a predicate symbol* $\triangleleft$ *on* $\mathbf{Exp} \times \mathcal{M}$. *The theory of aGKAT is the union of axioms of pod-monoid, and those of Fig. 1 and Fig. 3.*

$$(c_1 = c_2 \ \wedge\ c_1 \triangleleft \beta) \ \Rightarrow\ c_2 \triangleleft \beta \tag{14}$$
$$(c \triangleleft \beta_1 \ \wedge\ \beta_1 \leq \beta_2) \ \Rightarrow\ c \triangleleft \beta_2 \tag{15}$$
$$(c \cdot c_1 \triangleleft \beta_1 \ \wedge\ c \cdot c_2 \triangleleft \beta_2) \ \Rightarrow\ c \cdot (c_1 +_b c_2) \triangleleft \beta_1 + \beta_2 \tag{16}$$
$$(c_1 \triangleleft \beta_1 \ \wedge\ c_2 \triangleleft \beta_2) \ \Rightarrow\ c_1 \cdot c_2 \triangleleft \beta_1 \cdot \beta_2 \tag{17}$$
$$(c_1 \triangleleft \beta \ \wedge\ c_2 \triangleleft \beta) \ \Rightarrow\ c_1 +_b c_2 \triangleleft \beta \tag{18}$$
$$c \triangleleft 1 \tag{19}$$
$$0 \triangleleft 0 \tag{20}$$

Figure 3: Axioms on the relation $\triangleleft$

Recall that the intended meaning of $\triangleleft$ in the case where $\mathcal{M} = ([0, 1], \leq, \cdot, 1, +, 0)$ is that $c \triangleleft \beta$ holds if the probability of successful execution of program c is bounded by β. Observe that the $\triangleleft$-axioms of Fig. 3 are arguably simple, are Horn clauses and that none deals with guarded iteration $c^{(b)}$.

Let us explain some intuitions underlying these axioms. Axiom (19) simply says that any program has a probability of successful execution bounded by 1, while (20) states that program 0 (which is **assert false**) has probability 0 of successful execution. Axiom (15) says that the statement still holds if we

increase the probability β_1. Axiom (14) states that programs which are equal (up to the GKAT axioms of Fig. 1) admit the same probability of successful execution. Axiom (18) says that if the two branches of a conditional admit a bound β for their successful execution, so does the conditional itself. As to Axiom (17), its meaning is that the probability of successful execution of the composition of two programs c_1 and c_2 is bounded by the product of the probabilities of successful execution of respectively c_1 and c_2.

Maybe the less intuitive axiom is Axiom (16). Note that the difference with Axiom (18) is that for Axiom (18) any initial state s either satisfies b or $\overline{b}$, and so only one branch of $c_1 +_b c_2$ is explored. By contrast in Axiom (16) any initial state s might lead by the probabilistic execution of c *both* to states satisfying b and to states satisfying $\overline{b}$, so triggering both branches of the conditional. We will come back to this axiom in Remark 4.2 below.

After these intuitive considerations, we now formally define a semantic interpretation of aGKAT as follows:

Definition 4.3. *A probabilistic interpretation of aGKAT is obtained by extending a probabilistic interpretation $\mathcal{P}_i$ of GKAT given in Sect. 2.2 in the following way:*

- *we consider a triple $i = (\mathtt{State}, eval, sat)$ of Def. 2.1 interpreting GKAT,*
- *the pod-monoid $\mathcal{M}$ is interpreted as indicated above by $([0,1], \leq, \cdot, 1, +, 0)$ where $\cdot$ is the product and $+$ the truncated sum,*
- *the predicate $\triangleleft$ is interpreted by the relation between sub-Markov kernels f and $[0,1]$-reals β consisting in the pairs (f, β) satisfying $\forall\, s \in \mathtt{State}$, $\Sigma_{s' \in \mathtt{State}} f(s)(s') \leq \beta$, i.e. for any s, the total mass of the sub-distribution $f(s)$ is bounded by β.*

We still denote the interpretation of an expression c as $\mathcal{P}_i[\![c]\!]$.

If i is an interpretation and F a 1st-order formula on the signature consisting of terms in *Exp* and in real numbers and predicates $=$ and $\triangleleft$, we write $i \models F$ if F is valid in the model defined by i. By abuse we will simply write $\models F$ if i is clear from the context. So by the definition above we have in particular that $i \models c \triangleleft \beta$ if $\forall\, s \in \mathtt{State}$, $\sum\limits_{s' \in \mathtt{State}} \mathcal{P}_i[\![c]\!](s)(s') \leq \beta$.

We now establish the following proposition.

Proposition 4.1. *Any probabilistic interpretation of aGKAT is a model of its theory, i.e.:*

1. *the interpretation of aGKAT expressions satisfies the axioms of GKAT (Fig. 1) and that of $([0,1], \leq, \cdot, 1, +, 0)$ satisfies the axioms of pod-monoid,*
2. *the axioms of Fig. 3 (axioms (14) to (20)) are satisfied.*

Proof. (Prop. 4.1) See Appendix 13. The most delicate case is that of Axiom (16). □

Remark 4.1. *Proposition 4.1 implies that if i is a probabilistic interpretation and if a statement $c \triangleleft \beta$ is derivable from the aGKAT axioms and possibly some semantic hypothesis of the shape $i \models A$, then $i \models c \triangleleft \beta$ holds. Note that Prop. 4.1 implies in particular that if i is a probabilistic interpretation and $A \Rightarrow B$ is an instance of an axiom in Fig. 3, then if $i \models A$ we can deduce that $i \models B$. This is because as i is a model, classical logic rules are sound in it.*

Remark 4.2. *Note that by analogy with Axiom (18), one could have expected an axiom stronger than Axiom (16), namely that if $(c \cdot c_1 \triangleleft \beta \ \wedge \ c \cdot c_2 \triangleleft \beta)$ then one would have $c \cdot (c_1 +_b c_2) \triangleleft \beta$ (this would then generalize Axiom (18) when taking $c = 1$). However it turns out that this candidate additional axiom is not valid in the probabilistic model. A counter-example is given in the Appendix 14.*

Proposition 4.2. *The following property is derivable in aGKAT:*

$$(c \cdot b_1 \triangleleft \beta_1 \ \wedge \ c \cdot b_2 \triangleleft \beta_2) \Rightarrow c \cdot (b_1 + b_2) \triangleleft \beta_1 + \beta_2$$

Proof. Observe that $b_1 + b_2 = b_1 +_{b_1} b_2$ by Prop. 2.2 and use Axiom (16). □

The proposition below refines in some sense Axiom (16).

Proposition 4.3. *The following property is derivable in aGKAT:*

$$(c \cdot b \cdot c_1 \triangleleft \beta_1 \ \wedge \ c \cdot \overline{b} \cdot c_2 \triangleleft \beta_2) \Rightarrow c \cdot (c_1 +_b c_2) \triangleleft \beta_1 + \beta_2$$

Proof. Observe that $c_1 +_b c_2 = b \cdot c_1 +_b \bar{b} \cdot c_2$ by Axiom (4), Axiom (2), applied two times. Then apply Axiom (16) to c, $c_1' = bc_1$ and $c_2' = bc_2$. □

Remark 4.3 (Axiom (16), left distributivity and union bound)**.** *Recall that KAT [20] has an axiom of left distributivity $c \cdot (c_1 + c_2) = c \cdot c_1 + c \cdot c_2$. It does not hold in GKAT with the guarded sum $+_b$ though. In some sense axiom (16) (or its refinement Prop. 4.3) can be seen as a kind of compensation for this lack of left distributivity because it allows, when one is reasoning about an expression $c \cdot (c_1 +_b c_2)$ (in order to establish a bound β), to continue the proof with two branches, respectively on $c \cdot c_1$ and on $c \cdot c_2$. If one obtains two bounds $c \cdot c_i \triangleleft \beta_i$, for $i = 1, 2$ then one can deduce that $c \cdot (c_1 +_b c_2) \triangleleft \beta_1 + \beta_2$.*

Moreover if c_1 and c_2 are tests b_1 and b_2, then by Prop. 2.2 $b_1 + b_2 = b_1 +_{b_1} b_2$. So $c \cdot (b_1 + b_2) = c \cdot (b_1 +_{b_1} b_2) \triangleleft \beta_1 + \beta_2$. So the probability that after execution of c the test $(b_1 + b_2)$ is satisfied is inferior to the sum of the probability that b_1 is satisfied and of the probability that b_2 is satisfied. This is the application of the binary union bound principle on post-conditions, and it can easily be applied to an arbitrary union bound.

4.2 Semantic reasoning

When reasoning about concrete programs, we want to establish properties on their semantic interpretations. That might sometimes require, besides the axioms of aGKAT, the use of some semantic properties. One such example is that some actions can be commuted without changing the semantics of the program. We establish thus some notations:

Definition 4.4. *Given two GKAT program c and c' and a probabilistic interpretation i, we write $c \equiv c'$ if $i \models c = c'$, i.e. $\mathcal{P}_i[\![c]\!] = \mathcal{P}_i[\![c']\!]$.*

This definition is required to establish the following proposition.

Proposition 4.4. *Consider GKAT programs c and c', and a probabilistic interpretation i.*

1. *If b is a test which only depends on the values of some variables $x_1, \ldots, x_n$ and if c leaves the values of those variables unchanged, then we have $c \cdot b \equiv b \cdot c$,*
2. *If $c \equiv c'$ and $i \models c \triangleleft \beta$, then $i \models c' \triangleleft \beta$.*

Observe that (1) holds because the syntax of programs does not allow any form of aliasing and (2) because the property $i \models c \triangleleft \beta$ only depends on the semantic interpretation $\mathcal{P}_i[\![c]\!]$.

Let us now illustrate the use of aGKAT on a small example.

Example 4.1 (Double tossing)**.** *Consider the program c below:*

$$c \quad = \quad (x \overset{\$}{\leftarrow} Coin) \cdot (c_1 +_{(x=1)} y \leftarrow 0), \qquad \text{where } c_1 = (x \overset{\$}{\leftarrow} Coin) \cdot (y \leftarrow 1 +_{(x=1)} y \leftarrow 0)$$

Consider the interpretation i where Coin is the distribution of a fair coin, that takes value 0 (resp. 1) with probability 1/2 (resp. 1/2). This can be represented either by adding to the theory two axioms describing the behaviour of Coin, namely axioms $(x \overset{\$}{\leftarrow} Coin) \cdot (x = 1) \triangleleft 1/2$ and $(x \overset{\$}{\leftarrow} Coin) \cdot (x \neq 1) \triangleleft 1/2$, or by using the following semantic properties of i: $\models (x \overset{\$}{\leftarrow} Coin) \cdot (x = 1) \triangleleft 1/2$ and $\models (x \overset{\$}{\leftarrow} Coin) \cdot (x \neq 1) \triangleleft 1/2$. We want to prove that after the execution of c, the probability that y equals 0 is below $3/4$, and the probability that y equals 1 is below $1/4$, i.e. $\models c \cdot (y = 0) \triangleleft 3/4$ and $\models c \cdot (y = 1) \triangleleft 1/4$.

Recall first that as i is a model, by Prop. 4.1, it satisfies all axioms of Fig. 1 and Fig. 3, and all classical logic rules are sound in it (see Remark 4.1). Now, by using Axiom (5), $c \cdot (y = 0)$ can be rewritten as follows:

$$\begin{aligned} c \cdot (y = 0) \quad &= \quad (x \overset{\$}{\leftarrow} Coin) \cdot (c_1' +_{(x=1)} y \leftarrow 0(y = 0)) \\ c_1' \quad &= \quad (x \overset{\$}{\leftarrow} Coin) \cdot (y \leftarrow 1(y = 0) +_{(x=1)} y \leftarrow 0(y = 0)) \end{aligned}$$

Let us name the following expressions, corresponding to the various possible branches of executions of $c \cdot (y = 0)$:

$$\begin{aligned} c_2 \quad &= \quad (x \overset{\$}{\leftarrow} Coin) \cdot (x = 1) \cdot (x \overset{\$}{\leftarrow} Coin) \cdot (x = 1) \cdot (y \leftarrow 1) \cdot (y = 0) \\ c_3 \quad &= \quad (x \overset{\$}{\leftarrow} Coin) \cdot (x = 1) \cdot (x \overset{\$}{\leftarrow} Coin) \cdot (x \neq 1) \cdot (y \leftarrow 0) \cdot (y = 0) \\ c_4 \quad &= \quad (x \overset{\$}{\leftarrow} Coin) \cdot (x \neq 1) \cdot (y \leftarrow 0) \cdot (y = 0) \end{aligned}$$

First, from the model we know that $(y \leftarrow 1) \cdot (y = 0) \equiv 0$*, so* $c_2 \equiv 0$*, so* $\models c_2 \lhd 0$*.*

Then, as $\models (x \xleftarrow{\$} \mathit{Coin}) \cdot (x \neq 1) \lhd 1/2$*, by Axioms (19) and (17) we have* $\models c_4 \lhd 1/2$*.*

Then, as $\models (x \xleftarrow{\$} \mathit{Coin}) \cdot (x = 1) \lhd 1/2$*, by Axioms (17) and (19) we have* $\models c_3 \lhd 1/4$*.*

$$\textit{By applying Prop. 4.3 to } c_2 \textit{ and } c_3 \textit{ we get: } \models (x \xleftarrow{\$} \mathit{Coin}) \cdot (x = 1) \cdot c_1' \quad \lhd \quad 1/4 \tag{21}$$

By applying again Prop. 4.3, this time to (21) and by $\models c_4 \lhd 1/2$ *we finally obtain* $\models c \cdot (y = 0) \lhd 3/4$ $(= 1/4 + 1/2)$*. We give in Appendix 15 a step-by-step fully explicit version of the proof above. The proof that* $\models c \cdot (y = 1) \lhd 1/4$ *holds is similar.*

5 Encoding aHL in aGKAT

We want to relate deduction in aHL and reasoning in aGKAT, by following the approach of [21] on the encoding of propositional Hoare logic in KAT. We consider the programming language of Example 2.1 but the results remain valid if we consider extended grammars of terms, distributions, tests and commands, where the class of tests is closed by substitution of terms t (as in Example 2.1). We will encode aHL derivations consisting of judgements $\vdash_\beta c : \phi \Rightarrow \phi'$ where ϕ and ϕ' belong to the class of tests.

Concretely the idea will be to encode the aHL judgement $\vdash_\beta c : \phi \Rightarrow \phi'$ by the aGKAT statement $\phi \cdot c \cdot \overline{\phi'} \lhd \beta$. Similarly to [21], showing that an aHL rule is sound in aGKAT will consist in proving that the conjunction of the aGKAT equations encoding the premises of the aHL rule implies the equation encoding the conclusion of the rule.

Observe that similarly as for Hoare logic, some rules of aHL, namely axiom rules (*Assn*) and (*Rand*), do not depend on aHL judgements as premises but rather on an interpretation of actions and predicates, and possibly a semantic condition (for (*Rand*)). Thus we do not expect to derive their encoding as an equation valid in the theory of aGKAT. Instead, one could add new axioms corresponding to (*Assn*) and (*Rand*) for specific distributions (as mentioned in Example 4.1), or alternatively when dealing with examples consider a particular interpretation i and thus reason on equalities of expressions in the model.

Fig. 4 lists the interpretations of the rules of aHL (Figure 2) in aGKAT, by encoding aHL judgments as aGKAT equations. Note that the rule (*Assn*) uses the test $\phi[t/x]$ obtained by substituting the term t in ϕ, which does belong to the class of tests by definition.

- Skip:
$$\phi 1 \bar{\phi} \lhd 0 \tag{22}$$
- Assn:
$$\phi[t/x](x \leftarrow t)\bar{\phi} \lhd 0 \tag{23}$$
- Rand:
$$(\forall\, m \cdot m \models \phi \Rightarrow \mathcal{P}_i[\![x \xleftarrow{\$} d]\!](m)[\bar{\psi}] \leq \beta) \Rightarrow \phi(x \xleftarrow{\$} d)\bar{\psi} \lhd \beta \tag{24}$$
- Seq:
$$(\phi c \bar{\phi'} \lhd \beta) \wedge (\phi' c' \bar{\phi''} \lhd \beta') \Rightarrow \phi c c' \bar{\phi''} \lhd \beta + \beta' \tag{25}$$
- Cond:
$$(\phi b c \bar{\psi} \lhd \beta) \wedge (\phi \bar{b} c' \bar{\psi} \lhd \beta) \Rightarrow \phi(c +_b c')\bar{\psi} \lhd \beta \tag{28}$$
- Weak:
$$(\phi \leq \phi') \wedge (\phi c \bar{\psi} \lhd \beta) \wedge (\psi' \leq \psi) \wedge (\beta \leq \beta') \Rightarrow \phi' c \bar{\psi'} \lhd \beta' \tag{26}$$
- And:
$$(\phi c \bar{\psi} \lhd \beta) \wedge (\phi c \bar{\psi'} \lhd \beta') \Rightarrow \phi c \psi \bar{\psi'} \lhd \beta + \beta' \tag{29}$$
- Or:
$$(\phi c \bar{\psi} \lhd \beta) \wedge (\phi' c \bar{\psi} \lhd \beta) \Rightarrow (\phi + \phi') c \bar{\psi} \lhd \beta \tag{27}$$
- False:
$$\phi c \bar{\bot} \lhd 1 \tag{30}$$
- While:
$$(\models b_v \in \mathbb{N}) \wedge (\models (\phi \wedge (b_v \leq 0)) \to \bar{b}) \Rightarrow (\phi c \bar{\phi} \lhd \beta) \wedge (\forall_{\eta > 0} \cdot \phi b [b_v = \eta] c \overline{[b_v < \eta]} \lhd 0) \Rightarrow \phi[b_v \leq k] c^{(b)} \overline{\phi \bar{b}} \lhd k\beta \tag{31}$$

Figure 4: Interpretation of aHL rules in aGKAT

The next theorem establishes the main result of the paper.

Theorem 5.1. *All the rules of the system aHL, union bound logic, except* (Assn) *and* (Rand)*, have an aGKAT interpretation (in Fig. 4) that is derivable from the axioms of aGKAT.*

Note that the interpretation of the (**While**) rule, (31) in Fig. 4, is not a plain aGKAT formula, but has some semantic premises. This is because the aHL (**While**) rule itself is expressed with semantic premises. The proof of Theorem 5.1 can be found in Appendix 16. The most interesting cases are (**Seq**) and (**And**) rules, and the most difficult one is that of (**While**).

Observe that an interesting feature of aGKAT is that none of its $\lhd$-axioms (Fig. 3) refers to guarded iteration $c^{(b)}$, and nevertheless aGKAT is as expressive as aHL and allows to derive its (**While**) rule. Another interesting specificity of aGKAT w.r.t. to aHL is that in aGKAT the axioms for reasoning on program equivalence (those of GKAT, Fig. 1) are disjoint from those for reasoning on probabilities (Fig. 3).

6 An example showing that aGKAT is more expressive than aHL

We have shown that aGKAT allows to encode aHL, but in this section we will show that aGKAT is more expressive than aHL, in the sense that in can prove some bounds that aHL cannot.

Example 6.1 (While program)**.** *Let d be the distribution corresponding to a fair dice with three outcomes, that is to say that d has support $\{0,1,2\}$ and $d(0) = d(1) = d(2) = 1/3$.*

We consider the program c below:

$$c \quad = \quad (x \xleftarrow{\$} d) \cdot (x \xleftarrow{\$} d)^{([x=0])} \cdot [x = 1]$$

So c can be described as follows:

it samples d a first time and assigns the result to x; then until it obtains $(x \neq 0)$ it repeats sampling d and assigning the result to x; if at some point it obtains $(x \neq 0)$, then if $(x = 1)$ it terminates successfully, otherwise (that is to say if $(x = 2)$) it aborts.

The analysis of the probability of successful termination of c goes as follows:

with the first sample one obtains $(x = 1)$ with probability $1/3$ and then the program terminates successfully; or one obtains $(x = 2)$ with probability $1/3$ and then the program aborts; or one obtains $(x = 0)$ with probability $1/3$ and we execute c again.

So the probability of successful termination is:

$$\Sigma_{i=1}^{+\infty}(\frac{1}{3})^i = \frac{1}{3} \cdot \frac{1}{1-\frac{1}{3}} = \frac{1}{3} \cdot \frac{3}{2} = \frac{1}{2}$$

Let us now proceed with an analysis in aGKAT. We represent the properties of d with the following 4 axioms:

$$(x \xleftarrow{\$} d)[x = i] \lhd 1/3 \text{ for } i = 0,1,2 \ , (x \xleftarrow{\$} d)[x \neq 0,1,2] \lhd 0.$$

We can then derive the following proof by using GKAT axioms:

$$\begin{array}{lll}
c &= (x \xleftarrow{\$} d) \cdot (x \xleftarrow{\$} d)^{([x=0])} \cdot [x = 1] & \\
&= (x \xleftarrow{\$} d) \cdot ((x \xleftarrow{\$} d) \cdot (x \xleftarrow{\$} d)^{([x=0])} +_{[x=0]} 1) \cdot [x = 1] & \text{by ax. (11)} \\
&= (x \xleftarrow{\$} d) \cdot (1 +_{[x\neq 0]} (x \xleftarrow{\$} d) \cdot (x \xleftarrow{\$} d)^{([x=0])}) \cdot [x = 1] & \text{by ax. (2)} \\
&= (x \xleftarrow{\$} d) \cdot ([x \neq 0] +_{[x\neq 0]} [x = 0] \cdot (x \xleftarrow{\$} d) \cdot (x \xleftarrow{\$} d)^{([x=0])}) \cdot [x = 1] & \text{by ax. (4) and (2)} \\
&= (x \xleftarrow{\$} d) \cdot ([x \neq 0] \cdot [x = 1] +_{[x\neq 0]} [x = 0] \cdot (x \xleftarrow{\$} d) \cdot (x \xleftarrow{\$} d)^{([x=0])} \cdot [x = 1]) & \text{by ax. (5)} \\
&= (x \xleftarrow{\$} d) \cdot ([x = 1] +_{[x\neq 0]} [x = 0] \cdot (x \xleftarrow{\$} d) \cdot (x \xleftarrow{\$} d)^{([x=0])} \cdot [x = 1]) & \text{by properties of tests} \\
&= (x \xleftarrow{\$} d) \cdot ([x = 1] +_{[x\neq 0]} [x = 0] \cdot c) &
\end{array}$$

We know that $(x \xleftarrow{\$} d) \cdot [x = 1] \lhd 1/3$ and $(x \xleftarrow{\$} d) \cdot [x = 0] \lhd 1/3$. By using axioms (19) and (17) we deduce that $(x \xleftarrow{\$} d) \cdot [x = 0] \cdot c \lhd 1/3$.

By applying axiom (16) to the last line of the derivation we obtain $c \lhd 1/3 + 1/3 = 2/3$.

We can then refine this bound by proceeding by recursion. Denote $\beta_0 = 1$ and $\beta_{i+1} = \frac{1}{3} \cdot (1 + \beta_i)$ for $i \geq 0$. That is to say that $\beta_i = \Sigma_{j=1}^{i}(\frac{1}{3})^j + (\frac{1}{3})^i$. Let us prove by recursion on i that one can derive $c \lhd \beta_i$ for any $i \geq 0$.

The property holds for $i = 0$. Let us assume it holds for i. By applying as before axiom (16) to the last line of the derivation and by using the recursion hypothesis we can derive $c \lhd 1/3 + 1/3\beta_i = \beta_{i+1}$. So by recursion we conclude that for any $i \geq 0$ we can derive $c \lhd \beta_i$.

As moreover the sequence (β_i) converges to $\Sigma_{i=1}^{+\infty}(\frac{1}{3})^i = \frac{1}{2}$ we can deduce meta-theoretically that $c \lhd \frac{1}{2}$ (although we cannot derive this limit bound within our system).

However we can verify that one cannot derive in aHL the property $c \lhd \beta_2$, that is to say $c \lhd \frac{5}{9}$. Indeed the aHL judgement corresponding to $c \lhd \frac{5}{9}$ is $\vdash_{5/9} (x \xleftarrow{\$} d) \cdot (x \xleftarrow{\$} d)^{([x=0])} : T \Rightarrow [x \neq 1]$. However in

order to be able to apply a (*While*) rule in aHL one needs to have an integer variable b_v that strictly decreases at each execution of the body of the loop, which is not the case here. So one cannot apply any (*While*) rule, and thus one cannot prove this bound.

This example thus shows that aGKAT is more expressive than aHL, in the sense that it can prove probability bounds that aHL cannot.

7 Example of the Report-noisy-max algorithm

We now consider the example of the *Report-noisy-max* algorithm, which has been analysed in [5] with the logic aHL. Our analysis here using aGKAT will be similar, but the equational approach of aGKAT will simplify some steps. The full proof can be found in Appendix 17.

We consider a finite set $\mathcal{R}$ and a quality score function *qscore*, which takes as input a pair of an element r of $\mathcal{R}$ and a database d, and returns a real number. The goal of the algorithm is to find an element r^* of $\mathcal{R}$ which approximately minimizes the function *qscore* on d. The algorithm is randomized and only computes an approximate minimization because it is designed to satisfy a differential privacy property (see [9]). The algorithm proceeds by computing for each element r of $\mathcal{R}$ the quality score $qscore(r, d)$ and adding to it a Laplacian noise (according to the Laplace mechanism for differential privacy [9]) and returning the element r^* with the highest noisy value.

Here we do not deal with the privacy property of this program, but instead our objective is to study its *accuracy*, that is to say to bound the difference between the value of $qscore(r^*, d)$ and the real minimum of $qscore(\cdot, d)$ on $\mathcal{R}$. The algorithm is encoded in GKAT as the program $c = (flag \leftarrow 1);\ (best \leftarrow 0);\ (\mathcal{R}_0 \leftarrow \mathcal{R});\ (\mathcal{R}' \leftarrow \varnothing);\ c'^{[\mathcal{R}\neq\varnothing]};\ return(r^*)$ where

$$c' = (r \leftarrow pick(\mathcal{R}));\ (noisy[r] \overset{\$}{\leftarrow} \mathcal{L}_{\epsilon/2}(qscore(r, d)));\ (c_1 +_b 1);\ (\mathcal{R} \leftarrow \mathcal{R}\backslash\{r\});\ (\mathcal{R}' \leftarrow \mathcal{R}' \cup \{r\})$$
$$\text{where } c_1 = (flag \leftarrow 0);\ (r^* \leftarrow r);\ (best \leftarrow noisy[r]) \qquad b = (noisy[r] > best) + (flag == 1)$$

The variable *flag* has Boolean values ($\{0, 1\}$), $\mathcal{R}$, $\mathcal{R}_0$ and $\mathcal{R}'$ are sets, r, r^* range over elements of $\mathcal{R}$, $noisy[r]$ and *best* range over reals. The notation $noisy[r]$ is an array-like notation for representing n variables, where n is the size of the set $\mathcal{R}$. Note that variable $\mathcal{R}'$ does not play any role in the algorithm, it is just used to express properties of the execution.

This program uses the following kinds of actions and tests:

- actions for operations on sets: picking an (arbitrary) element r from a set $(r \leftarrow pick(\mathcal{R}))$, removing ($\mathcal{R} \leftarrow \mathcal{R}\backslash\{r\}$)) and adding an element $(\mathcal{R}' \leftarrow \mathcal{R}' \cup \{r\})$,
- sampling from a Laplacian distribution centered in a with parameter p : $(x \overset{\$}{\leftarrow} \mathcal{L}_p(a))$,
- tests: inequalities for reals, equality for Boolean value, comparison to empty set for sets $[\mathcal{R} \neq \varnothing]$; we will also need a finite number of additional tests for expressing properties on the execution, that we will see later.

We recall the following accuracy property of the Laplace distribution [5]:

Lemma 7.1. *Let $\beta \in [0, 1]$, ν a sample from $\mathcal{L}_p(a)$. Then $Pr_{\mathcal{L}_p(a)}[\mid \nu - a \mid > \frac{1}{p}\log(\frac{1}{\beta})] < \beta$.*

Therefore we have $\models (x \overset{\$}{\leftarrow} \mathcal{L}_p(a))[\mid x - a \mid > \frac{1}{p}\log(\frac{1}{\beta})] \triangleleft \beta$. Hence for the sampling in c:

$$\models (noisy[r] \overset{\$}{\leftarrow} \mathcal{L}_{\epsilon/2}(qscore(r, d))) \cdot \overline{b_1} \triangleleft \frac{\beta}{\mid \mathcal{R}_0 \mid} \tag{32}$$

where $\overline{b_1} = [\mid noisy[r] - qscore(r, d) \mid > \frac{2}{\epsilon}\log(\frac{|\mathcal{R}_0|}{\beta})]$.

Now we want to establish a property for the whole program c'. Observe that $\overline{b_1}$ only depends on the values of $noisy[r]$ and $qscore(r, d)$. Moreover $noisy[r]$ and $qscore(r, d)$ are not changed by the last 3 actions of c'. Therefore by applying Prop.4.4.1 we get $c';\ \overline{b_1} \equiv c''$, where c'' is obtained by inserting in c' the test $\overline{b_1}$ just after $(noisy[r] \overset{\$}{\leftarrow} \mathcal{L}_{\epsilon/2}(qscore(r, d)))$. As we know by (19) that for any c_0 we have $\models c_0 \triangleleft 1$, by combining this with axiom (17) and (32) we get $\models c'' \triangleleft \frac{\beta}{|\mathcal{R}_0|}$. This is a step where aGKAT has provided us a concise and simple reasoning. Therefore, as $c';\ \overline{b_1} \equiv c''$, we get by Prop. 4.4.2: $\models c' \cdot \overline{b_1} \triangleleft \frac{\beta}{|\mathcal{R}_0|}$.

We want to prove an invariant for the body c' of the *while* loop in c. For that consider the test b_2 corresponding to the predicate $\phi_2 = \forall\, r \in \mathcal{R}',\ \mid noisy[r] - qscore(r, d) \mid \leq \frac{2}{\epsilon}\log(\frac{|\mathcal{R}_0|}{\beta})$.

$$\text{We have:} \quad b_2 \cdot b_1 \cdot (\mathcal{R}' \leftarrow \mathcal{R}' \cup \{r\}) \cdot \overline{b_2} \quad \equiv \quad 0 \tag{33}$$

Let c_2 be c' deprived of the last action, i.e. $c' = c_2 \cdot (\mathcal{R}' \leftarrow \mathcal{R}' \cup \{r\})$. The reasoning we did on c' before can be repeated for c_2, and so as for c' we get: $\models c_2 \cdot \overline{b_1} \triangleleft \frac{\beta}{|\mathcal{R}_0|}$. Therefore by axioms (17) and (19) we get $\models b_2 \cdot c_2 \cdot \overline{b_1} \triangleleft \frac{\beta}{|\mathcal{R}_0|}$ (here again aGKAT helps us with conciseness). Moreover as c_2 does not modify $\mathcal{R}'$, by using Prop.4.4.1 we get $\models b_2 \cdot c_2 \cdot \overline{b_2} \triangleleft 0$. Thus by using the aGKAT encoding (Theorem 5.1) of the aHL rule (And) we obtain from the two previous statements: $\models b_2 \cdot c_2 \cdot \overline{b_1 \cdot b_2} \triangleleft \frac{\beta}{|\mathcal{R}_0|}$. Equation (33) gives us $\models (b_1 \cdot b_2) \cdot (\mathcal{R}' \leftarrow \mathcal{R}' \cup \{r\}) \cdot \overline{b_2} \triangleleft 0$.

By using the aGKAT encoding of the aHL rule (Seq) we get from the two last statements: $\models b_2 \cdot c' \cdot \overline{b_2} \triangleleft \frac{\beta}{|\mathcal{R}_0|}$. By using the aGKAT encoding of the aHL rule (While) we get from this last statement, since the loop runs for $\mid \mathcal{R}_0 \mid$ iterations, $\models b_2 \cdot c'^{[\mathcal{R} \neq \varnothing]} \cdot \overline{b_2} \triangleleft \beta$.

Finally as $(\mathcal{R}' \leftarrow \varnothing) \cdot \overline{b_2} \equiv 0$ we deduce from this statement using (Seq) that $\models c \cdot \overline{b_2} \triangleleft \beta$. So we have proven using aGKAT that the property corresponding to the following judgement holds: $\vdash_\beta c : \top \Rightarrow \forall\, r \in \mathcal{R}',\ \mid noisy[r] - qscore(r, d) \mid \leq \frac{2}{\epsilon} \log(\frac{|\mathcal{R}_0|}{\beta})$.

In Appendix 17 we continue the proof to finally obtain an accuracy bound for the algorithm.

8 Related work

Several works have explored the use of program logics for the verification of probabilistic programs. Some of these works have explored approaches based on Hoare-like logics [15] while some other ones have developed the approach of weakest-pre-expectations, e.g. [24, 16]. The paper [7] has extended standard Hoare logic to deal with a language containing a probabilistic choice operator, and in which predicates express claims about the state of a probabilistic program. In this work, a semantics for the language is given and a Hoare-style deduction system presented, and proven to be correct w.r.t. the semantics. Another example is the union bound logic aHL [5] that we already presented. More recently, Graded Hoare logic (GHL) was introduced in [10] as a parameterisable framework for extending Hoare logic with a preordered monoidal analysis, with a few examples of applications: the union bound logic aHL [5]; logics for analysis of computation time; or the logic for reasoning about program counter security [25]. Other works even explore extensions of propositional dynamic logic to probabilistic programs, as [22]. This article proposes a probabilistic analog of PDL, which generalises the non-deterministic logical constructs and proof rules in PDL to arithmetic analogs in the probabilistic version.

Other approaches to probabilistic program verification were also introduced in the literature, relying on algebraic structures to wrap the apparatus of logical systems into more elegant frameworks. Some of these approaches are extension of Kleene algebra with tests, of which we give a few examples. A probabilistic extension of GKAT (ProbGKAT) was introduced in [28] for reasoning about imperative programs with probabilistic branching. One difference to our approach is the syntactic introduction of the probability, by the operator $\oplus_r$, where r is a probability; we rather do it semantically, by taking samplings as basic instructions, more in the style of the probabilistic assignment operator of the language `pIMP` [14] for instance. Additionally, [28] provides an non-algebraic axiomatisation for the structure and proves its completeness.

Another KAT extension was given in [11] which aimed at capturing fuzzy programs by replacing the Boolean algebra of KAT by a lattice, with the goal of being able to reason on fuzzy (non Boolean) properties [13].

A variant of GKAT was introduced in [12] as a relational structure to reason about properties of pairs of probabilistic programs (e.g. non-interference between variables), in the style of the system BiKAT [3] in the non-probabilistic setting. Still in the domain of relational reasoning, we can mention relational differential dynamic logic [19], which is specifically designed for the verification of *cyber-physical systems*, in a process that the authors called *synchronizing* the dynamics for comparing two systems.

9 Discussion and future work

We believe that a promising aspect of aGKAT and of the axiomatic presentation we introduced in this paper, is that they can contribute to extend the range of applicability of (co)-algebraic techniques of verification illustrated e.g. in [1, 23, 29, 2] to the realm of approximate reasoning on program effects [5, 6, 10]. This suggests several exciting research directions, which we discuss below.

Towards decision procedures. Recall that the paper [29] has given a decision procedure for the equivalence of GKAT programs whose complexity is almost linear time, assuming that the number of tests is fixed. This procedure is based on a new automata construction. One could investigate in an analogous

way decision problems in aGKAT for statements of the form $c \triangleleft \beta$. The problem could be expressed with some semantic hypothesis, typically some probabilistic assumptions on the randomized primitives used by the program. Our axioms on the relation $\triangleleft$ (Fig. (3)) are quite promising in this respect since they are Horn clauses. Combining automata methods [29] and Horn clauses deduction techniques might lead to some efficient procedures. In [28], the authors present a decision procedure for demonstrating the existence of bisimilarity between two ProbGKAT expressions.

Extension to other effects. In the present paper we restricted ourselves to a specific pod-monoid with specific interval of values and set of operators, which were enough to capture aHL and handle the initial intended goals of reasoning on probabilistic properties. However we would like to push this approach further. By using a generic and external structure to the main algebraic model of programs, we could follow a parametric approach, and obtain more freedom on the structure chosen to capture a wider range of quantitative analysis of effects, like for instance: the analysis of computation time model, by taking the natural numbers and the arithmetic sum as the monoidal composition; the program counter security model, by taking a set of binary values and the string concatenation as the composition; and the union bound logic itself. Those are a few concrete models considered for a generic version of Hoare Logic, analysed in [10].

We want to stress however that it is not trivial to capture in the same generic setting both the union bound logic and the logic for analysis of computation time. The system aGKAT as it stands does not allow to do that. In particular axiom (19) implies that the neutral element of the first monoid is also maximal for the order; this is not the case in the monoid $(\mathbb{N}, +)$ (or even $(\mathbb{N}^\infty, +)$) used for the Hoare logic for analysis of computation time [10].

The framework of pRHL could also benefit from an algebraic approach, calling for a structure taking into account the parametric reasoning about judgments themselves. One would need to embed the parameters into the structure itself, resorting, for example, to a relation between algebraic terms and the elements from the structure which model these parameters

Towards a stronger completeness. Another possible direction for future work would be to study completeness of aGKAT with respect to some class of Horn clauses which embed aHL rules. That would mean to prove that the theory of aGKAT could always derive equations that represent valid aHL rules. We could draw inspiration from Kozen's classical work [21], in which an analogous result was proven for KAT with respect to a class of Horn clauses which embed propositional Hoare logic.

Towards relational properties. In this paper we have considered properties on single executions of a program, but some important questions can be expressed as relational properties on pairs of execution, for instance non-interference, continuity or sensitivity properties. An extension of Kleene algebra with tests called BiKAT for relational properties was introduced in [3] and another framework for probabilistic relational properties was proposed in [12]. It would interesting to explore if the approximation construction we defined in the present paper could be applied to the probabilistic relational setting of [12]. This would be analogous to the move in the relational Hoare logic setting, from pRHL to apRHL.

Non-determinism and probabilities. One of the advantages of the syntactical restriction of GKAT is to facilitate the inclusion of probabilistic models, by neglecting nondeterminism. While usually avoided, and always difficult, one possible direction for future work could be to consider a language with both nondeterminism and probabilities, capturing more application scenarios.

References

[1] D. Kozen A. Angus. Kleene algebra with tests and program schematology. Technical report, Computer Science Department, Cornell University, Ithaca, NY 14853-7501, USA, July 2001. Technical Report TR2001-1844.

[2] Carolyn Jane Anderson, Nate Foster, Arjun Guha, Jean-Baptiste Jeannin, Dexter Kozen, Cole Schlesinger, and David Walker. NetKAT: semantic foundations for networks. In Suresh Jagannathan and Peter Sewell, editors, *The 41st Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL '14, San Diego, CA, USA, January 20-21, 2014*, pages 113–126. ACM, 2014. doi:10.1145/2535838.2535862.

[3] Timos Antonopoulos, Eric Koskinen, Ton Chanh Le, Ramana Nagasamudram, David A. Naumann, and Minh Ngo. An algebra of alignment for relational verification. *CoRR*, abs/2202.04278, 2022. URL: https://arxiv.org/abs/2202.04278, arXiv:2202.04278.

[4] Gilles Barthe, François Dupressoir, Benjamin Grégoire, César Kunz, Benedikt Schmidt, and Pierre-Yves Strub. Easycrypt: A tutorial. In Alessandro Aldini, Javier López, and Fabio Martinelli, editors, *Foundations of Security Analysis and Design VII - FOSAD 2012/2013 Tutorial Lectures*, volume 8604 of *Lecture Notes in Computer Science*, pages 146–166. Springer, 2013. doi:10.1007/978-3-319-10082-1_6.

[5] Gilles Barthe, Marco Gaboardi, Benjamin Grégoire, Justin Hsu, and Pierre-Yves Strub. A program logic for union bounds. In Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi, editors, *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, volume 55 of *LIPIcs*, pages 107:1–107:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPIcs.ICALP.2016.107.

[6] Gilles Barthe, Boris Köpf, Federico Olmedo, and Santiago Zanella Béguelin. Probabilistic relational reasoning for differential privacy. *ACM Trans. Program. Lang. Syst.*, 35(3):9:1–9:49, 2013. doi:10.1145/2492061.

[7] Jerry den Hartog and Erik P. de Vink. Verifying probabilistic programs using a Hoare like logic. *Int. J. Found. Comput. Sci.*, 13(3):315–340, 2002. URL: http://dx.doi.org/10.1142/S012905410200114X, doi:10.1142/S012905410200114X.

[8] Easycrypt development team. Easycrypt, 2024. URL: https://formosa-crypto.org/projects/.

[9] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014. doi:10.1561/0400000042.

[10] Marco Gaboardi, Shin-ya Katsumata, Dominic Orchard, and Tetsuya Sato. Graded hoare logic and its categorical semantics. In Nobuko Yoshida, editor, *Programming Languages and Systems - 30th European Symposium on Programming, ESOP 2021, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2021, Luxembourg City, Luxembourg, March 27 - April 1, 2021, Proceedings*, volume 12648 of *Lecture Notes in Computer Science*, pages 234–263. Springer, 2021. doi:10.1007/978-3-030-72019-3_9.

[11] L. Gomes, A. Madeira, and L. S. Barbosa. Generalising KAT to verify weighted computations. *Scient. Annals of Comp. Sc.*, 29(2):141–184, 2019. doi:10.7561/SACS.2019.2.141.

[12] Leandro Gomes, Patrick Baillot, and Marco Gaboardi. BiGKAT: an algebraic framework for relational verification of probabilistic programs. working paper or preprint, March 2023. URL: https://hal.science/hal-04017128.

[13] Leandro Gomes, Alexandre Madeira, and Luís Soares Barbosa. A semantics and a logic for fuzzy arden syntax. *Soft Comput.*, 25(9):6789–6805, 2021. doi:10.1007/s00500-021-05593-9.

[14] Ichiro Hasuo, Yuichiro Oyabu, Clovis Eberhart, Kohei Suenaga, Kenta Cho, and Shin-ya Katsumata. Control-data separation and logical condition propagation for efficient inference on probabilistic programs. *J. Log. Algebraic Methods Program.*, 136:100922, 2024. doi:10.1016/J.JLAMP.2023.100922.

[15] Claire Jones. *Probabilistic non-determinism*. PhD thesis, University of Edinburgh, UK, 1990. URL: https://hdl.handle.net/1842/413.

[16] Benjamin Lucien Kaminski, Joost-Pieter Katoen, Christoph Matheja, and Federico Olmedo. Weakest precondition reasoning for expected runtimes of randomized algorithms. *J. ACM*, 65(5):30:1–30:68, 2018. doi:10.1145/3208102.

[17] Tobias Kappé, Paul Brunet, Alexandra Silva, Jana Wagemaker, and Fabio Zanasi. Concurrent kleene algebra with observations: From hypotheses to completeness. In *Proceedings of FOSSACS 2020*, volume 12077 of *LNCS*, pages 381–400. Springer, 2020. doi:10.1007/978-3-030-45231-5_20.

[18] Michael J. Kearns and Umesh V. Vazirani. *An Introduction to Computational Learning Theory*. MIT Press, 1994. URL: https://mitpress.mit.edu/books/introduction-computational-learning-theory.

[19] Juraj Kolcák, Jérémy Dubut, Ichiro Hasuo, Shin-ya Katsumata, David Sprunger, and Akihisa Yamada. Relational differential dynamic logic. In Armin Biere and David Parker, editors, *Tools and Algorithms for the Construction and Analysis of Systems - 26th International Conference, TACAS 2020, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2020, Dublin, Ireland, April 25-30, 2020, Proceedings, Part I*, volume 12078 of *Lecture Notes in Computer Science*, pages 191–208. Springer, 2020. doi:10.1007/978-3-030-45190-5_11.

[20] D. Kozen. Kleene algebra with tests. *ACM Trans. on Prog. Lang. and Systems*, 19(3):427–443, 1997. doi:10.1145/256167.256195.

[21] D. Kozen. On Hoare logic and Kleene algebra with tests. *ACM Trans. on Comp. Logic*, 1(212):1–14, 2000. doi:10.1109/LICS.1999.782610.

[22] Dexter Kozen. A probabilistic PDL. *J. Comput. Syst. Sci.*, 30(2):162–178, 1985. doi:10.1016/0022-0000(85)90012-1.

[23] Dexter Kozen and Maria-Christina Patron. Certification of compiler optimizations using Kleene algebra with tests. In John W. Lloyd, Verónica Dahl, Ulrich Furbach, Manfred Kerber, Kung-Kiu Lau, Catuscia Palamidessi, Luís Moniz Pereira, Yehoshua Sagiv, and Peter J. Stuckey, editors, *Computational Logic - CL 2000, First International Conference, London, UK, 24-28 July, 2000, Proceedings*, volume 1861 of *Lecture Notes in Computer Science*, pages 568–582. Springer, 2000. doi:10.1007/3-540-44957-4_38.

[24] Annabelle McIver and Carroll Morgan. *Abstraction, Refinement and Proof for Probabilistic Systems*. Monographs in Computer Science. Springer, 2005. URL: https://doi.org/10.1007/b138392, doi:10.1007/B138392.

[25] David Molnar, Matt Piotrowski, David Schultz, and David A. Wagner. The program counter security model: Automatic detection and removal of control-flow side channel attacks. In Dongho Won and Seungjoo Kim, editors, *Information Security and Cryptology - ICISC 2005, 8th International Conference, Seoul, Korea, December 1-2, 2005, Revised Selected Papers*, volume 3935 of *Lecture Notes in Computer Science*, pages 156–168. Springer, 2005. doi:10.1007/11734727_14.

[26] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995. doi:10.1017/cbo9780511814075.

[27] Damien Pous. Kleene algebra with tests and coq tools for while programs. In Sandrine Blazy, Christine Paulin-Mohring, and David Pichardie, editors, *Interactive Theorem Proving - 4th International Conference, ITP 2013, Rennes, France, July 22-26, 2013. Proceedings*, volume 7998 of *Lecture Notes in Computer Science*, pages 180–196. Springer, 2013. doi:10.1007/978-3-642-39634-2_15.

[28] Wojciech Rozowski, Tobias Kappé, Dexter Kozen, Todd Schmid, and Alexandra Silva. Probabilistic guarded KAT modulo bisimilarity: Completeness and complexity. In Kousha Etessami, Uriel Feige, and Gabriele Puppis, editors, *50th International Colloquium on Automata, Languages, and Programming, ICALP 2023, July 10-14, 2023, Paderborn, Germany*, volume 261 of *LIPIcs*, pages 136:1–136:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.ICALP.2023.136.

[29] Steffen Smolka, Nate Foster, Justin Hsu, Tobias Kappé, Dexter Kozen, and Alexandra Silva. Guarded Kleene algebra with tests: verification of uninterpreted programs in nearly linear time. *Proc. ACM Program. Lang.*, 4(POPL):61:1–61:28, 2020. doi:10.1145/3371129.

[30] Jana Wagemaker, Nate Foster, Tobias Kappé, Dexter Kozen, Jurriaan Rot, and Alexandra Silva. Concurrent NetKAT - modeling and analyzing stateful, concurrent networks. In Ilya Sergey, editor, *Programming Languages and Systems - 31st European Symposium on Programming, ESOP 2022, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2022, Munich, Germany, April 2-7, 2022, Proceedings*, volume 13240 of *Lecture Notes in Computer Science*, pages 575–602. Springer, 2022. doi:10.1007/978-3-030-99336-8_21.

[31] Cheng Zhang, Arthur Azevedo de Amorim, and Marco Gaboardi. On incorrectness logic and Kleene algebra with top and tests. *Proc. ACM Program. Lang.*, 6(POPL):1–30, 2022. doi:10.1145/3498690.

Appendix

A Derivable GKAT facts

$$c_1 +_{b_1} (c_2 +_{b_2} c_3) = (c_1 +_{b_1} c_2) +_{b_1+b_2} c_3 \quad (34)$$
$$c_1 +_b c_2 = c_1 +_b \neg\, b \cdot c_2 \quad (35)$$
$$b_1 \cdot (c_1 +_{b_2} c_2) = b_1 c_1 +_{b_2} b_1 c_2 \quad (36)$$
$$c +_b 0 = bc \quad (37)$$
$$c_1 +_0 c_2 = c_2 \quad (38)$$
$$b \cdot (c_1 +_b c_2) = bc_1 \quad (39)$$
$$c^{(b)} = c^{(b)} \cdot \neg\, b \quad (40)$$
$$c^{(b)} = (bc)^{(b)} \quad (41)$$
$$c^{(0)} = 1 \quad (42)$$
$$c^{(1)} = 0 \quad (43)$$
$$b_1^{(b_2)} = \neg\, b_2 \quad (44)$$
$$c^{(b_2)} = c^{(b_1 b_2)} \cdot c^{(b_2)} \quad (45)$$

Figure 5: Derivable GKAT facts

B Proof of Proposition 2.1

$$\begin{aligned}
& b \cdot b_1 + \neg\, b \cdot b_2 \\
= & \quad \{ (1) \} \\
& (b \cdot b_1 + \neg\, b \cdot b_2) +_b (b \cdot b_1 + \neg\, b \cdot b_2) \\
= & \quad \{ (4) \text{ and } (35) \} \\
& b \cdot (b \cdot b_1 + \neg\, b \cdot b_2) +_b \neg\, b \cdot (b \cdot b_1 + \neg\, b \cdot b_2) \\
= & \quad \{ 36 \} \\
& (b \cdot b \cdot b_1 + b \cdot \neg\, b \cdot b_2) +_b (\neg\, b \cdot b \cdot b_1 + \neg\, b \cdot \neg\, b \cdot b_2) \\
= & \quad \{ \text{B.A. and } (7) \} \\
& (b \cdot b_1 + 0) +_b (0 + \neg\, b \cdot b_2) \\
= & \quad \{ \text{B.A.} \} \\
& (b \cdot b_1) +_b (\neg\, b \cdot b_2) \\
= & \quad \{ (4) \text{ and } (35) \} \\
& b_1 +_b b_2
\end{aligned}$$

C Proof of Proposition 2.2

Because the tests form a Boolean algebra we have:

$$b_1 = b_1 + b_1 b_2 \quad \text{(absorption law)}$$

Now we have:

$$\begin{aligned}
b_1 +_{b_1} b_2 &= b_1 b_1 + \overline{b_1} b_2 \text{ by Prop. 2.1} \\
&= b_1 + \overline{b_1} b_2 \text{ by idempotency} \\
&= (b_1 + b_1 b_2) + \overline{b_1} b_2 \text{ by the equation above} \\
&= b_1 + (b_1 + \overline{b_1}) b_2 \\
&= b_1 + b_2
\end{aligned}$$

D Proof of Proposition 4.1

The fact that the probabilistic interpretation satisfies the axioms of GKAT is known from [29]. It remains to prove that the interpretation i satisfies the axioms of Fig. 3.

Consider a set of states S. Axioms (14) and (15) are trivial.

Axiom (17):

By assumption we know that:

$$\forall_{s\in S}, \sum_{s'\in S} \mathcal{P}_i[\![c_1]\!](s)(s') \leq \beta_1, \;\; \forall_{s\in S}, \sum_{s'\in S} \mathcal{P}_i[\![c_2]\!](s)(s') \leq \beta_2$$

So we have:

$$\begin{array}{rcl}
\forall_{s\in S}, \sum_{s'\in S} \mathcal{P}_i[\![c_1 \cdot c_2]\!](s)(s') & = & \sum_{s'\in S}\sum_{s''\in S} \mathcal{P}_i[\![c_1]\!](s)(s'') \cdot \mathcal{P}_i[\![c_2]\!](s'')(s') \\
& = & \sum_{s''\in S}\sum_{s'\in S} \mathcal{P}_i[\![c_1]\!](s)(s'') \cdot \mathcal{P}_i[\![c_2]\!](s'')(s') \\
& & \text{by commutativity of } + \\
& = & \sum_{s''\in S} (\mathcal{P}_i[\![c_1]\!](s)(s'') \cdot \sum_{s'\in S} \mathcal{P}_i[\![c_2]\!](s'')(s')) \\
& & \text{by distributivity of . over } + \\
& \leq & \sum_{s''\in S} \mathcal{P}_i[\![c_1]\!](s)(s'') \cdot \beta_2 \text{ by assumptions} \\
& \leq & \beta_1 \cdot \beta_2 \text{ by assumptions}
\end{array}$$

We then conclude, by Definition 4.3, $c_1 \cdot c_2 \triangleleft \beta_1 \cdot \beta_2$.

Axiom (18):

Assumptions: $c_1 \triangleleft \beta \Leftrightarrow \forall_{s\in S}, \sum_{s'\in S} \mathcal{P}_i[\![c_1]\!](s)(s') \leq \beta$ and $c_2 \triangleleft \beta \Leftrightarrow \forall_{s\in S}, \sum_{s'\in S} \mathcal{P}_i[\![c_2]\!](s)(s') \leq \beta$.

Goal: $c_1 +_b c_2 \triangleleft \beta \Leftrightarrow \forall_{s\in S}, \sum_{s'\in S} \mathcal{P}_i[\![c_1 +_b c_2]\!](s)(s') \leq \beta$

$$\sum_{s'\in S} \mathcal{P}_i[\![c_1 +_b c_2]\!](s)(s') = \begin{cases} \sum_{s'\in S} \mathcal{P}_i[\![c_1]\!](s)(s') & \text{if } s \in sat^\dagger(b) \\ \sum_{s'\in S} \mathcal{P}_i[\![c_2]\!](s)(s') & \text{if } s \in sat^\dagger(\neg\, b) \end{cases}$$

By assumptions, $\forall_{s\in S}, \sum_{s'\in S} \mathcal{P}_i[\![c_1]\!](s)(s') \leq \beta$ and $\sum_{s'\in S} \mathcal{P}_i[\![c_2]\!](s)(s') \leq \beta$.

Hence we conclude $\sum_{s'\in S} \mathcal{P}_i[\![c_1 +_b c_2]\!](s)(s') \leq \beta$, i.e. by Definition 4.3 $c_1 +_b c_2 \triangleleft \beta$.

Axiom (16):

By assumption we know that:

$\forall_{s\in S}, \sum_{s'\in S} \mathcal{P}_i[\![cc_1]\!](s)(s') \leq \beta_1$ and $\forall_{s\in S}, \sum_{s'\in S} \mathcal{P}_i[\![cc_2]\!](s)(s') \leq \beta_2$.

Now, we have:

$$\begin{array}{rl}
& \forall_{s\in S}, \sum_{s'\in S} \mathcal{P}_i[\![c(c_1 +_b c_2)]\!](s)(s') \\
= & \{ \text{ Definition 2.1}\} \\
& \sum_{s'\in S}\sum_{s''\in S} \mathcal{P}_i[\![c]\!](s)(s'') \cdot \mathcal{P}_i[\![c_1 +_b c_2]\!](s'')(s') \\
= & \{ \text{ commutativity of } +\} \\
& \sum_{s''\in S} \mathcal{P}_i[\![c]\!](s)(s'') \cdot \sum_{s'\in S} \mathcal{P}_i[\![c_1 +_b c_2]\!](s'')(s') \\
= & \{ \text{ (Definition 2.1)}\} \\
& \sum_{s''\in S} \mathcal{P}_i[\![c]\!](s)(s'') \cdot (\sum_{s'\in sat^\dagger(b)} \mathcal{P}_i[\![c_1]\!](s'')(s') + \sum_{s'\in sat^\dagger(\neg\, b)} \mathcal{P}_i[\![c_2]\!](s'')(s')) \\
= & \{ \text{ distributivity of } \cdot \text{ over } +\} \\
& \sum_{s''\in S} \mathcal{P}_i[\![c]\!](s)(s'') \cdot \sum_{s'\in sat^\dagger(b)} \mathcal{P}_i[\![c_1]\!](s'')(s') \\
+ & \sum_{s''\in S} \mathcal{P}_i[\![c]\!](s)(s'') \cdot \sum_{s'\in sat^\dagger(\neg\, b)} \mathcal{P}_i[\![c_2]\!](s'')(s')
\end{array}$$

$$
\begin{aligned}
&= && \{ \text{ commutativity of } + \} \\
&\quad \sum_{s' \in sat^\dagger(b)} \sum_{s'' \in S} \mathcal{P}_i[\![c]\!](s)(s'') \cdot \mathcal{P}_i[\![c_1]\!](s'')(s') \\
&+ \sum_{s' \in sat^\dagger(\neg\, b)} \sum_{s'' \in S} \mathcal{P}_i[\![c]\!](s)(s'') \cdot \mathcal{P}_i[\![c_2]\!](s'')(s') \\
&= && \{ \text{ Definition 2.1} \} \\
&\quad \sum_{s' \in sat^\dagger(b)} \mathcal{P}_i[\![cc_1]\!](s)(s') + \sum_{s' \in sat^\dagger(\neg\, b)} \mathcal{P}_i[\![cc_2]\!](s)(s') \\
&\leq && \{ \text{ assumptions} \} \\
&\quad \beta_1 + \beta_2
\end{aligned}
$$

Hence we have proven that $c(c_1 +_b c_2) \triangleleft \beta_1 + \beta_2$ holds in the model.
The validity of Axiom (19) comes directly from Definition 4.3.
Axiom (20):
$\forall_{s \in S}, \sum_{s' \in S} \mathcal{P}_i[\![0]\!](s)(s') = 0 \leq 0$. Hence, by Definition 4.3, $0 \triangleleft 0$.

E Counter-example for Remark 4.2 (Invalid statement in the probabilistic model)

Let us take as set of states $\texttt{State} = \{s_1, s_2\}$. We use a matrix notation for Markov kernels: the matrix $\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$ represents the Markov kernel f such that, for all $i, j \in \{1, 2\}$, $f(s_i)(s_j) = a_{i,j}$. Recall that intuitively this means that the execution of the program (represented by f) on initial state s_i has probability $a_{i,j}$ to terminate on state s_j. So in particular the sum over each line is less than or equal to 1. We abusively denote: $f = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$.

Let us define:

$$
c = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}
$$

$$
c_1 = \begin{pmatrix} 1/2 & 1/2 \\ 0 & 0 \end{pmatrix}
$$

$$
c_2 = \begin{pmatrix} 0 & 0 \\ 1/2 & 1/2 \end{pmatrix}
$$

Then we have:

$$
c \cdot c_1 = c \cdot c_2 = \begin{pmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}
$$

So we have: $c \cdot c_1 \triangleleft 1/2$ and $c \cdot c_2 \triangleleft 1/2$, because the sum over each line is equal to $1/2$.

Let us take for b the test on states s defined by $b = [s = s_1]$. Then we have:

$$
c_1 +_b c_2 = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}
$$

So:

$$
c \cdot (c_1 +_b c_2) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}
$$

So $c \cdot (c_1 +_b c_2) \triangleleft 1 (= \frac{1}{2} + \frac{1}{2})$ holds, Axiom (16) is satisfied. And the minimum β which satisfies $c \cdot (c_1 +_b c_2) \triangleleft \beta$ is 1, thus the candidate statement of Remark 4.2 is not valid.

F Proof in the double tossing Example 4.1

We give here a step-by-step explicit proof of the proof in Example 4.1.

$c \cdot (y = 0) = (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (c_1' +_{(x=1)} y \leftarrow 0(y = 0))$ by Axiom (5) (46)

$c \cdot (y = 0) = (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot ((x = 1) + (x \neq 1)) \cdot (c_1' +_{(x=1)} y \leftarrow 0(y = 0))$ by boolean properties (47)

$(x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x \neq 1) \cdot (c_1' +_{(x=1)} y \leftarrow 0(y = 0)) = (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x \neq 1) \cdot (y \leftarrow 0) \cdot (y = 0)$ (48)
by fact (39) in Appendix 10

$(x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot (c_1' +_{(x=1)} y \leftarrow 0(y = 0)) = (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot c_1'$ by fact (39) (49)

$(x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot c_1' =$
$(x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot ((x = 1) + (x \neq 1)) \cdot (y \leftarrow 1(y = 0) +_{(x=1)} y \leftarrow 0(y = 0))$ (50)
by boolean properties

$(x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot (y \leftarrow 1(y = 0) +_{(x=1)} y \leftarrow 0(y = 0)) =$
$(x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot y \leftarrow 1(y = 0)$ by fact (39) (51)

$(x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x \neq 1) \cdot (y \leftarrow 1(y = 0) +_{(x=1)} y \leftarrow 0(y = 0)) =$
$(x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x \neq 1) \cdot y \leftarrow 0(y = 0)$ by fact (39) (52)

$y \leftarrow 1(y = 0) \equiv 0$ by the model (53)

$\models c_2 = (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot (y \leftarrow 1) \cdot (y = 0) \triangleleft 0$ by (53) and axiom (17) (54)

$\models (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x \neq 1) \triangleleft 1/2$ by the choice of interpretation i (55)

$\models c_4 = (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x \neq 1) \cdot (y \leftarrow 0) \cdot (y = 0) \triangleleft 1/2$ by (55), axioms (19) and (17) (56)

$\models (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x \neq 1) \cdot (c_1' +_{(x=1)} y \leftarrow 0(y = 0)) \triangleleft 1/2$ by (48) and (56) (57)

$\models (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \triangleleft 1/2$ by the choice of interpretation i (58)

$\models c_3 = (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x \neq 1) \cdot (y \leftarrow 0) \cdot (y = 0) \triangleleft 1/4$ by (55), (58), axioms (19) and (17) (59)

$\models (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot c_1' \triangleleft 1/4$ by (54), (59), (51), (52) and Prop. 4.3 (60)

$\models (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot (x = 1) \cdot (c_1' +_{(x=1)} y \leftarrow 0(y = 0)) \triangleleft 1/4$ by (49) and (60) (61)

$\models (x \overset{\$}{\leftarrow} \mathit{Coin}) \cdot ((x = 1) + (x \neq 1)) \cdot (c_1' +_{(x=1)} y \leftarrow 0(y = 0)) \triangleleft 3/4$ by (57), (61) and Prop. 4.3 (62)

$\models c \cdot (y = 0) \triangleleft 3/4$ by (47) and (62) (63)

G Proof of Theorem 5.1

- **Skip**:
  Goal: $\phi 1 \neg\, \phi \triangleleft 0$

$$\begin{aligned} & \phi 1 \neg\, \phi \\ = & \quad \{ \text{(9)} \} \\ & \phi \neg\, \phi \\ = & \quad \{ \text{B.A.} \} \\ & 0 \end{aligned}$$

  By axiom (20), $0 \triangleleft 0$.

- **Seq**:
  Assumptions:

$$\phi c \neg\, \phi' \triangleleft \beta \tag{64}$$

and

$$\phi' c' \neg\, \phi'' \lhd \beta' \tag{65}$$

Goal: $\phi c c' \neg\, \phi'' \lhd \beta + \beta'$

$$\begin{array}{ll} & \phi c c' \neg\, \phi'' \\ = & \quad \{\text{ B.A.}\} \\ & \phi c 1 c' \neg\, \phi'' \\ = & \quad \{\text{ Proposition 2.1 and B.A.}\} \\ & \phi c (\phi' +_{\phi'} \neg\, \phi') c' \neg\, \phi'' \\ = & \quad \{\text{ (5)}\} \\ & \phi c (\phi' c' \neg\, \phi'' +_{\phi'} \neg\, \phi' c' \neg\, \phi'') \end{array}$$

We have $\phi c(\phi' c' \neg\, \phi'') \lhd \beta'$, by axiom (17) because $\phi c \lhd 1$ (axiom 19) and (65).

Additionally, we have $\phi c(\neg\, \phi' c' \neg\, \phi'') = (\phi c \neg\, \phi') c' \neg\, \phi'' \lhd \beta$, by axiom (17) because (64) and $c' \neg\, \phi'' \lhd 1$ (axiom 19).

Hence, by axiom (16), $\phi c(\phi' c' \neg\, \phi'' +_{\phi'} \neg\, \phi' c' \neg\, \phi'') \lhd \beta + \beta'$, so $\phi c c' \neg\, \phi'' \lhd \beta + \beta'$.

- **Cond**:

  Assumptions:

$$\phi b c \neg\, \psi \lhd \beta \tag{66}$$

  and

$$\phi \neg\, b c' \neg\, \psi \lhd \beta \tag{67}$$

  Goal: $\phi(c +_b c') \neg\, \psi \lhd \beta$

  By axiom (5) and fact (36), $\phi(c +_b c') \neg\, \psi = \phi c \neg\, \psi +_b \phi c' \neg\, \psi$ and by (66), (67) and axiom (18), we have $\phi c \neg\, \psi +_b \phi c' \neg\, \psi \lhd \beta$.

- **Weak**:

  Assumptions:

$$\phi c \neg\, \psi \lhd \beta \tag{68}$$

$$\phi' \phi = \phi' \tag{69}$$

$$\psi \psi' = \psi \tag{70}$$

  and

$$\beta \leq \beta' \tag{71}$$

  Goal: $\phi' c \neg\, \psi' \lhd \beta'$

  By Boolean algebra, we derive

$$\psi \psi' = \psi \Leftrightarrow \neg\, \psi' \neg\, \psi = \neg\, \psi' \tag{72}$$

  Hence, we reason,

$$\begin{array}{ll} & \phi' c \neg\, \psi' \\ = & \quad \{\text{ (69) and (73)}\} \\ & \phi' \phi c \neg\, \psi' \neg\, \psi \\ = & \quad \{\text{ (B.A.)}\} \\ & \phi' (\phi c \neg\, \psi) \neg\, \psi' \end{array}$$

By axiom (19), $\phi' \lhd 1$, $\neg\ \psi' \lhd 1$. So, by (68) and axiom (17), we have $\phi'(\phi c\neg\ \psi)\neg\ \psi' \lhd 1 \cdot \beta \cdot 1 = \beta$. By (71) and axiom (15), we conclude $\phi'(\phi c\neg\ \psi)\neg\ \psi' \lhd \beta'$.

- **And**:

  Assumptions:

  $$\phi c\neg\ \psi \lhd \beta \tag{73}$$

  and

  $$\phi' c\neg\ \psi' \lhd \beta' \tag{74}$$

  Goal: $\phi c\neg\ (\psi\psi') \lhd \beta + \beta'$

  By Boolean algebra and Proposition 2.1, we reason $\phi c\neg\ (\psi\psi') = \phi c(\neg\ \psi + \neg\ \psi') = \phi c(\neg\ \psi +_{\neg\ \psi} \neg\ \psi')$ and by premises (73) and (74), and by axiom (16) we deduce
  $\phi c(\neg\ \psi +_{\neg\ \psi} \neg\ \psi') \lhd \beta + \beta'$
  and by axiom (14) we have $\phi c\neg\ (\psi\psi') \lhd \beta + \beta'$.

- **Or**:

  Assumptions:

  $$\phi c\neg\ \psi \lhd \beta \tag{75}$$

  and

  $$\phi' c\neg\ \psi \lhd \beta \tag{76}$$

  Goal: $(\phi + \phi')c\neg\ \psi \lhd \beta$

  $$\begin{array}{ll} & (\phi + \phi')c\neg\ \psi \\ = & \quad \{\ (4)\} \\ & (1 +_\phi \phi')c\neg\ \psi \\ = & \quad \{\ (5)\} \\ & c\neg\ \psi +_\phi \phi' c\neg\ \psi \\ = & \quad \{\ (4)\} \\ & \phi c\neg\ \psi +_\phi \phi' c\neg\ \psi \end{array}$$

  By premises (75), (76) and axiom (18), $\phi c\neg\ \psi +_\phi \phi' c\neg\ \psi \lhd \beta$.

- **False**:

  $\phi c\neg\ 0 = \phi c1 = \phi c \lhd 1$ by axioms (19), (17) and (14).

- **While**. We assume that the left-hand side of the implications in (31) hold and we want to prove for any integer k the following formula: $\phi[b_v \leq k]c^{(b)}\neg\ (\phi\neg\ b) \lhd k\beta$. Note in particular that by assumption, b_v only takes integer values.

  Let us first show that we can simplify the expression in the property to prove:

  $$\begin{array}{rcl} \phi[b_v \leq k]c^{(b)}\neg\ (\phi\neg\ b) & = & \phi[b_v \leq k]c^{(b)}(\neg\ \phi + b) \\ & = & \phi[b_v \leq k]c^{(b)}\neg\ b(\neg\ \phi + b),\ \text{by (40)} \\ & = & \phi[b_v \leq k]c^{(b)}(\neg\ b\neg\ \phi + 0) \\ & = & \phi[b_v \leq k]c^{(b)}\neg\ b\neg\ \phi \\ & = & \phi[b_v \leq k]c^{(b)}\neg\ \phi,\ \text{by (40) again.} \end{array}$$

Therefore we only need to prove for any integer k the following property, which we will do by induction on k:

$$IH(k): \qquad \phi[b_v \leq k]c^{(b)}\neg\,\phi \triangleleft k\beta \tag{77}$$

– Case $k = 0$:
As $\models \phi \wedge b_v \leq 0 \rightarrow \neg\, b$ we have:

$$\phi[b_v \leq 0] = \phi[b_v \leq 0]\neg\, b.$$

Moreover $\neg\, bc^{(b)} = \neg\, b$, by Axiom (11).
Therefore:

$$\begin{array}{l}\phi[b_v \leq 0]c^{(b)}\neg\,\phi = \\ \phi[b_v \leq 0]\neg\, bc^{(b)}\neg\,\phi = \\ \phi[b_v \leq 0]\neg\, b\neg\,\phi = \\ \phi\neg\,\phi[b_v \leq 0]\neg\, b = 0 \triangleleft 0\end{array}$$

So $IH(0)$ holds.

– Now, assume $IH(k)$ holds and let us prove $IH(k+1)$.
As b_v takes integer values we have:

$$[b_v \leq k+1] = [b_v \leq k] + [b_v = k+1].$$

So:

$$\begin{array}{l}\phi[b_v \leq k+1]c^{(b)}\neg\,\phi = \\ \phi([b_v \leq k] + [b_v = k+1])c^{(b)}\neg\,\phi = \text{ by Prop. 2.2} \\ \phi([b_v \leq k] +_{[b_v \leq k]} [b_v = k+1])c^{(b)}\neg\,\phi = \\ \phi[b_v \leq k]c^{(b)}\neg\,\phi +_{[b_v \leq k]} \phi[b_v = k+1]c^{(b)}\neg\,\phi\end{array}$$

We know by $IH(k)$ that: $\phi[b_v \leq k]c^{(b)}\neg\,\phi \triangleleft k\beta$.
Assume for the moment the following lemma:

Lemma G.1.

$$\phi[b_v = k+1]c^{(b)}\neg\,\phi \triangleleft (k+1)\beta$$

Then by using axioms (15) and (17) we obtain:

$$\phi[b_v \leq k+1]c^{(b)}\neg\,\phi \triangleleft (k+1)\beta$$

So we have proved $IH(k+1)$.
We now thus only have to prove Lemma 16.1.
By assumption we have:

$$\forall\,\eta > 0, \qquad \phi b[b_v = \eta]c[b_v \geq \eta] \triangleleft 0 \tag{78}$$

We obtain by Axiom (11):

$$bc^{(b)} = bcc^{(b)} \tag{79}$$

$$\neg\, bc^{(b)} = \neg\, b \tag{80}$$

Moreover, since by assumption, b_v only takes values in $\mathbb{N}$, we have:

$$1 = [b_v \geq k+1] + [b_v \leq k] \tag{81}$$

$$= [b_v \geq k+1] +_{[b_v \geq k+1]} [b_v \leq k], \text{ by Prop. 2.2} \tag{82}$$

Let us now prove the statement of Lemma 16.1:

$$\begin{array}{l}\phi[b_v = k+1]c^{(b)}\neg\,\phi = \\ \phi[b_v = k+1](b + \neg\, b)c^{(b)}\neg\,\phi = \text{ by Prop. 2.2} \\ \phi[b_v = k+1](b +_b \neg\, b)c^{(b)}\neg\,\phi = \text{ by Ax. (5)} \\ \phi[b_v = k+1](bc^{(b)}\neg\,\phi +_b \neg\, bc^{(b)}\neg\,\phi) = \text{ by (80)} \\ \phi[b_v = k+1](bc^{(b)}\neg\,\phi +_b \neg\,\phi) = \text{ by (36)} \\ [b_v = k+1](\phi bc^{(b)}\neg\,\phi +_b \phi\neg\,\phi) = \\ [b_v = k+1](\phi bc^{(b)}\neg\,\phi +_b 0) = \\ [b_v = k+1]\phi bc^{(b)}\neg\,\phi = \text{ by (79)} \\ \phi[b_v = k+1]bcc^{(b)}\neg\,\phi, = \\ \phi b[b_v = k+1]cc^{(b)}\neg\,\phi = \\ \phi b[b_v = k+1]c([b_v \geq k+1] +_{[b_v \geq k+1]} [b_v \leq k])c^{(b)}\neg\,\phi = \\ \phi b[b_v = k+1]c([b_v \geq k+1]c^{(b)}\neg\,\phi +_{[b_v \geq k+1]} [b_v \leq k]c^{(b)}\neg\,\phi)\end{array}$$

In order to bound with $\triangleleft$ the expression above, it is sufficient by Axiom (16) to bound with $\triangleleft$ respectively the two following expressions:

$$\phi b[b_v = k+1]c[b_v \geq k+1]c^{(b)}\neg\,\phi$$
$$\phi b[b_v = k+1]c[b_v \leq k]c^{(b)}\neg\,\phi$$

For the first expression we can use (78) for $\eta = k+1$ and we obtain:

$$\phi b[b_v = k+1]c[b_v \geq k+1]c^{(b)}\neg\,\phi \triangleleft 0 \tag{83}$$

For the second expression, let us temporarily assume the following lemma:

Lemma G.2.

$$\phi b[b_v = k+1]c[b_v \leq k]c^{(b)}\neg\,\phi \triangleleft (k+1)\beta$$

Then by applying Axiom (16) to property (83) and Lemma 16.2 we obtain:

$$\phi b[b_v = k+1]c([b_v \geq k+1]\ +_{[b_v \geq k+1]}\ [b_v \leq k])c^{(b)}\neg\,\phi \triangleleft (k+1)\beta \tag{84}$$

and thus:

$$\phi[b_v = k+1]c^{(b)}\neg\,\phi \triangleleft (k+1)\beta \tag{85}$$

which proves Lemma 16.1.

There thus only remains to prove Lemma 16.2.

We have:

$$\begin{array}{l}\phi b[b_v = k+1]c[b_v \leq k]c^{(b)}\neg\,\phi = \\ \phi b[b_v = k+1]c(\phi + \neg\,\phi)[b_v \leq k]c^{(b)}\neg\,\phi = \text{ by Prop. 2.2} \\ \phi b[b_v = k+1]c(\phi +_\phi \neg\,\phi)[b_v \leq k]c^{(b)}\neg\,\phi = \text{ by Ax. (5)} \\ \phi b[b_v = k+1]c(\phi[b_v \leq k]c^{(b)}\neg\,\phi +_\phi \neg\,\phi[b_v \leq k]c^{(b)}\neg\,\phi)\end{array}$$

As before we want to use Axiom (16) to bound the latter expression.

First note that:

$$\begin{array}{rl}\phi[b_v \leq k]c^{(b)}\neg\,\phi \triangleleft k\beta, & \text{by } IH(k) \\ \phi b[b_v = k+1]c \triangleleft 1, & \text{by Axiom 19}\end{array}$$

From that we obtain:

$$\begin{array}{l}\phi b[b_v = k+1]c\phi[b_v \leq k]c^{(b)}\neg\,\phi = \\ (\phi b[b_v = k+1]c)(\phi[b_v \leq k]c^{(b)}\neg\,\phi) \triangleleft 1 \cdot (k\beta), \text{ by Ax. (17)}\end{array}$$

So

$$(\phi b[b_v = k+1]c)(\phi[b_v \leq k]c^{(b)}\neg\,\phi) \triangleleft k\beta \tag{86}$$

Moreover we have:

$$[b_v = k+1]b \triangleleft 1, \quad \text{by Axiom (19)} \tag{87}$$
$$\phi c\neg\,\phi \triangleleft \beta, \quad \text{by assumption (premise of the rule)} \tag{88}$$
$$[b_v \leq k]c^{(b)}\neg\,\phi \triangleleft 1, \text{ by Axiom (19)} \tag{89}$$

And thus:

$$\begin{array}{l}\phi b[b_v = k+1]c\neg\,\phi[b_v \leq k]c^{(b)}\neg\,\phi = \\ b[b_v = k+1]\phi c\neg\,\phi[b_v \leq k]c^{(b)}\neg\,\phi = \\ (b[b_v = k+1])(\phi c\neg\,\phi)([b_v \leq k]c^{(b)}\neg\,\phi) \triangleleft 1 \cdot \beta \cdot 1 \\ \qquad \text{by (17), (87), (88) and (89)}\end{array}$$

So

$$(b[b_v = k+1])(\phi c\neg\,\phi)([b_v \leq k]c^{(b)}\neg\,\phi) \triangleleft \beta \tag{90}$$

Now, recall that we had shown previously:

$$\begin{array}{l}\phi b[b_v = k+1]c[b_v \leq k]c^{(b)}\neg\,\phi = \\ \phi b[b_v = k+1]c(\phi[b_v \leq k]c^{(b)}\neg\,\phi +_\phi \neg\,\phi[b_v \leq k]c^{(b)}\neg\,\phi)\end{array}$$

By applying Axiom (16) to this latter expression, with (86) and (90) we get:

$$\phi b[b_v = k+1]c[b_v \leq k]c^{(b)}\neg\,\phi \triangleleft (k+1)\beta \tag{91}$$

We have thus proven Lemma 16.2 and hence concluded the proof that $IH(k)$ implies $IH(k+1)$.

Then by induction we can conclude that $IH(k)$ holds for all integers k. The case of rule (**While**) is completed.

H Example: Report-noisy-max algorithm

We will now consider the example of the *Report-noisy-max* algorithm, which has been analysed in [5] with the logic aHL (see also [9] for more background on this algorithm). Our analysis here using aGKAT will be similar to the previous one in aHL, but the equational approach of aGKAT will simplify some steps.

We consider a finite set $\mathcal{R}$ and a quality score function *qscore*, which takes as input a pair of an element r of $\mathcal{R}$ and a database d, and returns a real number. The goal of the Report-noisy-max algorithm is to find an element r^* of $\mathcal{R}$ which approximately minimizes the function *qscore* on d. The algorithm is randomized and only computes an approximate minimization because it is designed to satisfy a differential privacy property (see [9]).

Report-noisy-max proceeds by computing for each element r of $\mathcal{R}$ the quality score $qscore(r, d)$ and adding to it a Laplacian noise (according to the Laplace mechanism for differential privacy [9]) and returning the element r^* with the highest noisy value.

Here we do not deal with the differential privacy property of this program, but instead our objective is to study its *accuracy*, that is to say to bound the difference between the value of $qscore(r^*, d)$ and the real minimum of $qscore(\cdot, d)$ on $\mathcal{R}$.

The algorithm Report-noisy-max can be written as a GKAT program c as follows (we use intermediary notations c', c_1 and b for readability):

$$
\begin{array}{rcl}
c & = & (\mathit{flag} \leftarrow 1);\ (\mathit{best} \leftarrow 0);\ (\mathcal{R}_0 \leftarrow \mathcal{R});\ (\mathcal{R}' \leftarrow \varnothing);\ c'^{[\mathcal{R} \neq \varnothing]};\ \mathit{return}(r^*) \\
\text{where} & & \\
c' & = & (r \leftarrow \mathit{pick}(\mathcal{R}));\ (\mathit{noisy}[r] \xleftarrow{\$} \mathcal{L}_{\epsilon/2}(\mathit{qscore}(r, d))); \\
 & & (c_1 +_b 1); \\
 & & (\mathcal{R} \leftarrow \mathcal{R} \backslash \{r\});\ (\mathcal{R}' \leftarrow \mathcal{R}' \cup \{r\}) \\
c_1 & = & (\mathit{flag} \leftarrow 0);\ (r^* \leftarrow r);\ (\mathit{best} \leftarrow \mathit{noisy}[r]) \\
b & = & (\mathit{noisy}[r] > \mathit{best}) + (\mathit{flag} == 1)
\end{array}
$$

The variable *flag* has Boolean values ($\{0, 1\}$), $\mathcal{R}$, $\mathcal{R}_0$ and $\mathcal{R}'$ are sets, r, r^* range over elements of $\mathcal{R}$, $\mathit{noisy}[r]$ and *best* range over reals. The notation $\mathit{noisy}[r]$ is an array-like notation for representing n variables, where n is the size of the set $\mathcal{R}$. Note that the variable $\mathcal{R}'$ does not play any rôle in the algorithm, it will just be used to express properties of the execution.

This program uses the following kinds of actions and tests:

- actions representing basic operations on sets: picking an (arbitrary) element r from a set ($r \leftarrow \mathit{pick}(\mathcal{R})$), removing an element ($\mathcal{R} \leftarrow \mathcal{R} \backslash \{r\}$)) and adding an element ($\mathcal{R}' \leftarrow \mathcal{R}' \cup \{r\}$),
- sampling from a Laplacian distribution centered in a with parameter p : ($x \xleftarrow{\$} \mathcal{L}_p(a)$),
- tests: inequalities for reals, equality for Boolean value, comparison to empty set for sets $[\mathcal{R} \neq \varnothing]$; we will also need three additional tests for expressing properties on the execution, that we will see later.

We recall the following accuracy property of the Laplace distribution [5]:

Lemma H.1. *Assume β belongs to $[0, 1]$ and let ν be a sample from the Laplace distribution $\mathcal{L}_p(a)$, then we have:*

$$Pr_{\mathcal{L}_p(a)}[\mid \nu - a \mid > \frac{1}{p} \log(\frac{1}{\beta})] < \beta \tag{92}$$

Therefore the Laplacian distribution satisfies the following property:

$$\models (x \xleftarrow{\$} \mathcal{L}_p(a))[\mid x - a \mid > \frac{1}{p} \log(\frac{1}{\beta})] \triangleleft \beta \tag{93}$$

This corresponds to the following instance of the (Rand) axiom in aHL (see Fig. 2):

$$\frac{\forall\, m, \mathcal{P}_i[\![x \xleftarrow{\$} \mathcal{L}_p(a)]\!](m)[|\ x - a\ |> \frac{1}{p}\log(\frac{1}{\beta})] \leq \beta}{\vdash_\beta \textbf{do}\ .(x \xleftarrow{\$} \mathcal{L}_p(a)) : T \Rightarrow|\ x - a\ |\leq \frac{1}{p}\log(\frac{1}{\beta})}$$

Now, (93) gives us for the sampling in program c':

$$\models (noisy[r] \xleftarrow{\$} \mathcal{L}_{\epsilon/2}(qscore(r,d)))[|\ noisy[r] - qscore(r,d)\ |> \frac{2}{\epsilon}\log(\frac{|\ \mathcal{R}_0\ |}{\beta})] \triangleleft \frac{\beta}{|\ \mathcal{R}_0\ |} \tag{94}$$

We want to establish a property for the whole program c'.

Denote as b_1 the test $[|\ noisy[r] - qscore(r,d)\ |\leq \frac{2}{\epsilon}\log(\frac{|\mathcal{R}_0|}{\beta})]$. Observe that b_1 (and thus also $\overline{b_1}$) only depends on the values of $noisy[r]$ and $qscore(r,d)$. Moreover $noisy[r]$ and $qscore(r,d)$ are not changed by the last 3 actions of c'. Therefore by applying Prop.4.4.1 we get $c';\ \overline{b_1} \equiv c''$, where c'' is the expression obtained by replacing in c' $(noisy[r] \xleftarrow{\$} \mathcal{L}_{\epsilon/2}(qscore(r,d)))$ by $(noisy[r] \xleftarrow{\$} \mathcal{L}_{\epsilon/2}(qscore(r,d)));\ \overline{b_1}$.

So we know that:

$$\models (noisy[r] \xleftarrow{\$} \mathcal{L}_{\epsilon/2}(qscore(r,d)));\ \overline{b_1} \quad \triangleleft \quad \frac{\beta}{|\ \mathcal{R}_0\ |}\ \text{(this is (94))} \tag{95}$$

$$\models c_0 \quad \triangleleft \quad 1 \text{ for any } c_0\text{, by axiom (19)} \tag{96}$$

So by applying axiom (17) to (95) and (96) we obtain that $\models c'' \triangleleft \frac{\beta}{|\mathcal{R}_0|}$. Therefore as $c';\ \overline{b_1} \equiv c''$ we get by Prop. 4.4.2 that:

$$\models c' \cdot \overline{b_1} \triangleleft \frac{\beta}{|\ \mathcal{R}_0\ |} \tag{97}$$

We want to prove an invariant for the body c' of the *while* loop in c. For that consider the test b_2 corresponding to the following predicate:

$$\phi_2 = \forall\, r \in \mathcal{R}',\ |\ noisy[r] - qscore(r,d)\ |\leq \frac{2}{\epsilon}\log(\frac{|\ \mathcal{R}_0\ |}{\beta})$$

We have:

$$b_2 \cdot b_1 \cdot (\mathcal{R}' \leftarrow \mathcal{R}' \cup \{r\}) \cdot \overline{b_2} \quad \equiv \quad 0 \tag{98}$$

Let us denote as c_2 the expression c' deprived of the last action, that is to say: $c' = c_2 \cdot (\mathcal{R}' \leftarrow \mathcal{R}' \cup \{r\})$. The reasoning we did on c' before can be done for c_2, and so just as (97) we have:

$$\models c_2 \cdot \overline{b_1} \triangleleft \frac{\beta}{|\ \mathcal{R}_0\ |} \tag{99}$$

Therefore by axioms (17) and (19) we get:

$$\models b_2 \cdot c_2 \cdot \overline{b_1} \triangleleft \frac{\beta}{|\ \mathcal{R}_0\ |} \tag{100}$$

Moreover as c_2 does not modify $\mathcal{R}'$, by using Prop.4.4.1 we get:

$$\models b_2 \cdot c_2 \cdot \overline{b_2} \triangleleft 0 \tag{101}$$

Thus by using the aGKAT encoding (Theorem 5.1) of the aHL rule (And) we obtain from (100) and (101):

$$\models b_2 \cdot c_2 \cdot \overline{b_1 \cdot b_2} \triangleleft \frac{\beta}{|\ \mathcal{R}_0\ |} \tag{102}$$

Equation (98) gives us:

$$\models (b_1 \cdot b_2) \cdot (\mathcal{R}' \leftarrow \mathcal{R}' \cup \{r\}) \cdot \overline{b_2} \quad \triangleleft \quad 0 \tag{103}$$

By using the aGKAT encoding of the aHL rule (Seq) we get from (102) and (103):

$$\models b_2 \cdot c' \cdot \overline{b_2} \lhd \frac{\beta}{\mid \mathcal{R}_0 \mid} \tag{104}$$

By using the aGKAT encoding of the aHL rule (While) we get from (104), as the loop runs for $\mid \mathcal{R}_0 \mid$ iterations:

$$\models b_2 \cdot c'^{[\mathcal{R} \neq \varnothing]} \cdot \overline{b_2} \lhd \beta \tag{105}$$

Finally as $(\mathcal{R}' \leftarrow \varnothing) \cdot \overline{b_2} \equiv 0$ we deduce from (105) using (Seq) that:

$$\models c \cdot \overline{b_2} \lhd \beta \tag{106}$$

So we have proven in aGKAT the encoding of:

$$\vdash_\beta c : T \Rightarrow \forall\, r \in \mathcal{R}',\ \mid noisy[r] - qscore(r, d) \mid \leq \frac{2}{\epsilon} \log(\frac{\mid \mathcal{R}_0 \mid}{\beta}) \tag{107}$$

In order to obtain a property relating $qscore(r^*, d)$ to the values of $qscore(r, d)$ in order to have an accuracy result, we need to relate $noisy[r^*]$ to the $noisy[r]$. For that we will consider the following predicate:

$$\phi_3 = \forall\, r \in \mathcal{R}',\ (noisy[r^*] \geq noisy[r]) \wedge (best = noisy[r^*]) \tag{108}$$

Denote as b_3 the corresponding test. We have:

$$\models b_3 \cdot c' \cdot \overline{b_3} \quad \lhd \quad 0 \tag{109}$$

$$\models b_3 \cdot c'^{[\mathcal{R} \neq \varnothing]} \cdot \overline{b_3} \quad \lhd \quad 0 \text{ by (While) rule} \tag{110}$$

$$\models c \cdot \overline{b_3} \quad \lhd \quad 0 \text{ because } (\mathcal{R}' \leftarrow \varnothing) \cdot \overline{b_3} = 0 \tag{111}$$

Then from (106) and (111) we get with rule (And):

$$\models c \cdot \overline{b_2 \cdot b_3} \quad \lhd \quad \beta \tag{112}$$

By arithmetical reasoning we have that $\phi_2 \wedge \phi_3$ implies that for any r in $\mathcal{R}'$ we have:

$$qscore(r^*, d) \quad > \quad qscore(r, d) - \frac{2}{\epsilon} \log(\frac{\mid \mathcal{R}_0 \mid}{\beta}) \tag{113}$$

Therefore if b_4 denotes the test corresponding to

$$\phi_4 = \forall\, r \in \mathcal{R}',\ qscore(r^*, d) > qscore(r, d) - \frac{2}{\epsilon} \log(\frac{\mid \mathcal{R}_0 \mid}{\beta}) \tag{114}$$

we have:

$$b_2 \cdot b_3 \equiv (b_2 \cdot b_3) \cdot b_4 \tag{115}$$

Therefore from (112) and (115) we get by rule (Weak):

$$\models c \cdot \overline{b_4} \quad \lhd \quad \beta \tag{116}$$

So we have proven using aGKAT that the property corresponding to the following judgement holds::

$$\vdash_\beta c : T \Rightarrow \forall\, r \in \mathcal{R}',\ qscore(r^*, d) > qscore(r, d) - \frac{2}{\epsilon} \log(\frac{\mid \mathcal{R}_0 \mid}{\beta}) \tag{117}$$

This shows an accuracy property for c: with failure probability β, the result r^* of c gives a quality score which is not far below the quality score of any other element r of $\mathcal{R}'$ (that is to say $\mathcal{R}_0$).