Available online at www.sciencedirect.com

ScienceDirect

Electronic Notes in Theoretical Computer Science 325 (2016) 277–298

Electronic Notes in Theoretical Computer Science

www.elsevier.com/locate/entcs

Approximate Relational Hoare Logic for Continuous Random Samplings

Tetsuya Sato[1]

Research Institute for Mathematical Sciences, Kyoto University, Kyoto, 606-8502, Japan

Abstract

Approximate relational Hoare logic (apRHL) is a logic for formal verification of the differential privacy of databases written in the programming language pWHILE. Strictly speaking, however, this logic deals only with discrete random samplings. In this paper, we define the graded relational lifting of the subprobabilistic variant of Giry monad, which described differential privacy. We extend the logic apRHL with this graded lifting to deal with continuous random samplings. We give a generic method to give proof rules of apRHL for continuous random samplings.

Keywords: Differential privacy, Denotational semantics, Giry monad, Graded monad, Relational lifting

1 Introduction

Differential privacy is a *definition* of privacy of *randomised* databases proposed by Dwork, McSherry, Nissim and Smith [7]. A randomised database satisfies ε-differential privacy (written ε-differentially private) if for any two adjacent data, the difference of their output probability distributions is bounded by the privacy strength ε. Differential privacy guarantees high secrecy against database attacks regardless of the attackers' background knowledge, and it has the composition laws, with which we can calculate the privacy strength of a composite database from the privacy strengths of its components.

Approximate relational Hoare logic (apRHL) [2,17] is a probabilistic variant of the *relational Hoare logic* [4] for formal verification of the differential privacy of databases written in the programming language pWHILE. In the logic apRHL, a parametric relational lifting, which relate probability distributions, play a central role to describe differential privacy in the framework of verification. This parametric lifting is an extension of the relational lifting [10, Section 3] that captures

[1] Email:satoutet@kurims.kyoto-u.ac.jp

http://dx.doi.org/10.1016/j.entcs.2016.09.043

1571-0661/© 2016 The Author(s). Published by Elsevier B.V.

probabilistic bisimilarity of Markov chains [13] (see also [6, lemma 4]). The concept of differential privacy is described in the category of binary relation and mappings between them, and verified by the logic apRHL.

Strictly speaking, however, apRHL deals only with random samplings of *discrete* distributions, while the algorithms in many actual studies for differential privacy are modelled with *continuous* distributions, such as, the Laplacian distributions over real line. Therefore apRHL is desired to be extended to deal with random continuous samplings.

1.1 Contributions

Main contributions of this paper are the following two points:

- We define the graded relational lifting of sub-Giry monad describing differential privacy for continuous random samplings.
- We extend the logic apRHL [2,17] for continuous random samplings (we name *continuous apRHL*) .

This graded relational lifting is developed without witness distributions of probabilistic coupling, and hence is constructed in a different way from the coupling-based parametric lifting of relations given in the studies of apRHL [1,2,17].

In the continuous apRHL, we mainly extend the proof rules for relation compositions and the frame rule. We also develop a generic method to construct proof rules for random samplings. By importing the new rules added to apRHL+ in [1], we give a formal proof of the differential privacy of the *above-threshold algorithm* for real-valued queries [8, Section 3.6].

1.2 Preliminaries

We denote by **Set**, $\omega\mathbf{CPO}_\perp$, and **Meas** the categories of all sets and functions, all ω-complete partial orders with the least element and continuous functions between them, and all measurable spaces and measurable functions respectively. The category **Meas** is complete, cocomplete, and distributive. The forgetful functor $U\colon \mathbf{Meas} \to \mathbf{Set}$ preserves limits and colimits. For each measurable space X, we write Σ_X for the σ-algebra of X. For any $A \in \Sigma_X$, the *indicator function* $\chi_A\colon X \to [0,1]$ of A is defined by $\chi_A(x) = 1$ if $X \in A$ and $\chi_A(x) = 0$ otherwise.

The Category of Relations between Measurable Spaces

We introduce the category **BRel**(**Meas**) of binary relations between *measurable spaces* as follows:

- An object is a triple (X, Y, Φ) consisting of measurable spaces X and Y and a relation Φ between X and Y (i.e. $\Phi \subseteq UX \times UY$). We remark that Φ does not necessary to be a measurable subset of the product space $X \times Y$.
- An arrow $(f, g)\colon (X, Y, \Phi) \to (X', Y', \Phi')$ is a pair of measurable functions $f\colon X \to X'$ and $g\colon Y \to Y'$ such that $(Uf \times Ug)(\Phi) \subseteq \Phi'$.