Amplifying Contextual Distance in Higher-Order Languages, using the Law of Large Numbers

Raphaëlle Crubillé *[1,2] and Houssein Mansour†[2]

[1]CNRS
[2]Aix-Marseille Université

A central problem in the study of higher-order programming languages has been the question of *program equivalence*: when can we say that two programs, syntactically different, are nonetheless equivalent ? As an illustration, if we consider program transformations made for optimisation purposes, we would like to be able to say and prove that the output program is in some sense equivalent to the input program. A broadly accepted definition is *contextual equivalence*, that was first introduced by Morris [10]: two programs are equivalent when they behave exactly the same no matter in which *context* we put them, and this definition is made self-contained by deciding that contexts can be modelled as programs written in the same language. More formally, a generic definition of context equivalence for an higher-order language should be of the form:

$$M \sim N \text{ when } \forall C : \text{ context }, \mathrm{Obs}(C[M]) = \mathrm{Obs}(C[N]),$$

where Obs is a notion of *observation* associated to the language—for instance in a probabilistic language with Booleans as a ground type, we can say that the output type of contexts is Bool, and that Obs(C[M]) becomes the probability that $C[M]$ returns 0.

In a probabilistic setting, however, we would like to be also able to say that a program transformation returns a program that is not always equivalent to the input program, but that behaves the same with at least some given probability: this idea led to the concept of *programs distances*, for discrete randomised programming languages [9, 4, 5, 8, 6, . . .]. Similarly,

*raphaelle.crubille@lis-lab.fr
†houssein.mansour@etu.univ-amu.fr

quantitative ideas appear in *computational security* [3, 2]: two programs are *computationally indistinguishable* when no admissible adversary is able to distinguish them with more than *negligible probability.* Programs distances have also been investigated in the setting of *differential privacy* [7]. The quantitative case happens to be considerably less well behaved as the qualitative case, and by contrast with the case of equivalences there is no universally accepted notion of *contextual distance.* A natural generalisation of contextual equivalence could be [4]:

$$d^{ctx}(M,N) = \sup_{C \text{ context}} |\mathrm{Obs}(C[M]) - \mathrm{Obs}(C[N])|$$

It is a *refinement* of contextual equivalence, in the sense that we can recover the information about program equivalence from the pseudo-metric d^{ctx}: it is easy to see that two programs M, N are contextually equivalent if and only if $d^{ctx}(M,N) = 0$. The definition above, however, does not behave as expected as soon as contexts have *copying abilities*: in particular, if we start with a λ-calculus Λ with copying, and with a type system ensuring termination, a *trivialisation phenomenon* occurs [4]: the distance between two programs in Λ_{choice} is always either 0 or 1. We give below a quite informal statement of the trivialisation theorem:

Theorem 1 ([4])**.** *Let $\mathcal{L}$ a probabilistic language where all programs terminate, and where contexts have at least the following abilities: copying their argument (before and after evaluating it), and computing all functions on natural numbers built using primitive recursion. Then for any comparable pair of programs M, N, either they are contextually equivalent, or $d^{ctx}(M,N) = 1$.*

This result was shown in [4], with the crux of the proof done with techniques coming from real analysis: the main tool was Bernstein's theorem, that ensures that every continuous is the limit of a family of so-called *Bernstein's polynomials.* In the talk, we will present a *new* proof that uses a quite elementary result in *probability theory*, namely the *law of large numbers*–proved by Bernoulli [1] as early as 1713–that said that the empiric mean of a probability distribution *converges*–in some precise sense–to its *expectancy.* Our proof allows to understand the trivialisation theorem in terms of probability theory: as soon as the language we consider is expressive enough that contexts are able to compute the empiric mean of the probability distribution they take as argument, then they are able to distinguish *with probability* 1 any pair of distributions with different expectancy. It also opens interesting research directions about the rate of convergence of contextual distance towards the trivial metric when we allow contexts to use a bounded, but increasing, numbers of copies of their argument. Indeed, the rate of convergence of the weak law of large numbers is very well-studied mathematically, and we could potentially extract from there information about the rate of convergence of contextual distance.

References

[1] J. Bernoulli. *Ars coniectandi.* Impensis Thurnisiorum, fratrum, 1713.

[2] A. Broadbent and M. Karvonen. Categorical composable cryptography. In *International Conference on Foundations of Software Science and Computation Structures*, pages 161–183. Springer International Publishing Cham, 2022.

[3] A. Cappai and U. Dal Lago. On equivalences, metrics, and polynomial time. In *Proc. of FCT*, pages 311–323, 2015.

[4] R. Crubillé and U. Dal Lago. Metric reasoning about λ-terms: The affine case. In *Proc. of LICS*, pages 633–644, 2015.

[5] R. Crubillé and U. Dal Lago. Metric reasoning about λ-terms: The general case. In *European Symposium on Programming*, pages 341–367. Springer, 2017.

[6] U. Dal Lago, N. Hoshino, and P. Pistone. On the lattice of program metrics. In M. Gaboardi and F. van Raamsdonk, editors, *8th International Conference on Formal Structures for Computation and Deduction, FSCD 2023, July 3-6, 2023, Rome, Italy*, volume 260 of *LIPIcs*, pages 20:1–20:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023.

[7] A. A. de Amorim, M. Gaboardi, J. Hsu, and S.-y. Katsumata. Probabilistic relational reasoning via metrics. In *2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–19. IEEE, 2019.

[8] T. Ehrhard. Differentials and distances in probabilistic coherence spaces. *Log. Methods Comput. Sci.*, 18(3), 2022.

[9] D. Gebler, K. G. Larsen, and S. Tini. Compositional metric reasoning with probabilistic process calculi. In *Proc. of FoSSaCS*, pages 230–245, 2015.

[10] J. H. Morris Jr. *Lambda-calculus models of programming languages.* PhD thesis, Massachusetts Institute of Technology, 1969.