The Big-O Problem for Weighted Automata

Dmitry Chistikov
University of Warwick

Stefan Kiefer
University of Oxford

Andrzej S. Murawski
University of Oxford

David Purser
University of Warwick and MPI-SWS

Given two weighted automata $\mathcal{A}, \mathcal{B}$ over an algebraic structure $(\mathcal{S}, +, \times)$, the equivalence problem asks whether the two associated functions $f_{\mathcal{A}}, f_{\mathcal{B}} \colon \Sigma^* \to \mathcal{S}$ are equal: $f_{\mathcal{A}}(w) = f_{\mathcal{B}}(w)$ for all finite words w over the alphabet Σ. Over the ring $(\mathbb{Q}, +, \times)$, equivalence is decidable in polynomial time by the results of Schützenberger [9] and Tzeng [11].

Replacing = with $\leq$ makes the problem harder: even for the ring $(\mathbb{Q}, +, \times)$ the question of whether $f_{\mathcal{A}}(w) \leq f_{\mathcal{B}}(w)$ for all $w \in \Sigma^*$ is undecidable—even if $f_{\mathcal{A}}$ is constant [7].

We introduce and study another natural problem, in which the ordering is relaxed from inequality to inequality to within a constant factor.

Definition. *Given two WA $\mathcal{A}$ and $\mathcal{B}$, is it true that there exists a constant $c > 0$ such that*

$$f_{\mathcal{A}}(w) \leq c \cdot f_{\mathcal{B}}(w) \qquad \text{for all } w \in \Sigma^* \text{ ?}$$

Using standard mathematical notation, this condition asserts that $f_{\mathcal{A}}(w) = O(f_{\mathcal{B}}(w))$ as $|w| \to \infty$, and we refer to this problem as the *big-O* problem accordingly.

The *big-*Θ problem (which turns out to be computationally equivalent to the big-O problem), in line with the $\Theta(\cdot)$ notation in analysis of algorithms, asks whether $f_{\mathcal{A}} = O(f_{\mathcal{B}})$ and $f_{\mathcal{B}} = O(f_{\mathcal{A}})$.

We restrict our attention to the ring $(\mathbb{Q}, +, \times)$ and only consider *non-negative weighted automata*, i.e., those in which all transitions have non-negative weights. Our main findings are as follows:

- The big-O problem for non-negative WA is **undecidable in general**, by a reduction from nonemptiness for probabilistic automata. The result applies even for the special cases of labelled Markov chains and probabilistic automata.
- For **unambiguous automata**, (where every word has at most one accepting path) the big-O problem is decidable in polynomial time.
- In the **unary case**, i.e., if the input alphabet Σ is a singleton, the big-O problem is also decidable and, complete for the complexity class **coNP**. Our upper bound argument refines an analysis of growth of entries in powers of non-negative matrices by Friedland and Schneider [8], and the lower bound is by a reduction from unary NFA universality [10].
- In a more general **bounded case**, i.e., if the languages of all words positive weighted words are included in $w_1^* w_2^* \dots w_m^*$ for some finite words $w_1, \dots, w_m \in \Sigma^*$ the big-O problem is decidable subject to Schanuel's conjecture (a well-known conjecture in transcendental number theory [5], entailing that the first-order theory of the real numbers with the exponential function is decidable [6]).

Relation to total variation distances In the labelled Markov chain setting, the big-O problem can be reformulated as a boundedness problem for the following function. For two LMCs $\mathcal{A}$ and $\mathcal{B}$, define the (asymmetric) *ratio variation function* by

$$r(\mathcal{A}, \mathcal{B}) = \sup_{E \subseteq \Sigma^*} \frac{f_{\mathcal{A}}(E)}{f_{\mathcal{B}}(E)},$$

where $f_{\mathcal{A}}(E)$ and $f_{\mathcal{B}}(E)$ denote the total probability mass associated with an arbitrary set of finite words $E \subseteq \Sigma^*$ in $\mathcal{A}$ and $\mathcal{B}$, respectively. The supremum over $E \subseteq \Sigma^*$ can be replaced with supremum over $w \in \Sigma^*$. Consequently, the big-O problem for LMCs is equivalent to deciding whether $r(\mathcal{A}, \mathcal{B}) < \infty$.

Finding the value of r amounts to asking for the optimal (minimal) constant in the big-O notation.

r is a ratio-oriented variant of the classic *total variation distance tv*, defined by

$$tv(\mathcal{A}, \mathcal{B}) = \sup_{E \subseteq \Sigma^*} f_{\mathcal{A}}(E) - f_{\mathcal{B}}(E),$$

which is a well-established way of comparing two labelled Markov chains [1, 4]. We also consider the problem of approximating r to a given precision and the problem of comparing it with a given constant (threshold problem), showing that both are undecidable.

Application to privacy Consider a system $\mathcal{M}$, modelled by a single labelled Markov chain, where output words are observable to the environment but we want to protect the privacy of the starting configuration. We write f_s for the weighting function of $\mathcal{M}$ when executed from state s.

Definition. *[2, 3] Let $R \subseteq Q \times Q$ be a symmetric relation, which relates the starting configurations intended to remain indistinguishable. Given $\epsilon \geq 0$, $\mathcal{M}$ is ϵ-differentially private (with respect to R) if, for all pairs $s, s' \in Q$ such that $(s, s') \in R$, we have $f_s(E) \leq e^{\epsilon} \cdot f_{s'}(E)$ for every observable set of traces $E \subseteq \Sigma^*$.*

Note that there exists such an ϵ if and only if $r(s, s') < \infty$ for all $(s, s') \in R$ or, equivalently, (the LMC $\mathcal{M}$ executed from) the state s is big-O of (the LMC $\mathcal{M}$ executed from) the state s' for all $(s, s') \in R$. The minimal such ϵ satisfies $e^{\epsilon} = \max_{(s,s') \in R} r(s, s')$.

Hence, our results show that even deciding whether the ratio variation distance is finite or $+\infty$ is, in general, impossible. Likewise, it is undecidable whether a system modelled by a labelled Markov chain provides any degree of differential privacy, however low.

References

[1] Taolue Chen and Stefan Kiefer. 2014. On the total variation distance of labelled Markov chains. In *Proceedings of the Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. ACM, 33.

[2] Dmitry Chistikov, Andrzej S. Murawski, and David Purser. 2019. Asymmetric Distances for Approximate Differential Privacy. In *30th International Conference on Concurrency Theory, CONCUR 2019, August 27-30, 2019, Amsterdam, the Netherlands*. 10:1–10:17.

[3] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam D. Smith. 2006. Calibrating Noise to Sensitivity in Private Data Analysis. In *Theory of Cryptography, Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006, Proceedings*. 265–284. https://doi.org/10.1007/11681878_14

[4] Stefan Kiefer. 2018. On Computing the Total Variation Distance of Hidden Markov Models. In *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

[5] Serge Lang. 1966. *Introduction to transcendental numbers*. Addison-Wesley Pub. Co.

[6] Angus Macintyre and Alex J Wilkie. 1996. *On the decidability of the real exponential field*. A K Peters, 441–467.

[7] Azaria Paz. 2014. *Introduction to probabilistic automata*. Academic Press.

[8] Hans Schneider. 1986. The influence of the marked reduced graph of a nonnegative matrix on the Jordan form and on related properties: A survey. *Linear Algebra Appl.* 84 (1986), 161–189.

[9] Marcel Paul Schützenberger. 1961. On the definition of a family of automata. *Information and control* 4, 2-3 (1961), 245–270.

[10] Larry J Stockmeyer and Albert R Meyer. 1973. Word problems requiring exponential time (preliminary report). In *Proceedings of the fifth annual ACM symposium on Theory of computing*. ACM, 1–9.

[11] Wen-Guey Tzeng. 1992. A polynomial-time algorithm for the equivalence of probabilistic automata. *SIAM J. Comput.* 21, 2 (1992), 216–227.