




# A Complete Quantitative Axiomatisation of Behavioural Distance of Regular Expressions

Wojciech Różowski   

Department of Computer Science, University College London, UK

---

## Abstract

Deterministic automata have been traditionally studied through the point of view of language equivalence, but another perspective is given by the canonical notion of *shortest-distinguishing-word* distance quantifying the of states. Intuitively, the longer the word needed to observe a difference between two states, then the closer their behaviour is. In this paper, we give a sound and complete axiomatisation of *shortest-distinguishing-word* distance between regular languages. Our axiomatisation relies on a recently developed quantitative analogue of equational logic, allowing to manipulate rational-indexed judgements of the form  $e \equiv_\varepsilon f$  meaning *term  $e$  is approximately equivalent to term  $f$  within the error margin of  $\varepsilon$* . The technical core of the paper is dedicated to the completeness argument that draws techniques from order theory and Banach spaces to simplify the calculation of the behavioural distance to the point it can be then mimicked by axiomatic reasoning.

**2012 ACM Subject Classification** Theory of computation → Regular languages

**Keywords and phrases** Regular Expressions, Behavioural Distances, Quantitative Equational Theories

**Digital Object Identifier** 10.4230/LIPIcs.ICALP.2024.149

**Category** Track B: Automata, Logic, Semantics, and Theory of Programming

**Related Version** *Full Version*: <https://arxiv.org/abs/2404.13352> [28]

**Funding** This work was partially supported by ERC grant Autoprobe (grant agreement 101002697).

**Acknowledgements** The author wishes to thank Giorgio Bacci, Leo Lobski, Alexandra Silva and Mateo Torres-Ruiz for discussions and feedback, as well as anonymous reviewers for their comments.

## 1 Introduction

Transition systems have been widely employed to model computational phenomena. In theoretical computer science, it is customary to model computations as transition systems and subsequently reason about their equivalence or similarity. Classical examples include checking language equivalence of deterministic finite automata using Hopcroft and Karp’s algorithm [18] or constructing bisimulations between labelled transition systems [25]. Throughout the years, especially in the concurrency theory community, researchers have studied a plethora of different notions of behavioural equivalences and preorders one could impose on a transition system [41]. However, in many practical applications, especially when dealing with probabilistic or quantitative transition systems, asking about such classical notions of equivalence (or similarity) could be too strict, and it might be more reasonable to ask quantitative questions about how far apart the behaviour of the two states is.

A growing line of work on *behavioural distances* [38, 5, 39, 40, 13] answers this problem by equipping state-spaces of transition systems with (pseudo)metric structures quantifying the dissimilarity of states. In such a setting, states at distance zero are not necessarily the same, but rather equivalent with respect to some classical notion of behavioural equivalence. In a nutshell, equipping transition systems with such a notion of distance crucially relies on the possibility of *lifting* the distance between the states to the distance on the observable



© Wojciech Różowski;

licensed under Creative Commons License CC-BY 4.0

51st International Colloquium on Automata, Languages, and Programming (ICALP 2024).

Editors: Karl Bringmann, Martin Grohe, Gabriele Puppis, and Ola Svensson;

Article No. 149; pp. 149:1–149:20



Leibniz International Proceedings in Informatics

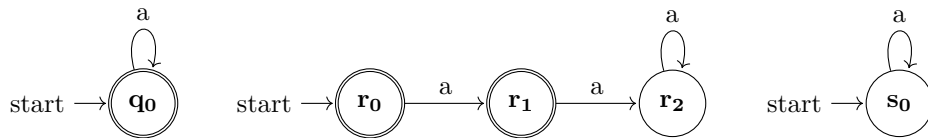
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



behaviour of the transition system. Behavioural distances were originally studied in the context of probabilistic transition systems [16, 39], where observable behaviour is in the form of probability distribution among possible transitions. In such a case, the necessary lifting of distances between states to distances between probability distributions of possible outcomes relies on the famous Kantorovich/Wasserstein liftings, studied traditionally in transportation theory [42]. In general, transition systems can be viewed more abstractly through a well-established category-theoretic framework of coalgebras for an endofunctor [27]. Recent work [5] generalised the Kantorovich/Wasserstein lifting to lifting endofunctors (modelling one-step behaviour of transition systems) from the category of sets and functions to the category of (pseudo)metric spaces and nonexpansive functions, thus allowing equipping a multitude of different kinds of transition systems with a sensible notion of behavioural distance.

Traditionally, besides looking at behavioural equivalence/similarity purely from the algorithmic point of view, one can look at those problems axiomatically, by describing behaviours of transition systems as expressions and by providing formal systems based on (in)equational logic for reasoning about equivalence/similarity of the transition systems described by the expressions. Classic examples include reasoning about language equivalence of Kleene's regular expressions representing deterministic finite automata using inference systems of Salomaa [30] or Kozen [19], or reasoning about bisimilarity of finite-state labelled transition systems through Milner's calculus of finite-state behaviours [24].

In this paper, we are interested in a similar axiomatic point of view, but in the case of behavioural distances. Unfortunately, the classical (in)equational logic cannot be applied here, as it has no way of talking about approximate equivalence. Instead, we rely on the quantitative analogue of equational logic [22], which deals with the statements of the form  $e \equiv_\varepsilon f$ , intuitively meaning *term  $e$  is within the distance of at most  $\varepsilon \in \mathbb{Q}^+$  from the term  $f$* . While the existing work [4, 2, 1] looked at quantitative axiomatisations of behavioural distance for probabilistic transition systems calculated through the Kantorovich/Wasserstein lifting, which can be thought of as a special case of the abstract coalgebraic framework relying on lifting endofunctors to the category of pseudometric spaces, the notions of behavioural distance for other kinds of transition systems have not been axiomatised before. It turns out that the approach to completeness used in [2] relies on properties which are not unique to distances obtained through the Kantorovich/Wasserstein lifting and can be employed to give complete axiomatisations of behavioural distances for other kinds of transition systems obtained through the coalgebraic framework [5]. In this paper, as a starting point, we look at one of the simplest instantiations of that abstract framework in the case of deterministic automata, yielding *shortest-distinguishing-word* distance. To illustrate that notion of distance, consider the following three deterministic finite automata:



Neither of the above automata are language equivalent. Their languages are respectively  $\{\epsilon, a, aa, aaa, \dots\}$ ,  $\{\epsilon, a\}$  and  $\emptyset$ . However, one could argue that the behaviour of the middle automaton is closer to the one on the left rather than the one on the right. In particular, languages of the left and middle automaton agree on all of the words of length less than two, while the left and right one disagree on all words. One can make this idea precise, by providing 1-bounded metric  $d_{\mathcal{L}} : \mathcal{P}(A^*) \times \mathcal{P}(A^*) \rightarrow [0, 1]$  on the set of all formal languages over some fixed alphabet  $A$  given by the following formula, where  $\lambda \in ]0, 1[$  and  $L, M \subseteq A^*$ :

$$d_{\mathcal{L}}(L, M) = \begin{cases} \lambda^{|w|} & w \text{ is the shortest word that belongs to only one of } L \text{ and } M \\ 0 & \text{if } L = M \end{cases} \quad (1)$$

If we set  $\lambda = \frac{1}{2}$ , then  $d_{\mathcal{L}}(\{\epsilon, a, aa, aaa, \dots\}, \{\epsilon, a\}) = \frac{1}{4}$  and  $d_{\mathcal{L}}(\{\epsilon, a, aa, aaa, \dots\}, \emptyset) = 1$ , which allows to formally state that the behaviour of the middle automaton is a better approximation of the left one, rather than the right one. Observe, that we excluded  $\lambda = 0$  and  $\lambda = 1$ , as in both cases  $d_{\mathcal{L}}$  would become a pseudometric setting all languages to be at distance zero or one. Automata in the example above correspond to the regular expressions  $a^*$ ,  $a + 1$  and  $0$  respectively. In order to determine the distance between arbitrary regular expressions  $e$  and  $f$  one would have to construct corresponding deterministic finite automata and calculate (or approximate) the distance between their languages. Instead, as a main contribution of this paper, we present a sound and complete quantitative inference system for reasoning about the shortest-distinguishing-word distance of languages denoted by regular expressions in question. Formally speaking, if  $\llbracket - \rrbracket : \text{Exp} \rightarrow \mathcal{P}(A^*)$  is a function taking regular expressions to their languages, then our inference system satisfies the following:

$$\vdash e \equiv_{\varepsilon} f \iff d_{\mathcal{L}}(\llbracket e \rrbracket, \llbracket f \rrbracket) \leq \varepsilon$$

Although much of our development is grounded in category theory and coalgebra, we spell out all the definitions and results concretely, without the need for specialised language. We organise the paper as follows:

1. In Section 2 we review basic definitions from automata theory and recall the semantics of regular expressions through Brzozowski derivatives [10]. Then, in order to talk about distances, we state basic definitions and properties surrounding (pseudo)metric spaces.
2. In Section 3 we instantiate the framework of coalgebraic behavioural metrics [5] to the concrete case of deterministic automata. We recall the abstract results from [5] in simple automata-theoretic terms.
3. In Section 4 we start by recalling the definitions surrounding the quantitative equational theories [22] from the literature. We then present the axioms of our inference system for the shortest-distinguishing-word distance of regular expressions, give soundness result and provide a discussion about the axioms. The interesting insight is that when relying on quantitative equational theories which contain an infinitary rule capturing the notion of convergence, there is no need for any fixpoint introduction rule. We illustrate this by axiomatically deriving Salomaa's fixpoint rule for regular expressions [30].
4. The key result of our paper is contained in Section 5, where we prove completeness of our inference system. The heart of the argument relies on showing that the behavioural distance of regular expressions can be approximated from above using Kleene's fixpoint theorem, which can be then mimicked through the means of axiomatic reasoning. This part of the paper makes heavy use of the order-theoretic and Banach space structures carried by the sets of pseudometrics over a given set.
5. We conclude in Section 6, review related literature, and sketch directions for future work. Omitted proofs appear in the full version [28].

## 2 Preliminaries

We start by recalling basic definitions surrounding deterministic automata, regular expressions and (pseudo)metric spaces from the literature.

**Deterministic automata.** A deterministic automaton  $\mathcal{M}$  with inputs in a finite alphabet  $A$  is a pair  $(M, \langle o_M, t_M \rangle)$  consisting of a set of states  $M$  and a pair of functions  $\langle o_M, t_M \rangle$ , where  $o_M : M \rightarrow \{0, 1\}$  is the *output* function which determines whether a state  $m$  is final ( $o_M(m) = 1$ ) or not ( $o_M(m) = 0$ ), and  $t : M \rightarrow M^A$  is the *transition* function, which, given an input letter  $a$  determines the next state. If the set  $M$  of states is finite, then we call an automaton  $\mathcal{M}$  a deterministic finite automaton (DFA). We will frequently write  $m_a$  to denote  $t_M(m)(a)$  and refer to  $m_a$  as the derivative of  $m$  for the input  $a$ . Definition of derivatives can be inductively extended to words  $w \in A^*$ , by setting  $m_\epsilon = m$  and  $m_{aw'} = (m_a)_{w'}$ . Note that our definition of deterministic automaton slightly differs from the most common one in the literature, by not explicitly including the initial state. Instead of talking about the language of the automaton, we will talk about the languages of particular states of the automaton. Given a state  $m \in M$ , we write  $L_{\mathcal{M}}(m) \subseteq A^*$  for its language, which is formally defined by  $L_{\mathcal{M}}(m) = \{w \in A^* \mid o(m_w) = 1\}$ . Given two deterministic automata  $(M, \langle o_M, t_M \rangle)$  and  $(N, \langle o_N, t_N \rangle)$ , a function  $h : M \rightarrow N$  is a homomorphism if it preserves outputs and input derivatives, that is  $o_N(h(m)) = o_M(m)$  and  $h(m)_a = h(m_a)$ . The set of all languages  $\mathcal{P}(A^*)$  over an alphabet  $A$  can be made into a deterministic automaton  $(\mathcal{P}(A^*), \langle o_L, t_L \rangle)$ , where for  $l \in \mathcal{P}(A^*)$  the output function is given by  $o_L(l) = [\epsilon \in l]$  and for all  $a \in A$  the input derivative is defined to be  $l_a = \{w \mid aw \in l\}$ . This automaton is *final*, that is for any other automaton  $\mathcal{M} = (M, \langle o_M, t_M \rangle)$  there exists a unique homomorphism from  $M$  to  $\mathcal{P}(A^*)$ , which is precisely given by the map  $L_{\mathcal{M}} : M \rightarrow \mathcal{P}(A^*)$  taking each state  $m \in S$  to its language. Given a set of states  $M' \subseteq M$ , we write  $\langle M' \rangle_{\mathcal{M}} \subseteq M$  for the smallest set of states reachable from  $M'$  through the transition function of the automaton  $\mathcal{M}$ . Clearly,  $(\langle M' \rangle_{\mathcal{M}}, \langle o_M, t_M \rangle)$  is a deterministic automaton. We will abuse the notation and write  $\langle M' \rangle_{\mathcal{M}}$  for  $(\langle M' \rangle_{\mathcal{M}}, \langle o_M, t_M \rangle)$ . The canonical inclusion map  $\iota : \langle M' \rangle_{\mathcal{M}} \hookrightarrow M$  given by  $\iota(m) = m$  for all  $m \in \langle M' \rangle_{\mathcal{M}}$  is a homomorphism from  $\langle M' \rangle_{\mathcal{M}}$  to  $\mathcal{M}$ . In the case of singleton and two-element sets of states, we will simplify the notation and write  $\langle m \rangle_{\mathcal{M}}$  and  $\langle m, m' \rangle_{\mathcal{M}}$ .

**Regular expressions.** We let  $e, f$  range over *regular expressions over  $A$*  generated by the following grammar:

$$e, f \in \text{Exp} ::= 0 \mid 1 \mid a \in A \mid e + f \mid e ; f \mid e^*$$

The standard interpretation of regular expressions  $\llbracket - \rrbracket : \text{Exp} \rightarrow \mathcal{P}(A^*)$  is inductively defined by the following:

$$\llbracket 0 \rrbracket = \emptyset \quad \llbracket 1 \rrbracket = \{\epsilon\} \quad \llbracket a \rrbracket = \{a\} \quad \llbracket e + f \rrbracket = \llbracket e \rrbracket \cup \llbracket f \rrbracket \quad \llbracket e ; f \rrbracket = \llbracket e \rrbracket \diamond \llbracket f \rrbracket \quad \llbracket e^* \rrbracket = \llbracket e \rrbracket^*$$

We write  $\epsilon$  for the empty word. Given  $L, M \subseteq A^*$ , we define  $L \diamond M = \{lm \mid l \in L, m \in M\}$ , where mere juxtaposition denotes concatenation of words.  $L^*$  denotes the *asterate* of the language  $L$  defined as  $L^* = \bigcup_{i \in \mathbb{N}} L^i$  with  $L^0 = \{\epsilon\}$  and  $L^{n+1} = L \diamond L^n$ .

**Brzowski derivatives.** The famous Kleene's theorem states that the formal languages accepted by DFA are in one-to-one correspondence with formal languages definable by regular expressions. One direction of this theorem involves constructing a DFA for an arbitrary regular expression. The most common way is via Thompson construction,  $\epsilon$ -transition removal and determinisation. Instead, we recall a direct construction due to Brzowski [10], in which the set  $\text{Exp}$  of regular expressions is equipped with a structure of deterministic automaton  $\mathcal{R} = (\text{Exp}, \langle o_{\mathcal{R}}, t_{\mathcal{R}} \rangle)$  through so-called Brzowski derivatives [10]. The output derivative  $o_{\mathcal{R}} : \text{Exp} \rightarrow \{0, 1\}$  is defined inductively by the following for  $a \in A$  and  $e, f \in \text{Exp}$ :

$$\begin{aligned} o_{\mathcal{R}}(0) = 0 \quad o_{\mathcal{R}}(1) = 1 \quad o_{\mathcal{R}}(a) = 0 \\ o_{\mathcal{R}}(e + f) = o_{\mathcal{R}}(e) \vee o_{\mathcal{R}}(f) \quad o_{\mathcal{R}}(e; f) = o_{\mathcal{R}}(e) \wedge o_{\mathcal{R}}(f) \quad o_{\mathcal{R}}(e^*) = 1 \end{aligned}$$

Similarly, the transition derivative  $t_{\mathcal{R}} \in \text{Exp} \rightarrow A \rightarrow \text{Exp}$  denoted  $t_{\mathcal{R}}(e)(a) = (e)_a$  is defined by the following:

$$(0)_a = 0 \quad (1)_a = 0 \quad (a')_a = \begin{cases} 1 & a = a' \\ 0 & a \neq a' \end{cases}$$

$$(e + f)_a = (e)_a + (f)_a \quad (e; f)_a = (e_a); f + o_{\mathcal{R}}(e); f \quad (e^*)_a = (e)_a; e^*$$

The canonical language-assigning homomorphism from  $\mathcal{R}$  to  $\mathcal{L}$ , happens to coincide with the semantics map  $\llbracket - \rrbracket$  assigning a language to each regular expression.

► **Lemma 1** ([33, Theorem 3.1.4]). *For all  $e \in \text{Exp}$ ,  $\llbracket e \rrbracket = L_{\mathcal{R}}(e)$*

Instead of looking at infinite-state automaton defined on the state-space of all regular expressions, we can restrict ourselves to the subautomaton  $\langle e \rangle_{\mathcal{R}}$  of  $\mathcal{R}$  while obtaining the semantics of  $e$ .

► **Lemma 2.** *For all  $e \in \text{Exp}$ ,  $\llbracket e \rrbracket = L_{\langle e \rangle_{\mathcal{R}}}(e)$*

Unfortunately, for an arbitrary regular expression  $e \in \text{Exp}$ , the automaton  $\langle e \rangle_{\mathcal{R}}$  is not guaranteed to have a finite set of states. However, simplifying the transition derivatives by removing duplicates in the expressions in the form  $e_1 + \dots + e_n$ , guarantees a finite number of reachable states from any expression. Formally speaking, let  $\doteq \subseteq \text{Exp} \times \text{Exp}$  be the least congruence relation closed under  $(e + f) + g \doteq e + (f + g)$  (Associativity),  $e + f \doteq f + e$  (Commutativity) and  $e \doteq e + e$  (Idempotence) for all  $e, f, g \in \text{Exp}$ . We will write  $\text{Exp}/\doteq$  for the quotient of  $\text{Exp}$  by the relation  $\doteq$  and  $[-]_{\doteq} : \text{Exp} \rightarrow \text{Exp}/\doteq$  for the canonical map taking each expression  $e \in \text{Exp}$  into its equivalence class  $[e]_{\doteq}$  modulo  $\doteq$ . Because of [27, Proposition 5.8],  $\text{Exp}/\doteq$  can be equipped with a structure of deterministic automaton  $\mathcal{Q} = (\text{Exp}/\doteq, \langle o_{\mathcal{Q}}, t_{\mathcal{Q}} \rangle)$ , where for all  $e \in \text{Exp}$ ,  $a \in A$ ,  $o_{\mathcal{Q}}([e]_{\doteq}) = o_{\mathcal{R}}(e)$  and  $([e]_{\doteq})_a = [e_a]_{\doteq}$ , which makes the quotient map  $[-]_{\doteq} : \text{Exp} \rightarrow \text{Exp}/\doteq$  into an automaton homomorphism from the Brzozowski automaton  $\mathcal{R}$  into  $\mathcal{Q}$ . This automaton enjoys the following property:

► **Lemma 3** ([10, Theorem 4.3]). *For any  $e \in \text{Exp}$ , the set  $\langle e \rangle_{\mathcal{Q}} \subseteq \text{Exp}/\doteq$  is finite.*

Through an identical line of reasoning as before (Lemma 2), we can show that:

► **Lemma 4.** *For all  $e \in \text{Exp}$ ,  $L_{\langle [e]_{\doteq} \rangle_{\mathcal{Q}}}([e]_{\doteq}) = \llbracket e \rrbracket$*

**(Pseudo)metric spaces.** Let  $\top \in ]0, \infty]$  be a fixed maximal element. A  $\top$ -bounded *pseudometric* on a set  $X$  (equivalently  $\top$ -*pseudometric* or even just a *pseudometric* if  $\top$  is clear from the context) is a function  $d : X \times X \rightarrow [0, \top]$  satisfying  $d(x, x) = 0$  (*reflexivity*),  $d(x, y) = d(y, x)$  (*symmetry*) and  $d(x, z) \leq d(x, y) + d(y, z)$  (*triangle inequality*) for all  $x, y, z \in X$ . If additionally  $d(x, y) = 0$  implies  $x = y$ ,  $d$  is called a  $\top$ -*metric*. A *(pseudo)metric space* is a pair  $(X, d)$  where  $X$  is a set and  $d$  is a (pseudo)metric on  $X$ . Given pseudometric spaces  $(X, d_X)$  and  $(Y, d_Y)$ , we call a map  $f : X \rightarrow Y$  *nonexpansive*, if for all  $x, x' \in X$ ,  $d_Y(f(x), f(x')) \leq d_X(x, x')$  and an *isometry* if  $d_Y(f(x), f(x')) = d_X(x, x')$ . A simple example of a pseudometric is the discrete metric which can be defined on any set  $X$  as  $d_X(x, x) = 0$  for all  $x \in X$  and  $d(x, y)_X = \top$  for  $x, y \in X$  such that  $x \neq y$ . The set  $D_X$  of (pseudo)metrics over some fixed set  $X$  can be equipped with a partial order structure given by the pointwise order, i.e.  $d \sqsubseteq d' \iff \forall x, x' \in X. d(x, y) \leq d'(x, y)$ .

► **Lemma 5** ([5, Lemma 3.2]).  $(D_X, \sqsubseteq)$  is a complete lattice. The join of an arbitrary set of pseudometrics  $D \subseteq D_X$  is taken pointwise, ie.  $(\sup D)(x, y) = \sup\{d(x, y) \mid d \in D\}$  for  $x, y \in X$ . The meet of  $D$  is defined to be  $\inf D = \sup\{d \mid d \in D_X, \forall d' \in D, d \sqsubseteq d'\}$ .

Crucially for our completeness proof, if we are dealing with descending chains, that is sequences  $\{d_i\}_{i \in \mathbb{N}}$ , such that  $d_i \sqsupseteq d_{i+1}$  for all  $i \in \mathbb{N}$ , then we can also calculate infima in the pointwise way<sup>1</sup>.

► **Lemma 6.** Let  $\{d_i\}_{i \in \mathbb{N}}$  be an infinite descending chain in the lattice  $(D_X, \sqsubseteq)$  of pseudometrics over some fixed set  $X$ . Then  $(\inf\{d_i \mid i \in \mathbb{N}\})(x, y) = \inf\{d_i(x, y) \mid i \in \mathbb{N}\}$  for any  $x, y \in X$ .

**Proof.** It suffices to argue that  $d(x, y) = \inf\{d_i(x, y) \mid i \in \mathbb{N}\}$  is a pseudometric. For reflexivity, observe that  $d(x, x) = \inf\{d_i(x, x) \mid i \in \mathbb{N}\} = \inf\{0\} = 0$  for all  $x \in X$ . For symmetry, we have that  $d(x, y) = \inf\{d_i(x, y) \mid i \in \mathbb{N}\} = \inf\{d_i(y, x) \mid i \in \mathbb{N}\} = d(y, x)$  for any  $x, y \in X$ . The only difficult case is triangle inequality. First, let  $i, j \in \mathbb{N}$  and define  $k = \max(i, j)$ . Since  $d_k \sqsubseteq d_i$  and  $d_k \sqsubseteq d_j$ , we have that  $d_k(x, y) + d_k(y, z) \leq d_i(x, y) + d_j(y, z)$ . Therefore  $\inf\{d_l(x, y) + d_l(y, z) \mid l \in \mathbb{N}\}$  is a lower bound of  $d_i(x, y) + d_j(y, z)$  for any  $i, j \in \mathbb{N}$  and hence it is below the greatest lower bound, that is  $\inf\{d_l(x, y) + d_l(y, z) \mid l \in \mathbb{N}\} \leq \inf\{d_i(x, y) + d_j(y, z) \mid i, j \in \mathbb{N}\}$ . We can use that property to show the following  $d(x, y) = \inf\{d_i(x, y) \mid i \in \mathbb{N}\} \leq \inf\{d_i(x, y) + d_i(y, z) \mid i \in \mathbb{N}\} \leq \inf\{d_i(x, y) + d_j(y, z) \mid i, j \in \mathbb{N}\} = \inf\{d_i(x, y) \mid i \in \mathbb{N}\} + \inf\{d_j(y, z) \mid j \in \mathbb{N}\} = d(x, y) + d(y, z)$ , which completes the proof. ◀

Additionally, the set of pseudometrics can be equipped with a norm. We write  $\overline{\mathbb{R}} = [-\infty, \infty]$  for the set of extended reals. For any set  $X$ , the set of functions  $\overline{\mathbb{R}}^{X \times X}$ , which is a superset of  $D_X$ , can be seen as a Banach space [26] (complete normed vector space) by means of the sup-norm  $\|d\| = \sup_{x, y \in X} |d(x, y)|$ . This structure will implicitly underly some of the claims used as intermediate steps in the proof of completeness.

### 3 Behavioural distance

We now instantiate the abstract coalgebraic framework [5] to the case of deterministic automata relying on the lifting described in [5, Example 5.33]. We concretise the generic results from that paper and spell them in simple automata-theoretic terms.

**Lifting pseudometrics.** Let  $\mathcal{M} = (M, \langle o_M, t_M \rangle)$  be a deterministic automaton. Its one-step observable behaviour (after applying the output and transition derivatives) can be seen as pairs of the type  $\{0, 1\} \times M^A$ , where the first component determines whether the given state is accepting or not and the second one gives successor state for each letter from the input alphabet. Let's say we have two such observations  $\langle o_1, f_1 \rangle, \langle o_2, f_2 \rangle \in \{0, 1\} \times M^A$ . If we had had some notion of a distance defined on the state-space of our automaton, or speaking more formally a 1-pseudometric  $d : M \times M \rightarrow [0, 1]$ , then we can *lift* this notion of distance, to a distance on observations, given by the following:

$$d_{\{0,1\} \times M^A}(\langle o_1, f_1 \rangle, \langle o_2, f_2 \rangle) = \max\{d_2(o_1, o_2), \lambda \cdot \max_{a \in A} d(f_1(a), f_2(a))\} \quad \lambda \in ]0, 1[$$

<sup>1</sup> Lemma 6 is one of the intermediate results used in the proof of [2, Lemma 5.6] that was communicated to us by the authors of [2]. As this result was excluded in the mentioned paper, we incorporated it along with its proof for the sake of completeness.

The definition above involves  $d_2$ , the discrete metric on the set  $\{0, 1\}$ . One can observe that  $d_{\{0,1\} \times M^A}$  is again a 1-pseudometric, but this time defined on the set  $\{0, 1\} \times M^A$  instead. We now move on to showing how one could use this lifting in order to equip a state-space of automaton  $\mathcal{M}$  with a sensible notion of behavioural pseudometric.

**Behavioural pseudometric.** If one gave us a 1-metric  $d : M \times M \rightarrow [0, 1]$  on the state-space of the automaton, we could use our lifting to produce a new pseudometric  $\Phi_{\mathcal{M}}(d) : M \times M \rightarrow [0, 1]$  on the same set, which would calculate a distance between an arbitrary pair of states by first applying the output and transition derivatives to obtain a pair of observations and then by calculating the distance between them using the aforementioned lifting of pseudometric  $d$  to the pseudometric defined on the set  $\{0, 1\} \times M^A$ . Formally speaking, define a map  $\Phi_{\mathcal{M}} : D_M \rightarrow D_M$  on the lattice of 1-pseudometrics on  $M$  given by the following:

$$\Phi_{\mathcal{M}}(d)(m, m') = \max\{d_2(o_M(m), o_M(m')), \lambda \cdot \max_{a \in A} d(m_a, m'_a)\} \quad \lambda \in ]0, 1[$$

The construction above only tells us how to construct new pseudometrics on the state-space of the automaton out of existing ones, but does not give one to start with. It turns out, that the map  $\Phi_{\mathcal{M}}$  is a monotone mapping on the lattice of 1-pseudometrics on the set  $M$  [5, Lemma 6.1]. Because of that, one can use the Knaster-Tarski fixpoint theorem [36] and construct its least fixed point, explicitly given by  $d_{\mathcal{M}} = \inf\{d \mid d \in D_M \wedge \Phi_{\mathcal{M}}(d) \sqsubseteq d\}$ . Pseudometrics, which are fixpoints of  $\Phi_{\mathcal{M}}$  intuitively interact well with the automaton structure, as they satisfy the property that the distance between two states is the same as the distance between their observable behaviour calculated using the lifting. Taking the least such pseudometric satisfies several desirable properties [5] and thus we will call  $d_{\mathcal{M}}$  a behavioural pseudometric on the automaton  $\mathcal{M}$ . First of all, preserving automaton transitions also preserves behavioural distances.

► **Proposition 7.** *Let  $\mathcal{M} = (M, \langle o_M, t_M \rangle)$  and  $\mathcal{N} = (N, \langle o_N, t_N \rangle)$  be deterministic automata. If  $h : M \rightarrow N$  is a homomorphism, then it is also an isometric mapping between pseudometric spaces  $(M, d_{\mathcal{M}})$  and  $(N, d_{\mathcal{N}})$ .*

If we look at the final automaton on the set of all formal languages over a fixed finite alphabet  $A$ , then one can easily verify that the behavioural distance given by the least fixpoint construction precisely corresponds to the Equation (1) defining the shortest-distinguishing-word distance we stated in Section 1. In general, states of an arbitrary deterministic automaton characterised by the behavioural pseudometric to be in distance zero are language equivalent. When we look at  $d_{\mathcal{L}}$  defined on the states of the final automaton, whose state-space consists of formal languages, then the language equivalence corresponds to the equality of states. In other words  $d_{\mathcal{L}}$  becomes a metric space.

► **Lemma 8.** *Let  $\mathcal{M} = (M, \langle o_M, t_M \rangle)$  be an arbitrary deterministic automaton and let  $\mathcal{L} = (\mathcal{P}(A^*), \langle o_L, t_L \rangle)$  be a deterministic automaton structure on the set of all languages over an alphabet  $A$ .*

1.  $(\mathcal{P}(A^*), d_{\mathcal{L}})$  is a metric space.
2. For any  $m, m' \in M$ ,  $d_{\mathcal{M}}(m, m') = 0 \iff L_{\mathcal{M}}(m) = L_{\mathcal{M}}(m')$ .

## 4 Quantitative Axiomatisation

In order to provide a quantitative inference system for reasoning about the behavioural distance of languages denoted by regular expressions, we first recall the definition of quantitative equational theories from the existing literature [22, 2] following the notational conventions

from [2]. We then present our axiomatisation and demonstrate its soundness. The interesting thing about our axiomatisation is the lack of any fixpoint introduction rule. We show that in the case of quantitative analogue of equational logic [22] containing the infinitary rule capturing the notion of convergence, we can use our axioms to derive Salomaa's fixpoint rule from his axiomatisation of language equivalence of regular expressions [30].

**Quantitative equational theories.** Let  $\Sigma$  be an algebraic signature (in the sense of universal algebra [11]) consisting of operation symbols  $f_n \in \Sigma$  of arity  $n \in \mathbb{N}$ . If we write  $X$  for the countable set of *metavariables*, then  $\mathbb{T}(\Sigma, X)$  denotes a set of freely generated terms over  $X$  built from the signature  $\Sigma$ . As a notational convention, we will use letters  $t, s, u, \dots \in \mathbb{T}(\Sigma, X)$  to denote terms. By a *substitution* we mean a function of the type  $\sigma: X \rightarrow \mathbb{T}(\Sigma, X)$  allowing to replace metavariables with terms. Each substitution can be inductively extended to terms in a unique way by setting  $\sigma(f(t_1, \dots, t_n)) = f(\sigma(t_1), \dots, \sigma(t_n))$  for each operation symbol  $f_n \in \Sigma$  from the signature. We will write  $\mathcal{S}(\Sigma)$  for the set of all substitutions. Given two terms  $t, s \in \mathbb{T}(\Sigma, X)$  and a nonnegative rational number  $\varepsilon \in \mathbb{Q}^+$  denoting the distance between the terms, we call  $t \equiv_\varepsilon s$  a *quantitative equation (of type  $\Sigma$ )*. Notation-wise, we will write  $\mathcal{E}(\Sigma)$  to denote the set of all quantitative equations (of type  $\Sigma$ ) and we will use the capital Greek letters  $\Gamma, \Theta, \dots \subseteq \mathcal{E}(\Sigma)$  to denote the subsets of  $\mathcal{E}(\Sigma)$ . By a *deducibility relation* we mean a binary relation denoted  $\vdash \subseteq \mathcal{P}(\mathcal{E}(\Sigma)) \times \mathcal{E}(\Sigma)$ . Similarly, to the classical equational logic, we will use the following notational shorthands  $\Gamma \vdash t \equiv_\varepsilon s \iff (\Gamma, t \equiv_\varepsilon s) \in \vdash$  and  $\vdash t \equiv_\varepsilon s \iff \emptyset \vdash t \equiv_\varepsilon s$ . Furthermore, following the usual notational conventions, we will write  $\Gamma \vdash \Theta$  as a shorthand for the situation when  $\Gamma \vdash t \equiv_\varepsilon s$  holds for all  $t \equiv_\varepsilon s \in \Theta$ . To call  $\vdash$  a *quantitative deduction system (of type  $\Sigma$ )* it needs to satisfy the following rules of inference:

- (Refl)  $\vdash t \equiv_0 t$ ,
- (Symm)  $\{t \equiv_\varepsilon s\} \vdash s \equiv_\varepsilon t$ ,
- (Triang)  $\{t \equiv_\varepsilon u, u \equiv_{\varepsilon'} s\} \vdash t \equiv_{\varepsilon+\varepsilon'} s$ ,
- (Max)  $\{t \equiv_\varepsilon s\} \vdash t \equiv_{\varepsilon+\varepsilon'} s$ , for all  $\varepsilon' > 0$ ,
- (Cont)  $\{t \equiv_{\varepsilon'} s \mid \varepsilon' > \varepsilon\} \vdash t \equiv_\varepsilon s$ ,
- (NExp)  $\{t_1 \equiv_\varepsilon s_1, \dots, t_n \equiv_\varepsilon s_n\} \vdash f(t_1, \dots, t_n) \equiv_\varepsilon f(s_1, \dots, s_n)$ , for all  $f_n \in \Sigma$ ,
- (Subst) If  $\Gamma \vdash t \equiv_\varepsilon s$ , then  $\sigma(\Gamma) \vdash \sigma(t) \equiv_\varepsilon \sigma(s)$ , for all  $\sigma \in \mathcal{S}(\Sigma)$ ,
- (Cut) If  $\Gamma \vdash \Theta$  and  $\Theta \vdash t \equiv_\varepsilon s$ , then  $\Gamma \vdash t \equiv_\varepsilon s$ ,
- (Assum) If  $t \equiv_\varepsilon s \in \Gamma$ , then  $\Gamma \vdash t \equiv_\varepsilon s$ .

where  $\sigma(\Gamma) = \{\sigma(t) \equiv_\varepsilon \sigma(s) \mid t \equiv_\varepsilon s \in \Gamma\}$ . Finally, by a *quantitative equational theory* we mean a set  $\mathcal{U}$  of universally quantified *quantitative inferences*  $\{t_1 \equiv_{\varepsilon_1} s_1, \dots, t_n \equiv_{\varepsilon_n} s_n\} \vdash t \equiv_\varepsilon s$ , with *finitely many premises*, closed under  $\vdash$ -derivability.

**Quantitative algebras.** Quantitative equational theories lie on the syntactic part of the picture. On the semantic side, we have their models called *quantitative algebras*, defined as follows.

► **Definition 9** ([22, Definition 3.1]). *A quantitative algebra is a tuple  $\mathcal{A} = (A, \Sigma^{\mathcal{A}}, d^{\mathcal{A}})$ , such that  $(A, \Sigma^{\mathcal{A}})$  is an algebra for the signature  $\Sigma$  and  $(A, d^{\mathcal{A}})$  is an  $\infty$ -pseudometric such that for all operation symbols  $f_n \in \Sigma$ , for all  $1 \leq i \leq n$ ,  $a_i, b_i \in A$ ,  $d^{\mathcal{A}}(a_i, b_i) \leq \varepsilon$  implies  $d^{\mathcal{A}}(f^{\mathcal{A}}(a_1, \dots, a_n), f^{\mathcal{A}}(b_1, \dots, b_n)) \leq \varepsilon$ .*



Consider a quantitative algebra  $\mathcal{A} = (A, \Sigma^A, d^A)$ . Given an assignment  $\iota: X \rightarrow A$  of meta-variables from  $X$  to elements of carrier  $A$ , one can inductively extend it to  $\Sigma$ -terms  $t \in \mathbb{T}(\Sigma, X)$  in a unique way. We will abuse the notation and just write  $\iota(t)$  for the interpretation of the term  $t$  in quantitative algebra  $\mathcal{A}$ . We will say that  $\mathcal{A}$  *satisfies* the quantitative inference  $\Gamma \vdash t \equiv_\varepsilon s$ , written  $\Gamma \models_{\mathcal{A}} t \equiv_\varepsilon s$ , if for any assignment of the meta-variables  $\iota: X \rightarrow A$  it is the case that for all  $t' \equiv_{\varepsilon'} s' \in \Gamma$  we have that  $d^A(\iota(t'), \iota(s')) \leq \varepsilon'$  implies  $d^A(\iota(t), \iota(s)) \leq \varepsilon$ . Finally, we say that a quantitative algebra  $\mathcal{A}$  *satisfies* (or is a *model* of) the quantitative theory  $\mathcal{U}$ , if whenever  $\Gamma \vdash t \equiv_\varepsilon s \in \mathcal{U}$ , then  $\Gamma \models_{\mathcal{A}} t \equiv_\varepsilon s$ .

**Quantitative algebra of regular expressions.** From now on, let's focus on the signature  $\Sigma = \{0_0, 1_0, +_2, ;_2, (-)_1^*, (-)_1\} \cup \{a_0 \mid a \in A\}$ , where  $A$  is a finite alphabet. This signature consists of all operations of regular expressions. We can easily interpret all those operations in the set  $\text{Exp}$  of all regular expressions, using trivial interpretation functions eg.  $+^{\mathcal{B}}(e, f) = e + f$ , which interpret the operations by simply constructing the appropriate terms. Formally speaking, we can do this because the set  $\text{Exp}$  is the carrier of initial algebra [11] (free algebra over the empty set of generators) for the signature  $\Sigma$ .

To make this algebra into a quantitative algebra, we first equip the set  $\text{Exp}$  with a  $\infty$ -pseudometric, given by  $d^{\mathcal{B}}(e, f) = d_{\mathcal{L}}(\llbracket e \rrbracket, \llbracket f \rrbracket)$  for all  $e, f \in \text{Exp}$ . Recall that  $d_{\mathcal{L}}$  used in the definition above is a behavioural pseudometric on the final deterministic automaton carried by the set  $\mathcal{P}(A^*)$  of all formal languages over an alphabet  $A$ . In other words, we define the distance between arbitrary expressions  $e$  and  $f$  to be the distance between formal languages  $\llbracket e \rrbracket$  and  $\llbracket f \rrbracket$  calculated through the shortest-distinguishing-word metric. It turns out, that in such a situation all the interpretation functions of  $\Sigma$ -algebra structure on  $\text{Exp}$  are non-expansive with respect to the pseudometric defined above. In other words, we have that:

► **Lemma 10.**  $\mathcal{B} = (\text{Exp}, \Sigma^{\mathcal{B}}, d^{\mathcal{B}})$  is a quantitative algebra.

**Axiomatisation.** In order to talk about the quantitative algebra  $\mathcal{B}$  of the behavioural distance of regular expressions in an axiomatic way, we introduce the quantitative equational theory REG (Figure 1). The first group of axioms capture properties of the nondeterministic

	<b>Nondeterministic choice</b>		<b>Sequential composition</b>
(SL1)	$\vdash e + e \equiv_0 e,$	(1S)	$\vdash 1; e \equiv_0 e,$
(SL2)	$\vdash e + f \equiv_0 f + e,$	(S)	$\vdash e; (f; g) \equiv_0 (e; f); g,$
(SL3)	$\vdash (e + f) + g \equiv_0 e + (f + g),$	(S1)	$\vdash e; 1 \equiv_0 e,$
(SL4)	$\vdash e + 0 \equiv_0 e,$	(0S)	$\vdash 0; e \equiv_0 0,$
(SL5)	$\{e \equiv_\varepsilon g, f \equiv_{\varepsilon'} h\}$ $\vdash e + f \equiv_{\max(\varepsilon, \varepsilon')} g + h,$	(S0)	$\vdash e; 0 \equiv_0 0,$
		(D1)	$\vdash e; (f + g) \equiv_0 e; f + e; g,$
		(D2)	$\vdash (e + f); g \equiv_0 e; g + f; g,$
	<b>Loops</b>		<b>Behavioural pseudometric</b>
(Unroll)	$\vdash e^* \equiv_0 e; e^* + 1,$	(Top)	$\vdash e \equiv_1 f,$
(Tight)	$\vdash (e + 1)^* \equiv_0 e^*,$	( $\lambda$ -Pref)	$\{e \equiv_\varepsilon f\} \vdash a; e \equiv_{\varepsilon'} a; f, \text{ for } \varepsilon' \geq \lambda \cdot \varepsilon$

■ **Figure 1** Axioms of the quantitative equational theory REG for  $e, f, g \in \text{Exp}$  and  $a \in A$ .

choice operator  $+$  (SL1-SL5). The first four axioms (SL1-SL4) are the usual laws of semilattices with bottom element 0. (SL5) is a quantitative axiom allowing one to reason about distances between sums of expressions in terms of distances between expressions being summed. Moreover, (SL1-SL5) are axioms of so-called *Quantitative Semilattices with zero*, which

have been shown to axiomatise the Hausdorff metric [22]. The sequencing axioms (1S), (S1), (S) state that the set  $\text{Exp}$  of regular expressions has the structure of a monoid (with neutral element 1) with absorbent element 0 (0S), (S0). Additionally, (D1-D2) talk about interaction of the nondeterministic choice operator  $+$  with sequential composition. The loop axioms (Unroll) and (Tight) are directly inherited from Salomaa's axiomatisation of language equivalence of regular expressions [30]. (Unroll) axiom associates loops with their intuitive behaviour of choosing, at each step, between successful termination and executing the loop body once. (Tight) states that the loop whose body might instantly terminate, causing the next loop iteration to be executed immediately is provably equivalent to a different loop, whose body does not contain immediate termination. The last remaining group are behavioural pseudometric axioms. (Top) states that any two expressions are at most in distance one from each other. Finally, ( $\lambda$ -Pref) captures the fact that prepending the same letter to arbitrary expressions shrinks the distance between them by the factor of  $\lambda \in ]0, 1[$  (used in the definition of  $d^{\mathcal{B}}$ ). This axiom is adapted from the axiomatisation of discounted probabilistic bisimilarity distance [2].

Through a simple induction on the length of derivation, one can verify that indeed  $\mathcal{B}$  is a model of the quantitative theory REG.

► **Theorem 11.** (*Soundness*) *The quantitative algebra  $\mathcal{B} = (\text{Exp}, \Sigma^{\mathcal{B}}, d^{\mathcal{B}})$  is a model of the quantitative theory REG. In other words, for any  $e, f \in \text{Exp}$  and  $\varepsilon \in \mathbb{Q}^+$ , if  $\Gamma \vdash e \equiv_{\varepsilon} f \in \text{REG}$ , then  $\Gamma \models_{\mathcal{B}} e \equiv_{\varepsilon} f$*

**Proof.** By the structural induction on the judgement  $\Gamma \vdash e \equiv_{\varepsilon} f \in \text{REG}$ . (Subst), (Cut) and (Assum) deduction rules from classical logic hold immediately. The soundness of (Refl), (Symm), (Triang), (Cont) and (Max) follows from the fact that  $d^{\mathcal{B}}$  is a pseudometric. (NExp) follows from the fact that interpretations of symbols from the algebraic signature are nonexpansive (Lemma 10). Recall that  $d^{\mathcal{B}} = d_{\mathcal{L}} \circ (\llbracket - \rrbracket \times \llbracket - \rrbracket)$ . The soundness of (Top) follows from the fact that  $d_{\mathcal{L}}$  is a 1-pseudometric. Additionally, for all axioms in the form  $\vdash e \equiv_0 f$  it suffices to show that  $\llbracket e \rrbracket = \llbracket f \rrbracket$ . (SL1), (SL2), (SL3), (SL4), (1S), (S), (S1), (0S), (S0), (D1), (D2), (Unroll) and (Tight) are taken from Salomaa's axiomatisation of language equivalence of regular expressions [30] and thus both sides of those equations denote the same formal languages [43, Theorem 5.2]. For ( $\lambda$ -Pref) assume that the premise is satisfied in the model, that is  $d_{\mathcal{L}}(\llbracket e \rrbracket, \llbracket f \rrbracket) \leq \varepsilon$ . Let  $\varepsilon' \geq \lambda \cdot \varepsilon$ . We show the following:

$$\begin{aligned}
d^{\mathcal{B}}(a ; e, a ; f) &= d_{\mathcal{L}}(\llbracket a ; e \rrbracket, \llbracket a ; f \rrbracket) && \text{(Def. of } d^{\mathcal{B}}\text{)} \\
&= \Phi_{\mathcal{L}}(d_L)(\llbracket a ; e \rrbracket, \llbracket a ; f \rrbracket) && (d_L \text{ is a fixpoint of } \Phi_{\mathcal{L}}) \\
&= \max\{d_2(o_{\mathcal{L}}(a ; e), o_{\mathcal{L}}(a ; f)), \lambda \cdot \max_{a' \in A} d_{\mathcal{L}}(\llbracket a ; e \rrbracket_{a'}, \llbracket a ; f \rrbracket_{a'})\} \\
&= \lambda \cdot d_{\mathcal{L}}(\llbracket e \rrbracket, \llbracket f \rrbracket) && \text{(Def. of final automaton)} \\
&\leq \lambda \cdot \varepsilon \leq \varepsilon' && \text{(Assumptions)}
\end{aligned}$$

Finally, (SL5) is derivable from other axioms<sup>2</sup>. If  $\varepsilon = \max(\varepsilon, \varepsilon')$  then  $\{e \equiv_{\varepsilon} g\} \vdash e \equiv_{\max(\varepsilon, \varepsilon')} g$  holds by (Assum). If  $\varepsilon < \max(\varepsilon, \varepsilon')$ , then we can derive the quantitative judgement above using (Max). By a similar line of reasoning, we can show that  $\{f \equiv_{\varepsilon'} h\} \vdash f \equiv_{\max(\varepsilon, \varepsilon')} h$ . Finally, using (Cut) and (NExp), we can show that  $\{e \equiv_{\varepsilon} g, f \equiv_{\varepsilon'} h\} \vdash e + f \equiv_{\max(\varepsilon, \varepsilon')} g + h$  as desired. ◀

<sup>2</sup> We included (SL5) as an axiom to highlight the similarity of our inference system with axiomatisations of language equivalence of regular expressions [30, 19] containing the axioms of semilattices with bottom. In the previous work [22], (SL1 – SL5) are precisely the axioms of *Quantitative Semilattices with zero* axiomatising the Hausdorff distance.

We now revisit the example from Section 1. Recall that states marked as initial of the left and middle automata can be respectively represented as  $a^*$  and  $a + 1$ . The shortest word distinguishing languages representing those expressions is  $aa$ . If we fix  $\lambda = \frac{1}{2}$ , then  $d_{\mathcal{L}}(\llbracket a^* \rrbracket, \llbracket a + 1 \rrbracket) = \frac{1}{4} = (\frac{1}{2})^{|aa|}$ . We can derive this distance through the means of axiomatic reasoning using the quantitative equational theory REG in the following way:

► **Example 12.**

$$\begin{array}{ll}
\vdash a^* \equiv_1 0 & \text{(Top)} \\
\vdash a; a^* \equiv_{\frac{1}{2}} a; 0 & (\lambda\text{-Pref}) \\
\vdash a; a^* + 1 \equiv_{\frac{1}{2}} a; 0 + 1 & (\vdash 1 \equiv_0 1 \text{ and SL5}) \\
\vdash a^* \equiv_{\frac{1}{2}} 1 & \text{(Triang, Unroll, S0 and SL4)} \\
\vdash a; a^* \equiv_{\frac{1}{4}} a; 1 & (\lambda\text{-Pref}) \\
\vdash a; a^* + 1 \equiv_{\frac{1}{4}} a; 1 + 1 & (\vdash 1 \equiv_0 1 \text{ and SL5}) \\
\vdash a^* \equiv_{\frac{1}{4}} a + 1 & \text{(Triang, Unroll and S1)}
\end{array}$$

**(The lack of) the fixpoint axiom.** Traditionally, completeness of inference systems for behavioural equivalence of languages of expressions featuring recursive constructs such as Kleene star or  $\mu$ -recursion [24] rely crucially on fixpoint introduction rules. Those allow showing that an expression is provably equivalent to a looping construct if it exhibits some form of self-similarity, typically subject to productivity constraints. As an illustration, Salomaa's axiomatisation of language equivalence of regular expressions incorporates the following inference rule:

$$\frac{g \equiv e; g + f \quad \epsilon \notin \llbracket e \rrbracket}{g \equiv e^*; f} \quad (2)$$

The side condition on the right states that the loop body is *productive*, that is a deterministic automaton corresponding to an expression  $e$  cannot immediately reach acceptance without performing any transitions. This is simply equivalent to the language  $\llbracket e \rrbracket$  not containing the empty word. It would be reasonable for one to expect REG to contain a similar rule to be complete, especially since it should be able to prove language equivalence of regular expressions (by proving that they are in distance zero from each other). Furthermore, all axioms of Salomaa except Equation (2) are contained in REG as rules for distance zero.

It turns out that in the presence of the infinitary continuity (**Cont**) rule of quantitative deduction systems and the ( $\lambda$ -Pref) axiom of REG, the Salomaa's inference rule (Equation (2)) becomes a derivable fact for distance zero. First of all, one can show that ( $\lambda$ -Pref) can be generalised from prepending single letters to prepending any regular expression satisfying the side condition from Equation (2).

► **Lemma 13.** *Let  $e, f, g \in \text{Exp}$ , such that  $\epsilon \notin \llbracket e \rrbracket$ . Then,  $\{f \equiv_{\epsilon} g\} \vdash e; f \equiv_{\epsilon'} e; g$  is derivable using the axioms of REG for all  $\epsilon' \geq \lambda \cdot \epsilon$ .*

With the above lemma in hand, one can inductively show that if  $g \equiv_0 e; g + f$  and  $\epsilon \notin \llbracket e \rrbracket$ , then  $g$  gets arbitrarily close to  $e^*; f$ . Intuitively, the more we unroll the loop in  $e^*; f$  using (Unroll) and the more we unroll the definition of  $g$ , then the closer both expressions become.

► **Lemma 14.** *Let  $e, f, g \in \text{Exp}$ , such that  $e \notin \llbracket e \rrbracket$  and let  $n \in \mathbb{N}$ . Then,  $\{g \equiv_0 e; g + f\} \vdash g \equiv_\varepsilon e^*; f$  is derivable using the axioms of REG for all  $\varepsilon \geq \lambda^n$ .*

Having the result above, we can now use the infinitary (Cont) rule capturing the limiting property of decreasing chain of overapproximations to the distance and show the derivability of Salomaa's inference rule.

► **Lemma 15.** *Let  $e, f, g \in \text{Exp}$ , such that  $e \notin \llbracket e \rrbracket$ . Then,  $\{g \equiv_0 e; g + f\} \vdash g \equiv_0 e^*; f$  is derivable using the axioms of REG.*

**Proof.** To deduce that  $\vdash g \equiv_0 e^*; f$  using (Cont) it suffices to show that  $\vdash g \equiv_\varepsilon e^*; f$  for all  $\varepsilon > 0$ . To do so, pick an arbitrary  $\varepsilon > 0$  and let  $N = \lceil \log_\lambda \varepsilon \rceil$ . Observe that  $\lambda^N = \lambda^{\lceil \log_\lambda \varepsilon \rceil} \leq \lambda^{\log_\lambda \varepsilon} = \varepsilon$ . Because of Lemma 14 we have that  $\vdash g \equiv_\varepsilon e^*; f$ , which completes the proof. ◀

## 5 Completeness

We now move on to the central result of this paper, which is the completeness of REG with respect to the shortest-distinguishing-word metric on languages denoting regular expressions. We use the strategy from the proof of completeness of quantitative axiomatisation of probabilistic bisimilarity distance [2]. It turns out that the results from [2] rely on properties that are not unique to the Kantorovich/Wassertstein lifting and can be also established for instances of the abstract coalgebraic framework [5].

The heart of our argument relies on the fact that the distance between languages denoting regular expressions can be calculated in a simpler way than applying the Knaster-Tarski fixpoint theorem while looking at the infinite-state final automaton of all formal languages over some fixed alphabet. In particular, regular expressions denote the behaviour of finite-state deterministic automata. Since automata homomorphisms are non-expansive mappings, the distance between languages  $\llbracket e \rrbracket$  and  $\llbracket f \rrbracket$  of some arbitrary regular expressions  $e, f \in \text{Exp}$  is the same as the distance between states in some DFA whose languages corresponds to  $\llbracket e \rrbracket$  and  $\llbracket f \rrbracket$ . To be precise, we will look at the finite subautomaton  $\langle [e]_{\equiv}, [f]_{\equiv} \rangle_{\mathcal{Q}}$  of the  $\equiv$  quotient of the Brzozowski automaton. The reason we care about deterministic finite automata is that it turns out that one can calculate the behavioural distance between two states through iterative approximation from above, which can be also derived axiomatically using the (Cont) rule of quantitative deduction systems. We start by showing how this simplification works and then we move on to establishing completeness.

**Behavioural distance on finite-state automata.** Consider a deterministic automaton  $\mathcal{M} = (M, \langle o_M, t_M \rangle)$ . The least fixpoint of a monotone endomap  $\Phi_{\mathcal{M}} : D_M \rightarrow D_M$  on the complete lattice of 1-pseudometrics on the set  $M$  results in  $d_{\mathcal{M}}$ , which is a behavioural pseudometric on the states of the automaton  $\mathcal{M}$ . It is noteworthy that  $\Phi_{\mathcal{M}}$  exhibits two generic properties. Firstly,  $\Phi_{\mathcal{M}}$  behaves well within the Banach space structure defined by the supremum norm.

► **Lemma 16.**  *$\Phi_{\mathcal{M}} : D_M \rightarrow D_M$  is nonexpansive with respect to the supremum norm. In other words, for all  $d, d' \in D_M$  we have that  $\|\Phi_{\mathcal{M}}(d') - \Phi_{\mathcal{M}}(d)\| \leq \|d' - d\|$ .*

**Proof.** We can safely assume that  $d \sqsubseteq d'$ , as other case will be symmetric. It suffices to show that for all  $m, m' \in M$ ,  $\Phi_{\mathcal{M}}(d')(m, m') - \Phi_{\mathcal{M}}(d)(m, m') \leq \|d' - d\|$ . First, let's consider the case when  $o_M(m) \neq o_M(m')$  and hence  $d_2(m, m') = 1$ . In such a scenario, it holds that  $\Phi_{\mathcal{M}}(d')(m, m') - \Phi_{\mathcal{M}}(d)(m, m') = 0 \leq \|d' - d\|$ . From now on, we will assume that  $o_M(m) = o_M(m')$  and hence  $d_2(m, m') = 0$ . We have the following

$$\begin{aligned}
\Phi_{\mathcal{M}}(d')(m, m') - \Phi_{\mathcal{M}}(d)(m, m') &= \lambda \cdot \max_{a \in A} d'(m_a, m'_a) - \lambda \cdot \max_{a \in A} d(m_a, m'_a) \\
&= \lambda \cdot \left( \max_{a \in A} d'(m_a, m'_a) - \max_{a \in A} d(m_a, m'_a) \right) \\
&\leq \lambda \cdot \left( \max_{a \in A} \{d'(m_a, m'_a) - d(m_a, m'_a)\} \right) \\
&\leq \lambda \cdot \sup_{n, n' \in M} \{d'(n, n') - d(n, n')\} \\
&= \lambda \cdot \|d' - d\| \leq \|d' - d\| \quad \blacktriangleleft
\end{aligned}$$

Secondly, it turns out that  $\Phi_{\mathcal{M}}$  has only one fixpoint. This means that if we want to calculate  $d_{\mathcal{M}}$  it suffices to look at any fixpoint of  $\Phi_{\mathcal{M}}$ . This will enable a simpler characterisation, than the one given by the Knaster-Tarski fixpoint theorem.

► **Lemma 17.**  $\Phi_{\mathcal{M}}$  has a unique fixed point.

**Proof.** Let  $d, d' \in D_M$  be two fixed points of  $\Phi_{\mathcal{M}}$ , that is  $\Phi_{\mathcal{M}}(d) = d$  and  $\Phi_{\mathcal{M}}(d') = d'$ . We can safely assume that  $d \sqsubseteq d'$ , as the other case is symmetric. We wish to show that  $d = d'$  and to do so we will use proof by contradiction.

Assume that  $d \neq d'$ , and hence there exist  $m, m' \in M$ , such that  $d(m, m') < d'(m, m')$  and  $\|d' - d\| = d'(m, m') - d(m, m') \neq 0$ . First, consider the case when  $o_M(m) \neq o_M(m')$ . In such a case both  $d(m, m')$  and  $d'(m, m')$  are equal to 1 and hence  $\|d' - d\| = 0$ , which leads to contradiction. From now, we can safely assume that  $o_M(m) = o_M(m')$ . Through an identical line of reasoning to the proof of Lemma 16, we can show that  $\|\Phi_{\mathcal{M}}(d') - \Phi_{\mathcal{M}}(d)\| \leq \lambda \cdot \|d' - d\|$ . Since both  $d$  and  $d'$  are fixed points, this would mean that  $\|d' - d\| \leq \lambda \cdot \|d' - d\|$ . Since  $\lambda \in ]0, 1[$ , this would imply that  $\|d' - d\| = 0$  leading again to contradiction. ◀

In particular, we will rely on the characterisation given by the Kleene fixpoint theorem [31, Theorem 2.8.5], which allows to obtain the greatest fixpoint of an endofunction on the lattice as the infimum of the decreasing sequence of finer approximations obtained by repeatedly applying the function to the top element of the lattice.

► **Theorem 18 (Kleene fixpoint theorem).** Let  $(X, \sqsubseteq)$  be a complete lattice with a top element  $\top$  and  $f : X \rightarrow X$  an endofunction that is  $\omega$ -cocontinuous or in other words for any decreasing chain  $\{x_i\}_{i \in \mathbb{N}}$  it holds that  $\inf_{i \in \mathbb{N}} \{f(x_i)\} = f(\inf_{i \in \mathbb{N}} x_i)$ . Then,  $f$  possesses a greatest fixpoint, given by  $\text{gfp}(f) = \inf_{i \in \mathbb{N}} \{f^{(i)}(\top)\}$  where  $f^{(n)}$  denotes  $n$ -fold self-composition of  $f$  given inductively by  $f^{(0)}(x) = x$  and  $f^{(n+1)}(x) = f^{(n)}(f(x))$  for all  $x \in X$ .

The theorem above requires the endomap to be  $\omega$ -cocontinuous. Luckily, it is the case for  $\Phi_{\mathcal{M}}$  if we restrict our attention to DFA. To show that, we directly follow the line of reasoning from [2, Lemma 5.6] generalising the similar line of reasoning for  $\omega$ -continuity from [37, Theorem 1]. First, using Lemma 6 we show that decreasing chains of pseudometrics over a finite set converge to their infimum. That result is a minor re-adaptation of [37, Theorem 1] implicitly used in [2, Lemma 5.6].

► **Lemma 19.** Let  $\{d_i\}_{i \in \mathbb{N}}$  be an infinite descending chain in the lattice  $(D_X, \sqsubseteq)$ , where  $X$  is a finite set. The sequence  $\{d_i\}_{i \in \mathbb{N}}$  converges (in the sense of convergence in the Banach space) to  $d(x, y) = \inf_{i \in \mathbb{N}} d_i(x, y)$ .

## 149:14 Quantitative Axiomatisation of Regular Expressions

**Proof.** Let  $\varepsilon > 0$  and let  $x, y \in X$ . Since  $d(x, y) = \inf_{i \in \mathbb{N}} d_i(x, y)$  there exists an index  $m_{x,y} \in \mathbb{N}$  such that for all  $n \geq m_{x,y}$ ,  $|d_n(x, y) - d(x, y)| < \varepsilon$ . Now, let  $N = \max\{m_{x,y} \mid x, y \in X\}$ . This is well-defined because  $X$  is finite. Therefore, for all  $n \geq N$  and  $x, y \in X$ ,  $|d_n(x, y) - d(x, y)| < \varepsilon$  and hence  $\|d_n - d\| < \varepsilon$ . ◀

We can now use the above to show the desired property, by re-adapting [37, Theorem 1].

► **Lemma 20.** *If  $\mathcal{M}$  is a deterministic finite automaton, then  $\Phi_{\mathcal{M}}$  is  $\omega$ -cocontinuous.*

**Proof.** By Lemma 19, the chain  $\{d_i\}_{i \in \mathbb{N}}$  converges to  $\inf_{i \in \mathbb{N}} d_i$ . Since  $\Phi_{\mathcal{M}}$  is nonexpansive (Lemma 16) it is also continuous (in the sense of the Banach space continuity) and therefore  $\{\Phi_{\mathcal{M}}(d_i)\}_{i \in \mathbb{N}}$  converges to  $\Phi_{\mathcal{M}}(\inf_{i \in \mathbb{N}} d_i)$ . Recall that  $\Phi_{\mathcal{M}}$  is monotone, which makes  $\{\Phi_{\mathcal{M}}(d_i)\}_{i \in \mathbb{N}}$  into a chain, which by Lemma 6 and Lemma 19 converges to  $\inf_{i \in \mathbb{N}} \{\Phi_{\mathcal{M}}(d_i)\}$ . Since limit points are unique,  $\inf_{i \in \mathbb{N}} \{\Phi_{\mathcal{M}}(d_i)\} = \Phi_{\mathcal{M}}(\inf_{i \in \mathbb{N}} d_i)$ . ◀

We can combine the preceding results and provide a straightforward characterisation of the distance between languages represented by arbitrary regular expressions, denoted as  $e, f \in \text{Exp}$ . Utilising a simple argument based on Proposition 7, which asserts that automata homomorphisms are nonexpansive, one can demonstrate that the distance between  $\llbracket e \rrbracket$  and  $\llbracket f \rrbracket$  in the final automaton is equivalent to the distance between  $[e]_{\equiv}$  and  $[f]_{\equiv}$  in  $\langle [e]_{\equiv}, [f]_{\equiv} \rangle_{\mathcal{Q}}$ . This is the least subautomaton of  $\mathcal{Q}$  that contains the derivatives (modulo  $\equiv$ ) reachable from  $[e]_{\equiv}$  and  $[f]_{\equiv}$ . Importantly, this automaton is finite (Lemma 3), allowing us to apply the Kleene fixpoint theorem to calculate the distance.

Let  $\Psi_{e,f}^{(0)}$  denote the discrete metric on the set  $\langle [e]_{\equiv}, [f]_{\equiv} \rangle_{\mathcal{Q}}$  (the top element of the lattice of pseudometrics over that set). Define  $\Psi_{e,f}^{(n+1)} = \Phi_{\langle [e]_{\equiv}, [f]_{\equiv} \rangle_{\mathcal{Q}}}(\Psi_{e,f}^{(n)})$ . Additionally, leveraging the fact that infima of decreasing chains are calculated pointwise (Lemma 6), we can conclude with the following:

► **Lemma 21.** *For all  $e, f \in \text{Exp}$ , the underlying pseudometric of the quantitative algebra  $\mathcal{B}$  can be given by  $d^{\mathcal{B}}(e, f) = \inf_{i \in \mathbb{N}} \left\{ \Psi_{e,f}^{(i)}([e]_{\equiv}, [f]_{\equiv}) \right\}$*

In simpler terms, we have demonstrated that the behavioural distance between a pair of arbitrary regular expressions can be calculated as the infimum of decreasing approximations of the actual distance from above. Alternatively, one could calculate the same distance as the supremum of increasing approximations from below using the Kleene fixpoint theorem for the least fixpoint. We chose the former approach because our proof of completeness relies on the (Cont) rule of quantitative deduction systems. This rule essentially states that to prove two terms are at a specific distance, we should be able to prove that for all approximations of that distance from above. This allows us to replicate the fixpoint calculation through axiomatic reasoning.

**Completeness result.** We start by recalling that regular expressions satisfy a certain decomposition property, stating that each expression can be reconstructed from its small-step semantics, up to  $\equiv_0$ . This property, often referred to as the fundamental theorem of Kleene Algebra/regular expressions (in analogy with the fundamental theorem of calculus and following the terminology of Rutten [27] and Silva [33]) is useful in further steps of the proof of completeness.

► **Theorem 22 (Fundamental Theorem).** *For any  $e \in \text{Exp}$ ,  $\vdash e_i \equiv_0 \sum_{a \in A} a; (e_i)_a + o_{\mathcal{R}}(e_i)$  is derivable using the axioms of REG.*

The theorem above makes use of the  $n$ -ary generalised sum operator, which is well defined because of (SL1-SL4) axioms of REG. Let's now say that we are interested in the distance between some expressions  $e, f \in \text{Exp}$ . As mentioned before, we will rely on  $\langle [e]_{\equiv}, [f]_{\equiv} \rangle_{\mathcal{Q}}$ , the least subautomaton of the  $\equiv$  quotient of the Brzozowski automaton containing states reachable from  $[e]_{\equiv}$  and  $[f]_{\equiv}$ . Recall that by Lemma 3 its state space is finite. It turns out that the approximations from above (from Lemma 21) to the distance between any pair of states in that automaton can be derived through the means of axiomatic reasoning.

► **Lemma 23.** *Let  $e, f \in \text{Exp}$  be arbitrary regular expressions and let  $[g]_{\equiv}, [h]_{\equiv} \in \langle [e]_{\equiv}, [f]_{\equiv} \rangle_{\mathcal{Q}}$ . For all  $i \in \mathbb{N}$ , and  $\varepsilon \geq \Psi_{e,f}^{(i)}([g]_{\equiv}, [h]_{\equiv})$ , one can derive  $\vdash g \equiv_{\varepsilon} h$  using the axioms of REG.*

**Proof.** We proceed by induction on  $i$ . For the base case, observe that  $\Psi_{e,f}^{(0)}$  is the discrete 1-pseudometric on the set  $\langle [e]_{\equiv}, [f]_{\equiv} \rangle_{\mathcal{Q}}$  such that  $\Psi_{e,f}^{(0)}([g]_{\equiv}, [h]_{\equiv}) = 0$  if and only if  $g \equiv h$ , or otherwise  $\Psi_{e,f}^{(0)}([g]_{\equiv}, [h]_{\equiv}) = 1$ . In the first case, we immediately have that  $g \equiv_0 h$ , because  $\equiv$  is contained in distance zero axioms of REG. In the latter case, we can just use (Top), to show that  $g \equiv_1 h$ . Then, in both cases, we can apply (Max) to obtain  $\vdash g \equiv_{\varepsilon} h$ , since  $\varepsilon \geq \Psi_{e,f}^{(0)}([g]_{\equiv}, [h]_{\equiv})$ . For the induction step, let  $i = j + 1$  and derive the following:

$$\begin{aligned} \varepsilon \geq \Psi_{e,f}^{(j+1)}([g]_{\equiv}, [h]_{\equiv}) &\iff \varepsilon \geq \Phi_{\langle [e]_{\equiv}, [f]_{\equiv} \rangle_{\mathcal{Q}}}(\Psi_{e,f}^{(j)}([g]_{\equiv}, [h]_{\equiv})) \quad (\text{Def. of } \Psi_{e,f}^{j+1}) \\ &\iff \varepsilon \geq \max \left\{ d_2(o_{\mathcal{Q}}([g]_{\equiv}), o_{\mathcal{Q}}([h]_{\equiv})), \lambda \cdot \max_{a \in A} \left\{ \Psi_{e,f}^{(j)}([g]_{\equiv_a}, [h]_{\equiv_a}) \right\} \right\} \quad (\text{Def. of } \Phi) \\ &\iff \varepsilon \geq \max \left\{ d_2(o_{\mathcal{R}}(g), o_{\mathcal{R}}(h)), \lambda \cdot \max_{a \in A} \left\{ \Phi_{\langle [e]_{\equiv}, [f]_{\equiv} \rangle_{\mathcal{Q}}}([g]_a, [h]_a) \right\} \right\} \\ &\hspace{15em} (\text{Def. of } \mathcal{Q}) \\ &\iff \varepsilon \geq d_2(o_{\mathcal{R}}(g), o_{\mathcal{R}}(h)) \text{ and for all } a \in A, \varepsilon \cdot \lambda^{-1} \geq \Psi_{e,f}^{(j)}([g]_a, [h]_a) \end{aligned}$$

Firstly, since  $d_2$  is the discrete 1-pseudometric on the set  $\{0, 1\}$ , we can use (Refl) or (Top) depending on whether  $o_{\mathcal{R}}(g) = o_{\mathcal{R}}(h)$  and then apply (Max) to derive  $\vdash o_{\mathcal{R}}(g) \equiv_{\varepsilon} o_{\mathcal{R}}(h)$ .

Let  $a \in A$ . We will show that  $\vdash a; (g)_a \equiv_{\varepsilon} a; (h)_a$ . Since  $\varepsilon \cdot \lambda^{-1}$  is not guaranteed to be rational, we cannot immediately apply the induction hypothesis. Instead, we rely on (Cont) rule. First, pick an arbitrary rational  $\varepsilon'$  strictly greater than  $\varepsilon$  and fix  $\{r_n\}_{n \in \mathbb{N}}$  to be any decreasing sequence of rationals that converges to  $\lambda^{-1}$ . Let  $r_N$  be an element of that sequence such that  $\varepsilon' \geq \lambda \cdot \varepsilon \cdot r_N$ . It is always possible to pick such element because  $\{\lambda \cdot r_n\}_{n \in \mathbb{N}}$  is a decreasing sequence that converges to 1 and  $\varepsilon' > \varepsilon$ . Since  $\varepsilon \cdot r_N \geq \varepsilon \cdot \lambda^{-1}$  and  $\varepsilon \cdot r_N \in \mathbb{Q}^+$ , we can use the induction hypothesis and derive  $\vdash (g)_a \equiv_{\varepsilon \cdot r_N} (h)_a$ . Then, by ( $\lambda$ -Pref) axiom we have that  $\vdash a; (g)_a \equiv_{\varepsilon'} a; (h)_a$ . Since we have shown it for arbitrary  $\varepsilon' > \varepsilon$ , by (Cont) rule we have that  $\vdash a; (g)_a \equiv_{\varepsilon} a; (h)_a$ . Using (SL5), we can combine all subexpressions involving the output and transition derivatives into the following:

$$\vdash \sum_{a \in A} a; (g)_a + o_{\mathcal{R}}(g) \equiv_{\varepsilon} \sum_{a \in A} a; (h)_a + o_{\mathcal{R}}(h)$$

Since both sides are normal forms of  $g$  and  $h$  existing because of Theorem 22, we can apply (Triang) on both sides and obtain  $\vdash g \equiv_{\varepsilon} h$  thus completing the proof. ◀

At this point, we have done all the hard work, and establishing completeness involves a straightforward argument that utilises the (Cont) rule and the lemma above.

► **Theorem 24 (Completeness).** *For any  $e, f \in \text{Exp}$  and  $\varepsilon \in \mathbb{Q}^+$ , if  $\models_{\mathcal{B}} e \equiv_{\varepsilon} f$ , then  $\vdash e \equiv_{\varepsilon} f \in \text{REG}$*

**Proof.** Assume that  $\models_{\mathcal{B}} e \equiv_{\varepsilon} f$ , which by the definition of  $\models_{\mathcal{B}}$  is equivalent to  $d^{\mathcal{B}}(e, f) \leq \varepsilon$ . In order to use (Cont) axiom to derive  $\vdash e \equiv_{\varepsilon} f$ , we need to be able to show  $\vdash e \equiv_{\varepsilon'} f$  for all  $\varepsilon' > \varepsilon$ . Because of iterative characterisation of  $d^{\mathcal{B}}$  from Lemma 21, we have that  $\inf_{i \in \mathbb{N}} \{\Psi_{e,f}^{(i)}([e]_{\equiv}, [f]_{\equiv})\} < \varepsilon'$ . Since  $\varepsilon'$  is strictly above the infimum of the descending chain of approximants, there exists a point  $i \in \mathbb{N}$ , such that  $\varepsilon' > \Psi_{e,f}^{(i)}([e]_{\equiv}, [f]_{\equiv})$ . We can show this by contradiction.

Assume that for all  $i \in \mathbb{N}$ ,  $\varepsilon' \leq \Psi_{e,f}^{(i)}([e]_{\equiv}, [f]_{\equiv})$ . This would make  $\varepsilon'$  into the lower bound of the chain  $\left\{ \Psi_{e,f}^{(i)}([e]_{\equiv}, [f]_{\equiv}) \right\}_{i \in \mathbb{N}}$  and in such a case  $\varepsilon'$  would be less than or equal to the infimum of that chain, which by assumption is less than or equal to  $\varepsilon$ . By transitivity, we could obtain  $\varepsilon' \leq \varepsilon$ . Since  $\varepsilon' > \varepsilon$ , by antisymmetry we could derive that  $\varepsilon' = \varepsilon$ , which would lead to the contradiction.

Using the fact shown above, we can use Lemma 23 to obtain  $\vdash e \equiv_{\varepsilon'} f \in \text{REG}$  for any  $\varepsilon' > \varepsilon$ , which completes the proof.  $\blacktriangleleft$

In simpler terms, the (Cont) rule enables us to demonstrate that two terms are at a specific distance by examining all strict overapproximations of that distance. Due to the iterative nature outlined in Lemma 21, this implies that we only need to consider finite approximants used in the Kleene fixpoint theorem. Each of those finite approximants can be axiomatically derived using Lemma 23.

## 6 Discussion

We have presented a sound and complete axiomatisation of the shortest-distinguishing word distance between languages representing regular expressions through a quantitative analogue of equational logic [22]. Before our paper, only axiomatised behavioural distances of probabilistic/weighted transition systems existed, through (variants of) the Kantorovich/Wasserstein lifting [21, 12, 4, 2, 1], while we looked at a behavioural distance obtained through a more general coalgebraic framework [5]. Outside of the coalgebra community, the shortest-distinguishing word distance and its variants also appear in the model checking [20] and in the automata learning [14] literature.

We have followed the strategy for proving completeness from [2]. The interesting insight about that strategy is that it relies on properties that are not exclusive to distances obtained through the Kantorovich/Wasserstein lifting and can be established for notions of behavioural distance for other kinds of transition systems stemming from the coalgebraic framework. In particular, one needs to show that the monotone map on the lattice of pseudometrics used in defining the distance of finite-state systems is nonexpansive with respect to the sup norm (and hence  $\omega$ -cocontinuous) and has a unique fixpoint, thus allowing to characterise the behavioural distance as the greatest fixpoint obtained through the Kleene fixpoint theorem. This point of view allows one to reconstruct the fixpoint calculation in terms of axiomatic manipulation involving the (Cont) rule, eventually leading to completeness.

We have additionally observed that in the presence of the infinitary (Cont) rule and the ( $\lambda$ -Pref) axiom, there is no need for a fixpoint rule, which is common place in all axiomatisations of regular expressions but also in other work on distances. In particular, the previous work on axiomatising a discounted probabilistic bisimilarity distance from [2] includes both ( $\lambda$ -Pref) and the fixpoint introduction rule, but its proof of completeness [2, Theorem 6.4] does not involve the fixpoint introduction rule at any point. We are highly confident that in the case of that axiomatisation, the fixpoint introduction rule could be derived from other axioms in a similar fashion to the way we derived Salomaa's rule for introducing the Kleene star [30]. Additionally, we are interested in how much this argument relates to the recent study of fixpoints in quantitative equational theories [23].



Moreover, the axiomatisations from [1, 2] rely on a slight modification of quantitative equational theories, which drop the requirement of all operations from the signature to be nonexpansive. This is dictated by the fact that the interpretation of  $\mu$ -recursion in Stark and Smolka's probabilistic process algebra [35] can increase the behavioural distances in the case of unguarded recursion, while in regular expressions recursive behaviour is introduced through Kleene's star, whose interpretation is non-expansive with respect to the shortest-distinguishing-word distance. This allowed us to fit instantly into the original framework of quantitative equational theories. The earlier work [4] focusing on Markov processes [8] also relies on quantitative equational theories, but its syntax does not involve any recursive primitives. Instead, the recursive behaviour is introduced through Cauchy completion of a pseudometric induced by the axioms. The earliest works on axiomatising behavioural distances of weighted [21] and probabilistic [12] transition systems, studied before the introduction of quantitative equational theories, rely on ad-hoc inference systems that cannot be easily generalised.

The pioneering works [13, 39, 37, 40, 38, 3] laid foundations for behavioural (pseudo)metrics of various flavours of probabilistic transition systems. The coalgebraic point of view [5] allowed to generalise these ideas to a wide range of transition systems by moving from the Kantorovich/Wasserstein lifting to the abstract setting of lifting endofunctors from the category of sets to the category of pseudometric spaces. Building upon this theory, further lines of work were dedicated to asymmetric distances (called hemimetrics) through the theory of quasi-lax liftings [44], fuzzy analogues of Hennessy-Milner logic characterising behavioural distance [17, 6], fibrational generalisations involving quantale-enriched categories [7], up-to techniques allowing for efficient approximation of behavioural distances [9] and quantitative analogues of van Glabbe's linear-time branching-time spectrum [15].

In this paper, we have focused on the simplest and most intuitive instantiation of the coalgebraic framework in the case of deterministic automata, but the natural next step would be to generalise our results to a wider class of transition systems. A good starting point could be to consider coalgebras for *polynomial* endofunctors, in the fashion of the framework of *Kleene Coalgebra* [33]. Alternatively, it would be interesting to look at recent work on a family of process algebras parametric on an equational theory representing the branching constructs [32] and study its generalisations to quantitative equational theories. A related and interesting avenue for future work are equational axiomatisations of behavioural equivalence of Guarded Kleene Algebra with Tests (GKAT) [34, 32] and its probabilistic extension (ProbGKAT) [29], whose completeness results rely on a powerful uniqueness of solutions axiom (UA). The soundness of UA in both cases is shown through an involved argument relying on equipping the transition systems giving the operational semantics with a form of behavioural distance and showing that recursive specifications describing finite-state systems correspond to certain contractive mappings. It may be more sensible, particularly for ProbGKAT to consider quantitative axiomatisations in the first place and give the proofs of completeness through the pattern explored in this paper.

---

## References

- 1 Giorgio Bacci, Giovanni Bacci, Kim G. Larsen, and Radu Mardare. Complete Axiomatization for the Total Variation Distance of Markov Chains. In Sam Staton, editor, *Proceedings of the Thirty-Fourth Conference on the Mathematical Foundations of Programming Semantics, MFPS 2018, Dalhousie University, Halifax, Canada, June 6-9, 2018*, volume 341 of *Electronic Notes in Theoretical Computer Science*, pages 27–39. Elsevier, 2018. doi:10.1016/J.ENTCS.2018.03.014.

- 2 Giorgio Bacci, Giovanni Bacci, Kim G. Larsen, and Radu Mardare. A Complete Quantitative Deduction system for the Bisimilarity Distance on Markov Chains. *Log. Methods Comput. Sci.*, 14(4), 2018. doi:10.23638/LMCS-14(4:15)2018.
- 3 Giorgio Bacci, Giovanni Bacci, Kim G. Larsen, and Radu Mardare. Converging from branching to linear metrics on Markov chains. *Math. Struct. Comput. Sci.*, 29(1):3–37, 2019. doi:10.1017/S0960129517000160.
- 4 Giorgio Bacci, Radu Mardare, Prakash Panangaden, and Gordon D. Plotkin. An Algebraic Theory of Markov Processes. In Anuj Dawar and Erich Grädel, editors, *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2018, Oxford, UK, July 09–12, 2018*, pages 679–688. ACM, 2018. doi:10.1145/3209108.3209177.
- 5 Paolo Baldan, Filippo Bonchi, Henning Kerstan, and Barbara König. Coalgebraic Behavioral Metrics. *Log. Methods Comput. Sci.*, 14(3), 2018. doi:10.23638/LMCS-14(3:20)2018.
- 6 Harsh Beohar, Sebastian Gurke, Barbara König, and Karla Messing. Hennessy-Milner Theorems via Galois Connections. In Bartek Klin and Elaine Pimentel, editors, *31st EACSL Annual Conference on Computer Science Logic, CSL 2023, February 13–16, 2023, Warsaw, Poland*, volume 252 of *LIPICs*, pages 12:1–12:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPICs.CSL.2023.12.
- 7 Harsh Beohar, Sebastian Gurke, Barbara König, Karla Messing, Jonas Forster, Lutz Schröder, and Paul Wild. Expressive Quantale-valued Logics for Coalgebras: an Adjunction-based Approach. *CoRR*, abs/2310.05711, 2023. doi:10.48550/arXiv.2310.05711.
- 8 Richard Blute, Josée Desharnais, Abbas Edalat, and Prakash Panangaden. Bisimulation for Labelled Markov Processes. In *Proceedings, 12th Annual IEEE Symposium on Logic in Computer Science, Warsaw, Poland, June 29 - July 2, 1997*, pages 149–158. IEEE Computer Society, 1997. doi:10.1109/LICS.1997.614943.
- 9 Filippo Bonchi, Barbara König, and Daniela Petrisan. Up-To Techniques for Behavioural Metrics via Fibrations. In Sven Schewe and Lijun Zhang, editors, *29th International Conference on Concurrency Theory, CONCUR 2018, September 4–7, 2018, Beijing, China*, volume 118 of *LIPICs*, pages 17:1–17:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPICs.CONCUR.2018.17.
- 10 Janusz A. Brzozowski. Derivatives of Regular Expressions. *J. ACM*, 11(4):481–494, 1964. doi:10.1145/321239.321249.
- 11 Stanley Burris and H P Sankappanavar. *A Course in Universal Algebra*. Lecture Notes in Statistics. Springer, New York, NY, November 1981.
- 12 Pedro R. D’Argenio, Daniel Gebler, and Matias David Lee. Axiomatizing Bisimulation Equivalences and Metrics from Probabilistic SOS rules. In Anca Muscholl, editor, *Foundations of Software Science and Computation Structures - 17th International Conference, FOSSACS 2014, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2014, Grenoble, France, April 5–13, 2014, Proceedings*, volume 8412 of *Lecture Notes in Computer Science*, pages 289–303. Springer, 2014. doi:10.1007/978-3-642-54830-7\_19.
- 13 Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled markov processes. *Theor. Comput. Sci.*, 318(3):323–354, 2004. doi:10.1016/J.TCS.2003.09.013.
- 14 Tiago Ferreira, Gerco van Heerdt, and Alexandra Silva. Tree-Based Adaptive Model Learning. In Nils Jansen, Mariëlle Stoelinga, and Petra van den Bos, editors, *A Journey from Process Algebra via Timed Automata to Model Learning - Essays Dedicated to Frits Vaandrager on the Occasion of His 60th Birthday*, volume 13560 of *Lecture Notes in Computer Science*, pages 164–179. Springer, 2022. doi:10.1007/978-3-031-15629-8\_10.
- 15 Jonas Forster, Lutz Schröder, and Paul Wild. Quantitative Graded Semantics and Spectra of Behavioural Metrics. *CoRR*, abs/2306.01487, 2023. doi:10.48550/arXiv.2306.01487.
- 16 Alessandro Giacalone, Chi-Chang Jou, and Scott A. Smolka. Algebraic Reasoning for Probabilistic Concurrent Systems. In Manfred Broy and Cliff B. Jones, editors, *Programming concepts and methods: Proceedings of the IFIP Working Group 2.2, 2.3 Working Conference on Programming Concepts and Methods, Sea of Galilee, Israel, 2–5 April, 1990*, pages 443–458. North-Holland, 1990.

- 17 Sergey Goncharov, Dirk Hofmann, Pedro Nora, Lutz Schröder, and Paul Wild. Kantorovich Functors and Characteristic Logics for Behavioural Distances. In Orna Kupferman and Pawel Sobocinski, editors, *Foundations of Software Science and Computation Structures - 26th International Conference, FoSSaCS 2023, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2023, Paris, France, April 22-27, 2023, Proceedings*, volume 13992 of *Lecture Notes in Computer Science*, pages 46–67. Springer, 2023. doi:10.1007/978-3-031-30829-1\_3.
- 18 John E. Hopcroft and Richard M. Karp. A Linear Algorithm for Testing Equivalence of Finite Automata, 1971. URL: <https://api.semanticscholar.org/CorpusID:120207847>.
- 19 Dexter Kozen. A Completeness Theorem for Kleene Algebras and the Algebra of Regular Events. *Inf. Comput.*, 110(2):366–390, 1994. doi:10.1006/INCO.1994.1037.
- 20 Marta Z. Kwiatkowska. A Metric for Traces. *Inf. Process. Lett.*, 35(3):129–135, 1990. doi:10.1016/0020-0190(90)90061-2.
- 21 Kim G. Larsen, Uli Fahrenberg, and Claus R. Thrane. Metrics for weighted transition systems: Axiomatization and complexity. *Theor. Comput. Sci.*, 412(28):3358–3369, 2011. doi:10.1016/J.TCS.2011.04.003.
- 22 Radu Mardare, Prakash Panangaden, and Gordon D. Plotkin. Quantitative Algebraic Reasoning. In Martin Grohe, Eric Koskinen, and Natarajan Shankar, editors, *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016*, pages 700–709. ACM, 2016. doi:10.1145/2933575.2934518.
- 23 Radu Mardare, Prakash Panangaden, and Gordon D. Plotkin. Fixed-points for Quantitative Equational Logics. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021, Rome, Italy, June 29 - July 2, 2021*, pages 1–13. IEEE, 2021. doi:10.1109/LICS52264.2021.9470662.
- 24 Robin Milner. A Complete Inference System for a Class of Regular Behaviours. *J. Comput. Syst. Sci.*, 28(3):439–466, 1984. doi:10.1016/0022-0000(84)90023-0.
- 25 David Michael Ritchie Park. Concurrency and Automata on Infinite Sequences. In *Theoretical Computer Science*, 1981. URL: <https://api.semanticscholar.org/CorpusID:206841958>.
- 26 Walter Rudin. *Functional Analysis*. International Series in Pure & Applied Mathematics. McGraw Hill Higher Education, Maidenhead, England, 2 edition, October 1990.
- 27 Jan J. M. M. Rutten. Universal coalgebra: a theory of systems. *Theor. Comput. Sci.*, 249(1):3–80, 2000. doi:10.1016/S0304-3975(00)00056-6.
- 28 Wojciech Różowski. A Complete Quantitative Axiomatisation of Behavioural Distance of Regular Expressions, 2024. arXiv:2404.13352.
- 29 Wojciech Różowski, 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.
- 30 Arto Salomaa. Two Complete Axiom Systems for the Algebra of Regular Events. *J. ACM*, 13(1):158–169, 1966. doi:10.1145/321312.321326.
- 31 Davide Sangiorgi. *Coinduction and the duality with induction*, pages 28–88. Cambridge University Press, 2011.
- 32 Todd Schmid, Wojciech Różowski, Alexandra Silva, and Jurriaan Rot. Processes Parametrised by an Algebraic Theory. In Mikolaj Bojanczyk, Emanuela Merelli, and David P. Woodruff, editors, *49th International Colloquium on Automata, Languages, and Programming, ICALP 2022, July 4-8, 2022, Paris, France*, volume 229 of *LIPICs*, pages 132:1–132:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPICs.ICALP.2022.132.
- 33 A.M Silva. *Kleene coalgebra*. PhD thesis, Radboud Universiteit Nijmegen, 2010.

- 34 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.
- 35 Eugene W. Stark and Scott A. Smolka. A complete axiom system for finite-state probabilistic processes. In Gordon D. Plotkin, Colin Stirling, and Mads Tofte, editors, *Proof, Language, and Interaction, Essays in Honour of Robin Milner*, pages 571–596. The MIT Press, 2000.
- 36 Alfred Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5(2):285–309, 1955.
- 37 Franck van Breugel. On behavioural pseudometrics and closure ordinals. *Inf. Process. Lett.*, 112(19):715–718, 2012. doi:10.1016/J.IPL.2012.06.019.
- 38 Franck van Breugel. Probabilistic bisimilarity distances. *ACM SIGLOG News*, 4(4):33–51, 2017. doi:10.1145/3157831.3157837.
- 39 Franck van Breugel and James Worrell. Towards Quantitative Verification of Probabilistic Transition Systems. In Fernando Orejas, Paul G. Spirakis, and Jan van Leeuwen, editors, *Automata, Languages and Programming, 28th International Colloquium, ICALP 2001, Crete, Greece, July 8-12, 2001, Proceedings*, volume 2076 of *Lecture Notes in Computer Science*, pages 421–432. Springer, 2001. doi:10.1007/3-540-48224-5\_35.
- 40 Franck van Breugel and James Worrell. Approximating and computing behavioural distances in probabilistic transition systems. *Theor. Comput. Sci.*, 360(1-3):373–385, 2006. doi:10.1016/J.TCS.2006.05.021.
- 41 Rob J. van Glabbeek. The Linear Time-Branching Time Spectrum (Extended Abstract). In Jos C. M. Baeten and Jan Willem Klop, editors, *CONCUR '90, Theories of Concurrency: Unification and Extension, Amsterdam, The Netherlands, August 27-30, 1990, Proceedings*, volume 458 of *Lecture Notes in Computer Science*, pages 278–297. Springer, 1990. doi:10.1007/BFB0039066.
- 42 Cédric Villani. *Optimal Transport*. Springer Berlin Heidelberg, 2009. doi:10.1007/978-3-540-71050-9.
- 43 Jana Wagemaker, Marcello M. Bonsangue, Tobias Kappé, Jurriaan Rot, and Alexandra Silva. Completeness and Incompleteness of Synchronous Kleene Algebra. In Graham Hutton, editor, *Mathematics of Program Construction - 13th International Conference, MPC 2019, Porto, Portugal, October 7-9, 2019, Proceedings*, volume 11825 of *Lecture Notes in Computer Science*, pages 385–413. Springer, 2019. doi:10.1007/978-3-030-33636-3\_14.
- 44 Paul Wild and Lutz Schröder. Characteristic Logics for Behavioural Hemimetrics via Fuzzy Lax Extensions. *Log. Methods Comput. Sci.*, 18(2), 2022. doi:10.46298/LMCS-18(2:19)2022.