FRACTALS FROM REGULAR BEHAVIOURS

TODD SCHMID ●[a], VICTORIA NOQUEZ ●[b], AND LAWRENCE S. MOSS ●[c]

[a] Bucknell University, USA
   e-mail address: t.schmid@bucknell.edu

[b] St. Mary's College of California, USA
   e-mail address: vln1@stmarys-ca.edu

[c] Indiana University, USA
   e-mail address: lmoss@iu.edu

ABSTRACT. We forge connections between the theory of fractal sets obtained as attractors of iterated function systems and process calculi. To this end, we reinterpret Milner's expressions for processes as contraction operators on a complete metric space. When the space is, for example, the plane, the denotations of fixed point terms correspond to familiar fractal sets. We give a sound and complete axiomatization of fractal equivalence, the congruence on terms consisting of pairs that construct identical self-similar sets in all interpretations. We further make connections to labelled Markov chains and to invariant measures. In all of this work, we use important results from process calculi. For example, we use Rabinovich's completeness theorem for trace equivalence in our own completeness theorem. In addition to our results, we also raise many questions related to both fractals and process calculi.

1. INTRODUCTION

Hutchinson noticed in [Hut81] that many familiar examples of fractals can be captured as the set-wise fixed-point of a finite family of contraction (i.e., distance shrinking) operators on a metric space. He called these spaces *self-similar*, since the intuition behind the contraction operators is that they are witnesses for the appearance of the fractal in a proper (smaller) subset of itself. For example, the famous Sierpiński gasket is the unique non-empty compact subset of the plane left fixed by the union of the three operators $\sigma_a, \sigma_b, \sigma_c : \mathbb{R}^2 \to \mathbb{R}^2$ in Figure 1.

The self-similarity of Hutchinson's fractals hints at an algorithm for constructing them: Each point in a self-similar set is the limit of a sequence of points obtained by applying the contraction operators one after the other to an initial point. In the Sierpiński gasket, the point $(1/4, \sqrt{3}/4)$ is the limit of the sequence

$$p, \ \sigma_b(p), \ \sigma_b\sigma_a(p), \ \sigma_b\sigma_a\sigma_a(p), \ \sigma_b\sigma_a\sigma_a\sigma_a(p), \ \ldots \tag{1.1}$$

where the initial point p is an arbitrary element of $\mathbb{R}^2$ (note that σ_b is applied last). Hutchinson showed in [Hut81] that the self-similar set corresponding to a given family of contraction operators is precisely the collection of points obtained in the manner just

described. The limit of the sequence in (1.1) does not depend on the initial point p because $\sigma_a, \sigma_b, \sigma_c$ are contractions. Much like digit expansions to real numbers, every stream of a's, b's, and c's corresponds to a unique point in the Sierpiński gasket, as we have seen in (1.1). The point $(1/4, \sqrt{3}/4)$, for example, corresponds to the stream $(b, a, a, a, \dots)$ ending in an infinite sequence of a's. Conversely, every point in the Sierpiński gasket comes from (in general more than one) corresponding stream.

From a computer science perspective, the languages of streams considered by Hutchinson are the *traces* observed by one-state labelled transition systems, like the one in Figure 1. We investigate whether one could achieve a similar effect with languages of streams obtained from labelled transition systems having more than one state. Observe, for example, Figure 2. These twisted versions of the Sierpiński gasket are constructed from the streams emitted by the two-state labelled transition system in Figure 2, starting from the states x and y respectively.

Each point in a twisted Sierpiński gasket corresponds to a stream of a's, b's, and c's, but not every stream corresponds to a point in the set: for example, the limit corresponding to the stream $(a, c, a, b, a, a, a, \dots)$ is $p = (19/32, 11\sqrt{3}/32)$. This stream is not emitted by either of the states x or y, and consequently the point p does not appear in either of the twisted Sierpiński gaskets generated by x or y.

In analogy with the theory of regular languages, we call the fractals generated by finite labelled transition systems *regular subfractals*, and give a logic for deciding if two labelled transition systems represent the same recipe under all interpretations of the labels, allowing both the underlying space and the chosen contractions to vary. By identifying points in the fractal set generated by a labelled transition system with traces observed by the labelled transition system, it is reasonable to suspect that two labelled transition systems represent equivalent fractal recipes—i.e., they represent the same fractal under every interpretation—if and only if they are trace equivalent. This is the content of Theorem 4.3, which allows us to connect the theory of fractal sets to mainstream topics in computer science.

Labelled transition systems are a staple of theoretical computer science, especially in the area of process algebra [Bae05], where a vast array of different notions of equivalence and axiomatization problems have been studied. We specifically use a syntax introduced by Milner in [Mil84] to express labelled transition systems as terms in an expression language with recursion. This leads us to a fragment of Milner's calculus consisting of just the terms that constitute recipes for fractal constructions. Using a logic of Rabinovich [Rab93] for

$$a, b, c \circlearrowleft \boxed{x}$$

$$\sigma_a \begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} \frac{1}{2}r + \frac{1}{4} \\ \frac{1}{2}s + \frac{\sqrt{3}}{4} \end{bmatrix} \qquad \sigma_b \begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} \frac{1}{2}r \\ \frac{1}{2}s \end{bmatrix}$$

$$\sigma_c \begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} \frac{1}{2}r + \frac{1}{2} \\ \frac{1}{2}s \end{bmatrix}$$
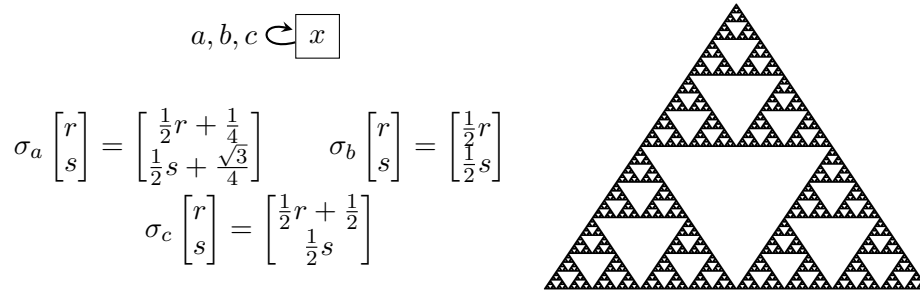
Figure 1: The Sierpiński gasket is the unique non-empty compact subset $\mathbf{S}$ of $\mathbb{R}^2$ such that $\mathbf{S} = \sigma_a(\mathbf{S}) \cup \sigma_b(\mathbf{S}) \cup \sigma_c(\mathbf{S})$. Each of its points corresponds to a stream emitted by the state x.
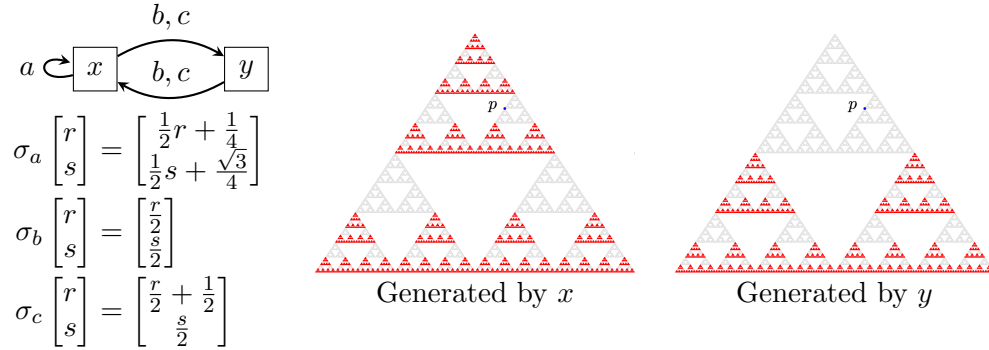
$$\sigma_a \begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} \frac{1}{2}r + \frac{1}{4} \\ \frac{1}{2}s + \frac{\sqrt{3}}{4} \end{bmatrix}$$

$$\sigma_b \begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} \frac{r}{2} \\ \frac{s}{2} \end{bmatrix}$$

$$\sigma_c \begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} \frac{r}{2} + \frac{1}{2} \\ \frac{s}{2} \end{bmatrix}$$

Generated by x Generated by y

Figure 2: Twisted Sierpińksi gaskets, depicted in red, and the point $p = (\frac{19}{32}, \frac{11\sqrt{3}}{32})$. The fractal set generated by x is the unique compact set fixed by taking the union of the image of σ_a with the images of the four compositions $\sigma_b\sigma_b$, $\sigma_b\sigma_c$, $\sigma_c\sigma_b$, and $\sigma_c\sigma_c$. Intuitively, this is the set obtained from the Sierpinski gasket by recursively removing the upper triangle in the bottom two thirds. The set generated by y is the union of σ_b and σ_c applied to the set generated by x.

deciding trace equivalence in Milner's calculus, we obtain a complete axiomatization of fractal recipe equivalence. As a consequence of this completeness theorem, we are also able to show that Rabinovich's logic is complete for the specific interpretation of process terms as fractal subsets of the Cantor set.

In his study of self-similar sets, Hutchinson also makes use of probability measures supported on self-similar sets, called *invariant measures*. Each invariant measure is specified by a probability distribution on the set of contractions generating its support. In the last technical section of the paper, we adapt the construction of invariant measures to a probabilistic version of labelled transition systems called *labelled Markov chains*. These allow us to give a measure-theoretic semantics to terms in a probabilistic version of Milner's specification language, the calculus introduced by Stark and Smolka [SS00]. Specifically, we interpret each probabilistic process term as a probability measure supported by a regular subfractal. Our measure-theoretic semantics of probabilistic process terms can be seen as a generalization of the trace measure semantics of Kerstan and König [KK13]. We offer a sound axiomatization of equivalence under this semantics and, calling on a result proved in a separate paper [CMN+25] using other techniques, also prove completeness. Finally, we show that our proposed axiomatization of fractal measure equivalence is also complete for the specific interpretation of probabilistic process terms as fractal measures on the Cantor set.

In sum, the contributions of this paper are as follows.

- In Section 3, we give a fractal recipe semantics to process terms using a generalization of iterated function systems.
- In Section 4, we show that two process terms agree on all fractal interpretations if and only if they are trace equivalent. This implies that fractal recipe equivalence is decidable for process terms, and it allows us to derive a complete axiomatization of fractal recipe equivalence from Rabinovich's axiomatization [Rab93] of trace equivalence of process terms.

- We adapt the fractal semantics of process terms to the probabilistic setting in Section 5. We furthermore show that probabilistic fractal recipe equivalence coincides with Kerstan and König's notion of trace measure equivalence [KK13].
- Finally, we propose an axiomatization of probabilistic fractal recipe equivalence in Sections 6 and 7. We prove that our axioms are sound and complete with respect to our semantics.

2. LABELLED TRANSITION SYSTEMS AND TRACE SEMANTICS

Labelled transition systems are a widely used model of nondeterminism. We are given a fixed finite set A of *action labels*. A *labelled transition system* (LTS) is a pair (X, α) consisting of a set X of *states* (also called the *state space*) and a *transition function* $\alpha \colon X \to \mathcal{P}(A \times X)$. We generally write $x \xrightarrow{a}_\alpha y$ if $(a, y) \in \alpha(x)$, or simply $x \xrightarrow{a} y$ if α is clear from context, and say that x *emits a and transitions to y*.

Given a state x of an LTS (X, α), we write $\langle x \rangle_\alpha$ for the LTS obtained by restricting the relations $\xrightarrow{a}$ to the set of states *reachable* from x, meaning there exists a *path* of the form $x \xrightarrow{a_1} x_1 \to \cdots \to x_{n-1} \xrightarrow{a_n} x_n$. We refer to $\langle x \rangle_\alpha$ as either the LTS *generated by x*, or the *process starting at x*. An LTS (X, α) is *locally finite* if $\langle x \rangle_\alpha$ is finite for every state x. The reader should note that there exist locally finite LTSs with infinite state spaces, and that one such LTS will be important in this paper; see Definition 2.5.

Traces. In the context of the current work, nondeterminism occurs when a process branches into multiple threads that execute in parallel. Under this interpretation, to an outside observer (without direct access to the implementation details of an LTS), two processes that emit the same set of sequences of action labels are indistinguishable.

Formally, let A^* be the set of finite words formed from the alphabet A. Given a state x of an LTS (X, α), the set $\mathrm{tr}_\alpha(x)$ of *traces emitted by x* is the set of words $a_1 \cdots a_n \in A^*$ such that there is a path of the form $x \xrightarrow{a_1} x_1 \xrightarrow{a_2} \cdots \xrightarrow{a_{n-1}} x_{n-1} \xrightarrow{a_n} x_n$ through (X, α). Given LTSs (X, α) and (Y, β), states $x \in X$ and $y \in Y$ are called *trace equivalent* if $\mathrm{tr}_\alpha(x) = \mathrm{tr}_\beta(y)$. If the transition structure is clear from context, we drop it from the notation and write simply $\mathrm{tr}(x)$ for $\mathrm{tr}_\alpha(x)$. It should be noted that every trace language $\mathrm{tr}_\alpha(x)$ is *prefix-closed*, which for a language L means that $w \in L$ whenever $wa \in L$ for some $a \in A$.

Trace equivalence is a well-documented notion of equivalence for processes, dating back at least to the work of Hoare and Milner in the late 1970s [Hoa78, Mil78, BPS01, Bae05]. We shall see it in our work on fractals as well.

Definition 2.1. A *stream* is an infinite sequence $(a_i)_{i \in \mathbb{N}} = (a_1, a_2, \dots)$ of letters from A. The set of streams is denoted A^ω. A state x in an LTS (X, α) *emits* a stream $(a_i)_{i \in \mathbb{N}}$ if there is an infinite path of the form $x \xrightarrow{a_1} x_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} x_n \xrightarrow{a_{n+1}} \cdots$ through (X, α). We write $\mathrm{str}_\alpha(x)$ for the set of streams emitted by x. Two states x, y are said to be *stream equivalent* if $\mathrm{str}_\alpha(x) = \mathrm{str}_\alpha(y)$.

Remark 2.2. We use both $(a_i)_{i \in \mathbb{N}}$ and $(a_1, a_2, \dots)$ as notations for streams. These two are intended to denote the same stream of letters. Note that indexing starts at 1 rather than 0.

In our construction of fractals from LTSs, points are represented only by (infinite) streams. We therefore focus primarily on LTSs with the property that for all states x, $\mathrm{tr}(x)$ is precisely the set of finite prefixes of streams emitted by x. We refer to an LTS (X, α)

satisfying this condition as *productive*. Productivity is equivalent to the absence of *deadlock* states, which are states with no outgoing transitions.

Lemma 2.3. *Let (X, α) be a finitely branching LTS. Then the following are equivalent:*

(1) (X, α) is productive, i.e., for any $x \in X$,

$$\mathrm{tr}_\alpha(x) = \{a_1 \ldots a_n \in A^* \mid \exists (b_i)_{i \in \mathbb{N}} \in \mathrm{str}_\alpha(x) \text{ such that } b_1 = a_1, \ldots, b_n = a_n\} \qquad (2.1)$$

$$\mathrm{str}_\alpha(x) = \{(a_i)_{i \in \mathbb{N}} \in A^\omega \mid \forall n \in \mathbb{N}, \ a_1 \cdots a_n \in \mathrm{tr}_\alpha(x)\} \qquad (2.2)$$

(2) For any $x \in X$, $\alpha(x) \neq \emptyset$.

Consequently, if (X, α) is productive and finitely branching, and $x, y \in X$, then $\mathrm{str}_\alpha(x) = \mathrm{str}_\alpha(y)$ if and only if $\mathrm{tr}_\alpha(x) = \mathrm{tr}_\alpha(y)$. In other words, for productive and finitely branching LTSs, trace equivalence and stream equivalence coincide.

Proof. To see that 1 implies 2, observe that trivially, $\epsilon \in \mathrm{tr}_\alpha(x)$ for any $x \in X$ (ϵ is the empty word). This implies there is a stream $(a_i)_{i \in \mathbb{N}} \in \mathrm{str}_\alpha(x)$ such that ϵ is an initial segment of $(a_i)_{i \in \mathbb{N}}$, or in other words, $\mathrm{str}_\alpha(x) \neq \emptyset$. Hence, $\alpha(x) \neq \emptyset$, as x emits a stream.

To see that 2 implies 1, assume $\alpha(x) \neq \emptyset$ for all $x \in X$. Then for any $x \in X$, a simple inductive argument shows that there are paths of arbitrary length starting from x. It follows from König's lemma [Kle02] (every infinite but finitely branching tree has an infinite branch) that there is therefore an infinite path starting at x. This allows us to argue as follows: To see (2.1), it suffices to show that $\mathrm{tr}_\alpha(x)$ is contained in the right hand side, as the reverse containment is clear. So, let $x \in X$ and suppose that $a_1 \ldots a_n \in \mathrm{tr}_\alpha(x)$. Then there is a path $x \xrightarrow{a_1} \cdots \xrightarrow{a_n} x_n$, and by the observation above, there is an infinite path $x_n \xrightarrow{a_{n+1}} x_{n+1} \xrightarrow{a_{n+1}} \cdots$. Hence, $(a_1, \ldots, a_n, a_{n+1}, a_{n+2}, \ldots) \in \mathrm{str}_\alpha(x)$, as desired. This establishes (2.1).

To see (2.2), it suffices to see that $\mathrm{str}_\alpha(x)$ contains the right hand side, as the forward containment is clear. So, let $(a_i)_{i \in \mathbb{N}} \in A^\omega$ such that for any $n \in \mathbb{N}$, $a_1 \cdots a_n \in \mathrm{tr}_\alpha(x)$. We are going to use König's lemma again: Let $x \in X$ and let $\mathfrak{P}$ be the set of all paths of the form $x \xrightarrow{a_1} \cdots \xrightarrow{a_n} x_n$. We define a tree structure on $\mathfrak{P}$ as follows: Given $P_1, P_2 \in \mathfrak{P}$, write $P_1 \to P_2$ if $P_1 = (x \xrightarrow{a_1} \cdots \xrightarrow{a_n} x_n)$ and $P_2 = (x \xrightarrow{a_1} \cdots \xrightarrow{a_n} x_n \xrightarrow{a_{n+1}} x_{n+1})$ for some x_{n+1}. Then $(\mathfrak{P}, \to)$ is an infinite but finitely branching tree, since every initial segment of $(a_i)_{i \in \mathbb{N}}$ is in $\mathrm{tr}_\alpha(x)$ and (X, α) is finitely branching. By König's lemma, there is an infinite path $P_1 \to P_2 \to \cdots$ in $(\mathfrak{P}, \to)$. By definition, this means there is an infinite path $x \xrightarrow{a_1} x_1 \xrightarrow{a_2} \cdots$ in (X, α). Hence, $(a_i)_{i \in \mathbb{N}} \in \mathrm{str}_\alpha(x)$. This establishes (2.2).

Let us now establish the stated consequence. In one direction, let $\mathrm{tr}_\alpha(x) = \mathrm{tr}_\alpha(y)$. Then

$$\begin{aligned}
\mathrm{str}_\alpha(x) &= \{(a_i)_{i \in \mathbb{N}} \in A^\omega \mid \forall n \in \mathbb{N}, \ a_1 \cdots a_n \in \mathrm{tr}_\alpha(x)\} \\
&= \{(a_i)_{i \in \mathbb{N}} \in A^\omega \mid \forall n \in \mathbb{N}, \ a_1 \cdots a_n \in \mathrm{tr}_\alpha(y)\} \\
&= \mathrm{str}_\alpha(y)
\end{aligned}$$

In the other, if $\mathrm{str}_\alpha(x) = \mathrm{str}_\alpha(y)$, then

$$\begin{aligned}
\mathrm{tr}_\alpha(x) &= \{a_1 \ldots a_n \in A^* \mid \exists (b_i)_{i \in \mathbb{N}} \in \mathrm{str}_\alpha(x) \text{ such that } b_1 = a_1, \ldots, b_n = a_n\} \\
&= \{a_1 \ldots a_n \in A^* \mid \exists (b_i)_{i \in \mathbb{N}} \in \mathrm{str}_\alpha(y) \text{ such that } b_1 = a_1, \ldots, b_n = a_n\} \\
&= \mathrm{tr}_\alpha(y) \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \square
\end{aligned}$$

Remark 2.4. The finite branching condition is necessary in Lemma 2.3. For example, consider the LTSs in Figure 3. Between x_0 and y_0, only y_0 emits the stream $(a, b, a, b, a, b, \ldots)$ of alternating letters a and b, despite both states emitting the trace language $\{(ab)^n a^m \mid$

$m, n \in \mathbb{N}\}$. This is possible because the LTS on the left is not finitely branching: Since $x_0 \xrightarrow{a} x_{i,1}$ for every $i \in \mathbb{N}$, x_0 has infinitely many outgoing paths.
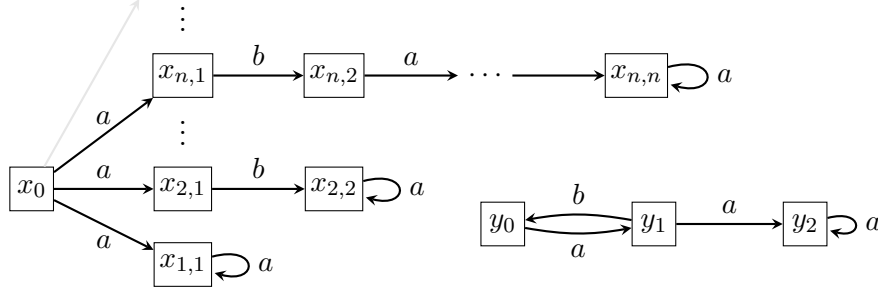

Figure 3: (A hedgehog model) In the LTS on the left in the above picture, the paths starting from x_0 are all of the form $x_0 \xrightarrow{a} x_{n,1} \xrightarrow{b} x_{n,2} \xrightarrow{a} x_{n,3} \xrightarrow{b} \cdots \rightarrow x_{n,n} \xrightarrow{a} \cdots$ (beginning with alternating as and bs). Thus, the state x_0 emits every finite initial segment of $(a, b, a, b, a, b, \dots)$, but does not emit this stream. Therefore, $\mathrm{tr}(x_0) = \mathrm{tr}(y_0)$ and yet $\mathrm{str}(x_0) \neq \mathrm{str}(y_0)$.

Specification. We use the following language for specifying processes: Starting with a fixed countably infinite set $\{v_1, v_2, \dots\}$ of *variables*, the set of *μ-expressions* is generated by the grammar

$$v \mid ae \mid e_1 + e_2 \mid \mu v\, e$$

where v is v_i for some $i \in \mathbb{N}$, $a \in A$, and e, e_1, e_2 are μ-expressions.

Intuitively, the process ae emits a and then turns into e, and $e_1 + e_2$ is the process that nondeterministically branches into e_1 and e_2. The process $\mu v\, e$ is like e, but with instances of v that appear free in e acting like `goto` expressions that return the process to $\mu v\, e$.

Definition 2.5. A *process term* is a μ-expression e in which every occurrence of a variable v appears both within the scope of a *recursive call* $\mu v\, (-)$ (e is *closed*) and within the scope of an *action prefix operator* $a(-)$ (e is *guarded*). For short, we call these *terms*. The set of terms is written Term. It is the set underlying the LTS (Term, γ) defined in Figure 4.

$$ae \xrightarrow{a} e \qquad \frac{e_1 \xrightarrow{a} f}{e_1 + e_2 \xrightarrow{a} f} \qquad \frac{e_2 \xrightarrow{a} f}{e_1 + e_2 \xrightarrow{a} f} \qquad \frac{e[\mu v\, e/v] \xrightarrow{a} f}{\mu v\, e \xrightarrow{a} f}$$

Figure 4: The relation $\xrightarrow{a} \subseteq \mathsf{Term} \times \mathsf{Term}$ defining (Term, γ).

Call a variable v appearing in an expression e *bound* if it appears withing the scope of a recursive call $\mu v\, (-)$ and *free* if it does not. In Figure 4, we use the notation $e[g/v]$ to denote the expression obtained by replacing each free occurrence of v in e with the expression g. Note that the expression $e[g/v]$ is only well-defined if no free variable in g appears bound in e. Given $e \in \mathsf{Term}$, the *process specified by* e is the LTS $\langle e \rangle_\gamma$.

Example 2.6. Let $e = av$, and let $f = \mu v\ e = \mu v\ av$. Then $e[f/v] = af \xrightarrow{a} f$. By the last rule in Figure 4, $\mu v\ e \xrightarrow{a} f = \mu v\ e$. By looking at the other rules, we see that none apply. Hence, if $\mu v\ e \xrightarrow{b} g$, then $b = a$ and $g = \mu v\ e$. The upshot is that $\langle f \rangle_\gamma$, the process specified by f, is the one-point process which can perform a but no other action: its state set is $\{f\}$, and $\gamma(f) = \{(a, f)\}$.

$$a \,\circlearrowright\, \boxed{\mu v\ av} \qquad \gamma(\mu v\ av) = \gamma(a(\mu v\ av)) = \{(a, \mu v\ av)\}$$

Remark 2.7. The set of process terms, as we have named them, is the fragment of Milner's fixed-point calculus from [Mil84] consisting of only the μ-expressions that specify productive LTSs.

Labelled transition systems specified by process terms are finite and productive, and conversely, every finite productive process is trace-equivalent to some process term.

Lemma 2.8. *The LTS* (Term, γ) *is productive and locally finite, i.e., for any* $e \in \mathsf{Term}$, *the set of terms reachable from* e *in* (Term, γ) *is finite. Conversely, if* x *is a state in a productive and locally finite LTS* (X, α), *then there is a process term* e *such that* $\mathrm{tr}(e) = \mathrm{tr}_\alpha(x)$.

Proof. The proof that (Term, γ) is productive is a routine induction on $e \in \mathsf{Term}$ that shows that $\gamma(e) \neq \emptyset$. To see that (Term, γ) is locally finite, we start by reproducing Milner's [Mil84, Proposition 5.1] below, to which we have made only small changes to reflect the minor differences between our syntax and Milner's.

Consider the LTS consisting of arbitrary μ-expressions (as opposed to just process terms). Let us write $(E, \bar{\gamma})$ for the set of μ-expressions E equipped with the transition structure $\bar{\gamma}$ defined as follows: $\bar{\gamma}(v_i) = \emptyset$ for all $i \in \mathbb{N}$, and all other transition relations are those derivable from the rules in Figure 4. Write $\to^+$ and $\to^*$ for the transitive and reflexive-transitive closures of the transition relation $\to$ respectively. For any $e \in E$, define the set of all μ-expressions reachable from e by $\langle e \rangle = \{f \in E \mid e \to^* f\}$. We proceed with a proof by induction on μ-expressions that $\langle e \rangle$ is finite for every $e \in E$, i.e., $(E, \bar{\gamma})$ is locally finite.[1]

In the base case, we have $e = v_i$ for some $i \in \mathbb{N}$. Here, $|\langle v_i \rangle| = |\{v_i\}| = 1$. For the inductive step, let e, e_1, e_2 be guarded μ-expressions and assume that $|\langle e_i \rangle|$ for $i \in \{1, 2\}$ and $|\langle e \rangle|$ are all finite.

- By definition of $\bar{\gamma}(ae)$, $|\langle ae \rangle| = |\{ae\} \cup \langle e \rangle| \leq 1 + |\langle e \rangle|$. The induction hypothesis therefore tells us that $\langle ae \rangle$ is finite.
- The set $\langle e_1 + e_2 \rangle$ precisely consists of $e_1 + e_2$ and all of the μ-expressions f such that either $e_1 \to^+ f$ or $e_2 \to^+ f$. It follows that

$$|\langle e_1 + e_2 \rangle| = |\{e_1 + e_2\} \cup \langle e_1 \rangle \cup \langle e_2 \rangle| \leq 1 + |\langle e_1 \rangle| + |\langle e_2 \rangle|$$

  By the induction hypothesis, $\langle e_1 + e_2 \rangle$ is finite.
- For the final induction step, observe that $\mu v\ e \to^+ f$ if and only if $e[\mu v\ e/v] \to^+ f$. It follows that $\langle \mu v\ e \rangle \setminus \{\mu v\ e, e[\mu v\ e/v]\} = \langle e[\mu v\ e/v] \rangle \setminus \{\mu v\ e, e[\mu v\ e/v]\}$. Clearly then, $\langle \mu v\ e \rangle$ is finite if and only if $\langle e[\mu v\ e/v] \rangle$ is finite. Now, every expression f for which $e[\mu v\ e/v] \to^* f$ is necessarily either of the form $g[\mu v\ e/v]$ with $e \to^* g$ or is $\mu v\ e$ itself. This tells us that

$$|\langle e[\mu v\ e/v] \rangle| \leq |\{\mu v\ e\} \cup \{g[\mu v\ e/v] \mid e \to^* g\}| = 1 + |\{g \mid e \to^* g\}| = 1 + |\langle e \rangle|$$

---

[1]And in fact, the size of $\langle e \rangle$ is linear with respect to the length of the μ-expression e.

$$
\begin{array}{ll}
\text{(I)} & e + e \equiv e \\
\text{(C)} & e_2 + e_1 \equiv e_1 + e_2 \\
\text{(A)} & e_1 + (e_2 + e_3) \equiv (e_1 + e_2) + e_3 \\
\text{(D)} & a(e_1 + e_2) \equiv ae_1 + ae_2 \\
\text{(F)} & \mu v\ e \equiv e[\mu v\ e/v]
\end{array}
$$

$$
\text{(α)} \quad \mu w\ e \equiv \mu v\ e[v/w]
$$

$$
\text{(Cong)} \quad \frac{e_1 \equiv f_1\ \cdots\ e_n \equiv f_n}{g[\vec{e}/\vec{v}] \equiv g[\vec{f}/\vec{v}]}
$$

$$
\text{(RSP)} \quad \frac{g \equiv e[g/v]}{g \equiv \mu v\ e}
$$

Figure 5: The names for the axioms on the left stand for *idempotence, commutativity, associativity, distributivity*, and *fixed point* respectively. On the right, we have *α-equivalence, congruence*, and the *recursive specification principle*. The axioms and rules of the provable equivalence relation are those listed here, in addition to those of equational logic (not shown). Here, $e, e_i, f, f_i, g \in \mathsf{Term}$ for all i. In (Cong), g has precisely the free variables $v_1, \ldots, v_n$, and no variable that appears free in f_i is bound in g for any i. In (α), v does not appear free in e.

The induction hypothesis now tells us that $\langle e[\mu v\ e/v]\rangle$ is finite, and therefore $\langle \mu v\ e\rangle$ is also finite.

This concludes the proof that $(E, \bar{\gamma})$ is locally finite. To see that (Term, γ) is locally finite, observe from the definition of $\bar{\gamma}$ that for each process term $e \in \mathsf{Term}$ we have $\bar{\gamma}(e) = \gamma(e) \subseteq A \times \mathsf{Term}$. It follows that $\langle e \rangle \subseteq \mathsf{Term}$ for each $e \in \mathsf{Term}$. Since $(E, \bar{\gamma})$ is locally finite, $\langle e \rangle$ is finite. This shows that (Term, γ) is locally finite.

The converse statement, that if x is a state in a finite LTS (X, α), then $\mathrm{tr}_\alpha(x) = \mathrm{tr}(e)$ for some $e \in \mathsf{PTerm}$, is a direct consequence of [Mil84, Theorem 5.7 and Corollary 5.8].[2] $\square$

Axiomatization of trace equivalence. Given an interpretation of process terms as states in an LTS, and given the notion of trace equivalence, one might ask if there is an algebraic or proof-theoretic account of trace equivalence of process terms. Rabinovich showed in [Rab93] that a complete inference system for trace equivalence can be obtained by adapting earlier work of Milner [Mil84]. The axioms of the complete inference system include equations like $e_1 + e_2 = e_2 + e_1$ and $a(e_1 + e_2) = ae_1 + ae_2$, which are intuitively true for trace equivalence.

Definition 2.9. Given $e_1, e_2 \in \mathsf{Term}$, we say that e_1 and e_2 are *provably equivalent* if $e_1 \equiv e_2$ can be derived from the axioms in Figure 5, and call $\equiv$ *provable equivalence*.

Theorem 2.10 Rabinovich [Rab93]. *Let $e_1, e_2 \in \mathsf{Term}$. Then $e_1 \equiv e_2$ iff $\mathrm{tr}(e_1) = \mathrm{tr}(e_2)$.*

Example 2.11. Consider the processes specified by $e_1 = \mu w\ \mu v\ (a_1 a_2 v + a_1 a_3 w)$ and $e_2 = \mu v\ (a_1(a_2 v + a_3 v))$ below.
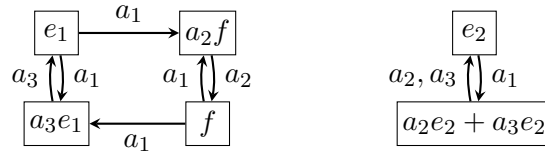
The term f is $\mu v\ (a_1 a_2 v + a_1 a_3 e_1)$. The traces emitted by both e_1 and e_2 are those that alternate between a_1 and either a_2 or a_3.

---

[2]In fact, the mentioned proof shows something stronger: that every state in a productive and locally finite LTS is *bisimilar* to a process term. It is well-known that bisimilarity implies trace equivalence [BPS01].

By Theorem 2.10, it should be possible to prove that $e_1 \equiv e_2$. Indeed, this can be done as follows: Applying (F) to $e_1 = \mu w \ \mu v \ (a_1 a_2 v + a_1 a_3 w)$, we start with

$$e_1 = \mu w \ \mu v \ (a_1 a_2 v + a_1 a_3 w) \stackrel{(\mathsf{F})}{\equiv} \mu v \ (a_1 a_2 v + a_1 a_3 e_1) = f$$

Applying (F) again, and then (Cong), we get

$$f \stackrel{(\mathsf{F})}{\equiv} a_1 a_2 f + a_1 a_3 e_1 \stackrel{(\mathsf{Cong})}{\equiv} a_1 a_2 e_1 + a_1 a_3 e_1$$

Applying (D) and using our substitution notation gives us

$$a_1 a_2 e_1 + a_1 a_3 e_1 \stackrel{(\mathsf{D})}{\equiv} a_1 (a_2 e_1 + a_3 e_1) = (a_1 (a_2 v + a_3 v))[e_1 / v]$$

Thus $e_1 \equiv (a_1 (a_2 v + a_3 v))[e_1 / v]$. We apply the rule (RSP) to see that

$$e_1 \equiv \mu v \ (a_1 (a_2 v + a_3 v)) = e_2$$

as desired.

Rabinovich's theorem tells us that, up to provable equivalence, our specification language consisting of process terms is really a specification language for languages of traces. In what follows, we are going to give an alternative semantics to process terms by using LTSs to generate fractal subsets of complete metric spaces. The main result of our paper is that the equivalence relations coming from these two semantics coincide: Two process terms are trace equivalent if and only if they generate the same fractals in all spaces. This is the content of Sections 3 and 4 below.

3. Fractals from Labelled Transition Systems

In the Sierpiński gasket $\mathbf{S}$ from Figure 1, every point of $\mathbf{S}$ corresponds to a stream of letters from the alphabet $\{a, b, c\}$, and every stream corresponds to a unique point. To obtain the point corresponding to a particular stream $(a_1, a_2, a_3, \dots)$ with each $a_i \in \{a, b, c\}$, start with any $p \in \mathbb{R}^2$ and compute the limit $\lim_{n \to \infty} \sigma_{a_1} \cdots \sigma_{a_n}(p)$. The point in the fractal corresponding to $(a_1, a_2, a_3, \dots)$ does not depend on p because $\sigma_a, \sigma_b, \sigma_c$ in Figure 1 are *contraction operators*.

Definition 3.1. Given a metric space M, a *contraction operator* on M is a function $h : M \to M$ such that for some $c \in [0, 1)$, $d(h(x), h(y)) \leq c \, d(x, y)$ for any $x, y \in M$. The number c is called a *contraction coefficient* of h. The set of contraction operators on M is written $\mathrm{Con}(M)$.[3]

For example, with the Sierpiński gasket (Figure 1) associated to the contractions σ_a, σ_b, and σ_c, $r = 1/2$ is a contraction coefficient for all three maps. Now, given $p, q \in \mathbb{R}^2$,

$$d(\sigma_{a_1} \cdots \sigma_{a_n}(p), \sigma_{a_1} \cdots \sigma_{a_n}(q)) \leq \frac{1}{2^n} \, d(p, q)$$

for all n, so it follows that $\lim_{n \to \infty} \sigma_{a_1} \cdots \sigma_{a_n}(p) = \lim_{n \to \infty} \sigma_{a_1} \cdots \sigma_{a_n}(q)$. For any finite set of contraction operators indexed by A and acting on a complete metric space M, every stream from A corresponds to a unique point in M.

---

[3]In general, with a metric space (M, d), we usually omit the metric and just write M. Also note that our notation $\mathrm{Con}(M)$ should not be read as suggesting that Con is a functor. It is not. Finally, we shall always assume that our metric spaces are non-empty.

Definition 3.2. A *contraction operator interpretation* is a function $\sigma \colon A \to \mathrm{Con}(M)$. We usually write $\sigma_a = \sigma(a)$. Given $\sigma \colon A \to \mathrm{Con}(M)$ and a stream $(a_i)_{i \in \mathbb{N}}$ from A, define

$$\sigma_\omega \colon A^\omega \to M \qquad \sigma_\omega(a_1, a_2, \dots) = \lim_{n \to \infty} \sigma_{a_1} \cdots \sigma_{a_n}(p) \tag{3.1}$$

where $p \in M$ is arbitrary. The *self-similar set* corresponding to a contraction operator interpretation σ is the set

$$\mathbf{S}_\sigma = \{\sigma_\omega(a_1, a_2, \dots) \mid (a_1, a_2, \dots) \text{ is a stream from } A\}. \tag{3.2}$$

Remark 3.3. Note that in (3.1), the contraction operators corresponding to the initial trace $(a_1, \dots, a_n)$ are applied in *reverse* order. That is, σ_{a_n} is applied before $\sigma_{a_{n-1}}$, $\sigma_{a_{n-2}}$ is applied before $\sigma_{a_{n-1}}$, and so on, and σ_{a_1} is applied last.

It should also be noted that contractions are *continuous*, which implies that for any stream $(a_1, a_2, \dots) \in A^\omega$, $\sigma \colon A \to \mathrm{Con}(M)$, and $p \in M$,

$$\sigma_\omega(a_1, a_2, \dots) = \lim_{n \to \infty} \sigma_{a_1} \sigma_{a_2} \cdots \sigma_{a_n}(p) = \sigma_{a_1}(\lim_{n \to \infty} \sigma_{a_2} \cdots \sigma_{a_n}(p)) = \sigma_{a_1} \circ \sigma_\omega(a_2, \dots)$$

If we write $a(-)$ for the prefixing map

$$a(a_i)_{i \in \mathbb{N}} = (a, a_1, \dots), \tag{3.3}$$

this implies the following statement: Let M be a complete metric space and $\sigma \colon A \to \mathrm{Con}(M)$. Then for any $a \in A$,

$$
\begin{array}{ccc}
A^\omega & \xrightarrow{\ a\ } & A^\omega \\
{\scriptstyle \sigma_\omega} \downarrow & & \downarrow {\scriptstyle \sigma_\omega} \\
M & \xrightarrow{\ \sigma_a\ } & M
\end{array}
\qquad \sigma_\omega \circ a(-) = \sigma_a \circ \sigma_\omega \tag{3.4}
$$

Regular Subfractals. Generalizing the fractals of Mandelbrot [Man77], Hutchinson introduced self-similar sets in [Hut81] and gave a comprehensive account of their theory. In op. cit., Hutchinson defines a self-similar set to be the invariant set of an *iterated function system* (or *IFS*). In our terminology, an iterated function system is equivalent to a contraction operator interpretation of a finite set A of actions, and the invariant set is the total set of points obtained from all streams from A. The fractals constructed from an LTS paired with a contraction operator interpretation generalize Hutchinson's self-similar sets to non-empty compact sets of points obtained from certain subsets of the streams, namely the subsets emitted by single states in the LTS.

Definition 3.4. Given a productive LTS (X, α) and a contraction operator interpretation $\sigma : A \to \mathrm{Con}(M)$, we define

$$[\![x]\!]_{\alpha, \sigma} = \sigma_\omega(\mathrm{str}_\alpha(x)) = \{\sigma_\omega((a_i)_{i \in \mathbb{N}}) \mid (a_i)_{i \in \mathbb{N}} \text{ is emitted by } x\} \tag{3.5}$$

and call this the set *generated by the state x*.

It is worth noting that $[\![x]\!]_{\alpha, \sigma}$ is always a subset of $\mathbf{S}_\sigma$, since $\mathrm{str}_\alpha(x) \subseteq A^\omega$ and therefore $[\![x]\!]_{\alpha, \sigma} = \sigma_\omega(\mathrm{str}_\alpha(x)) \subseteq \sigma_\omega(A^\omega) = \mathbf{S}_\sigma$.

Definition 3.5. Given a process term $e \in \mathsf{Term}$ and a contraction operator interpretation $\sigma : A \to \mathrm{Con}(M)$, the *regular subfractal semantics of e* corresponding to σ is $[\![e]\!]_\sigma = [\![e]\!]_{\gamma, \sigma}$, where γ is the structure map on Term given in Definition 2.5.

Example 3.6. The regular subfractal generated by x in Figure 2 is the regular subfractal semantics of $\mu v\ (av + b(bv + cv) + c(bv + cv))$ corresponding to the interpretation σ given in that figure. The regular subfractal semantics of e is a proper subset of the Sierpiński gasket, as it does not contain the point $p = (19/32, 11\sqrt{3}/32)$ corresponding to $(a, c, a, b, a, a, \dots)$.

Example 3.7. For another example involving the Sierpiński gasket, see Figure 6. Let $g = \mu u\ (bu + bv + cv)$, and note that v is free in g. The asymmetrical regular subfractal appearing in Figure 6 is the regular subfractal semantics of the term $e = \mu v\ (av + bg + cg)$ corresponding to the same contraction operator interpretation as in Figures 1 and 2.
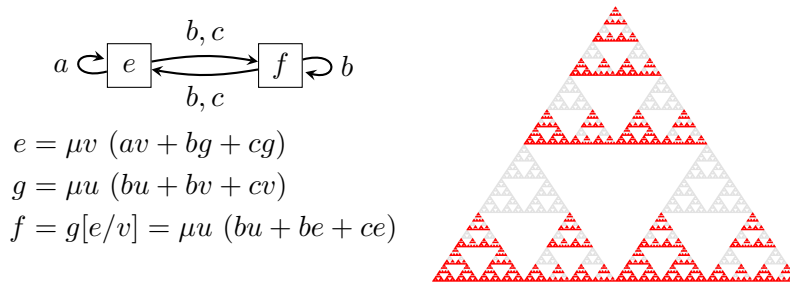

$$e = \mu v\ (av + bg + cg)$$
$$g = \mu u\ (bu + bv + cv)$$
$$f = g[e/v] = \mu u\ (bu + be + ce)$$

Figure 6: An asymmetrical regular subfractal of the Sierpiński gasket generated by the process term $e = \mu v\ (av + bg + cg)$, where $g = \mu u\ (bu + bv + cv)$, corresponding to the same contraction operator interpretation as in Figure 2.

Systems and Solutions. Self-similar sets are often characterized as the unique non-empty compact sets that solve systems of equations of the form

$$K = \sigma_1(K) \cup \cdots \cup \sigma_n(K)$$

with each σ_i a contraction operator on a complete metric space. For example, the Sierpiński gasket is the unique non-empty compact solution to $K = \sigma_a(K) \cup \sigma_b(K) \cup \sigma_c(K)$ in Figure 1. In this section, we are going to provide a similar characterization for regular subfractals that will play an important role in the completeness proof in Section 4. The main connection of the characterization above to trace semantics arises from thinking about an n-state LTS (X, α) as a system of formal equations for the trace sets associated to the states:

$$x_i = a_{i1}x_{i1} + a_{i2}x_{i2} + \cdots + a_{im_i}x_{im_i}$$

indexed by $X = \{x_1, \dots, x_n\}$, where $x_i \xrightarrow{a_{ij}}_\alpha x_{ij}$ for $i \leq n$ and $j \leq m_i$.

For the next few definitions and results, fix a non-empty metric space M.

Definition 3.8. Write $\mathbf{K}(M)$ for the set of non-empty compact subsets of M. We equip $\mathbf{K}(M)$ with the structure of a metric space using the *Hausdorff metric* on $\mathbf{K}(M)$ [Gro15], defined by

$$d_H(K_1, K_2) = \max\left\{ \sup_{u \in K_1} \inf_{v \in K_2} d(u, v),\ \sup_{v \in K_2} \inf_{u \in K_1} d(u, v) \right\} \tag{3.6}$$

If M is complete, so is $\mathbf{K}(M)$. Incidentally, we need to restrict to *non-empty* sets in (3.6), since otherwise we would worry about infima over the empty set. This is the primary motivation for the guardedness condition which we imposed on our terms.

We also recall the *Banach fixed-point theorem*, which allows for the computation of fixed-points by iteration.

Theorem 3.9 Banach [Ban22]. *Let $f\colon M \to M$ be a contraction map.*

(1) *The map f has at most one fixed-point. That is, if $p = f(p)$ and $q = f(q)$, then $p = q$.*
(2) *If M is complete, then for any $p \in M$, $\lim_{n\to\infty} f^n(p)$ is the unique fixed-point of f.*

Definition 3.10. Given a contraction operator interpretation $\sigma : A \to \mathrm{Con}(M)$, and an LTS (X, α), we call a function $\varphi : X \to \mathbf{K}(M)$ a *(σ-)solution* to (X, α) if for any $x \in X$,

$$\varphi(x) = \bigcup_{x \xrightarrow{a} y} \sigma_a(\varphi(y))$$

Example 3.11. Let $\mathbf{S}$ be the Sierpiński gasket as a subset of $\mathbb{R}^2$. Let (X, α) be the LTS in Figure 1. Then we have a single state, x, with $x \xrightarrow{a,b,c} x$. The function $\varphi\colon X \to \mathbf{K}(\mathbb{R}^2)$ given by $\varphi(x) = \mathbf{S}$ is a solution to (X, α), because $\mathbf{S} = \sigma_a(\mathbf{S}) \cup \sigma_b(\mathbf{S}) \cup \sigma_c(\mathbf{S})$. Self-similar sets, in the sense of Hutchinson [Hut81], are regular subfractals generated by one-state LTSs.

We are very interested in the existence and uniqueness of solutions to finite productive LTSs in $\mathbf{K}(M)$ for a given space M, since these are our connection to subfractal subsets of M. We shall obtain a general result in this direction in Theorem 3.13 below. In fact, we shall show that *locally finite* productive LTSs have unique solutions. Our work is an application of the Banach Fixed Point Theorem, and it also calls on general metric and topological properties of the set A^ω of streams of actions.

Topology of Streams. Recall that, given a contraction operator interpretation $\sigma\colon A \to \mathrm{Con}(M)$, the subset of M generated by a state x in a LTS (X, α) is given by taking the image of the set of streams $\mathrm{str}_\alpha(x)$ emitted by x under the map $\sigma_\omega\colon A^\omega \to M$ defined in (3.1). As we are about to see, if M is complete, then subsets of M generated by states of (X, α) are always compact in the topology of M induced by the metric. Topological results like these follow from familiar topological properties of A^ω and the sets $\mathrm{str}_\alpha(x)$, which we discuss next.

Definition 3.12. Let A be a finite set of actions. For each $a \in A$, let $0 < c_a < 1$, and write $c\colon A \to (0,1)$ for the map $a \mapsto c_a$. Let $c_{\mathsf{dm}} > 0$. Define (A^ω, d_c) to be the space of streams from A equipped with the metric d_c, defined

$$d_c((a_1, a_2, \dots), (b_1, b_2, \dots)) = \begin{cases} 0 & \text{if } a_n = b_n \text{ for all } n \\ c_{\mathsf{dm}} \prod_{i=1}^n c_{a_i} & n = \min\{n \mid a_{n+1} \neq b_{n+1}\}, \text{ otherwise} \end{cases} \tag{3.7}$$

We call (A^ω, d_c) the *diameter-c_{dm} Cantor set equipped with the c-metric*.

For any choice in $c\colon A \to (0,1)$ and $c_{\mathsf{dm}} > 0$, (A^ω, d_c) is a compact totally disconnected metric space with no isolated points [LAS78]. The diameter-1 Cantor set equipped with the $1/2$-metric ($c_a = 1/2$ for all $a \in A$) is referred to as simply *the Cantor set*. The topology on A^ω induced by any c-metric has a basis of open sets consisting of the *cylinder sets*, subsets of A^ω of the form

$$B_{a_1 \cdots a_n} = \left\{ (b_i)_{i \in \mathbb{N}} \mid b_1 = a_1, \dots, b_n = a_n \right\} \tag{3.8}$$

Note that above, the case of $n = 0$ has $B_\epsilon = A^\omega$, where ϵ is the empty word.

Theorem 3.13. *Let (X, α) be a productive LTS, M be a complete metric space, and $\sigma\colon A \to \mathrm{Con}(M)$ a contraction operator interpretation. Then*

(1) *For all $x \in X$, $\mathrm{str}_\alpha(x)$ is non-empty and closed in A^ω.*
(2) *For all $x \in X$, $[\![x]\!]_{\alpha,\sigma}$ is non-empty and compact in M.*
(3) $[\![-]\!]_{\alpha,\sigma} : X \to \mathbf{K}(M)$ *is a solution to (X,α).*
(4) *If (X,α) is locally finite, then $[\![-]\!]_{\alpha,\sigma} : X \to \mathbf{K}(M)$ is the unique solution to (X,α).*

Proof. To see item 1, first observe that $\mathrm{str}_\alpha(x) \neq \emptyset$ because (X,α) is productive. Let $c \colon A \to (0,1)$ be any function and consider the diameter-1 Cantor set (A^ω, d_c). Now consider a Cauchy sequence $\{(a_1^{(i)}, a_2^{(i)}, \dots)\}_{i \in \mathbb{N}}$ in $\mathrm{str}_\alpha(x)$, and let $(a_i)_{i \in \mathbb{N}}$ be its limit in (A^ω, d_c). Then x emits every finite initial segment of $(a_i)_{i \in \mathbb{N}}$, because for any $N \in \mathbb{N}$, there is an $m > N$ such that $(a_1, \dots, a_m, a_{m+1}^{(m)}, \dots) \in \mathrm{str}_\alpha(x)$. Since (X,α) is finitely branching and productive, Lemma 2.3 tells us that $(a_1, a_2, \dots) \in \mathrm{str}_\alpha(x)$. Therefore, $\mathrm{str}_\alpha(x)$ is a non-empty and also closed set.

Item 2 is standard in the literature on fractals: it follows from the fact that (in our terminology) σ_ω is a surjective map whose image is $\mathbf{S}_\sigma \subseteq M$ as defined in (3.2), and this last set is the self-similar set defined by a contraction operator interpretation (in terminology from the fractals literature, an *iterated function system*). Such sets are known to be compact. Nevertheless, for the convenience of the reader we provide a proof from first principles.

For each $a \in A$, let $c_a \in (0,1)$ be a contraction coefficient for σ_a and let $m_a \in M$ be the fixed point of σ_a. Let

$$c_{\mathsf{max}} = \max_{a \in A} c_a \qquad \ell_{\mathsf{max}} = \max_{a,b \in A} d(m_a, m_b)$$

For each stream $(b_i)_{i \in \mathbb{N}} \in A^\omega$, $p \in M$, and $k \in \mathbb{N}$, let

$$h(p, (b_i)_{i \in \mathbb{N}}, k) = \sigma_{b_1} \cdots \sigma_{b_k}(p)$$

Then by definition, $\sigma_\omega((b_i)_{i \in \mathbb{N}}) = \lim_{k \to \infty} h(p, (b_i)_{i \in \mathbb{N}}, k)$ for all p. Now fix a point $p \in M$ and let

$$\ell_p = \min\{d(p, m_b) \mid b \in A\}$$

Fix $a \in A$ such that $d(p, m_a) = \ell_p$. We show by induction on $k \geq 1$ that for $(b_i)_{i \in \mathbb{N}} \in A^\omega$,

$$d(h(p, (b_i)_{i \in \mathbb{N}}, k), m_{b_1}) \leq c_{\mathsf{max}}^k \ell_p + \sum_{i=0}^{k-1} c_{\mathsf{max}}^i \ell_{\mathsf{max}} \tag{3.9}$$

Here is the argument for $k = 1$: We have $h(p, (b_i)_{i \in \mathbb{N}}, 1) = \sigma_{b_1}(p)$ and

$$d(\sigma_{b_1}(p), m_{b_1}) = d(\sigma_{b_1}(p), \sigma_{b_1}(m_{b_1})) \leq c_{\mathsf{max}}(d(p, m_a) + d(m_a, m_{b_1})) \leq c_{\mathsf{max}}(\ell_p + \ell_{\mathsf{max}})$$

Assuming that (3.9) holds for k, we fix $(b_i)_{i \in \mathbb{N}}$ and apply our induction hypothesis to the tail $(b_2, b_3, \dots)$. We have $d(h(p, (b_2, b_3, \dots), k), m_{b_2}) \leq c_{\mathsf{max}}^k \ell_p + \sum_{i=0}^{k-1} c_{\mathsf{max}}^i \ell_{\mathsf{max}}$. Then

$$
\begin{aligned}
&d(h(p, (b_i)_{i \in \mathbb{N}}, k+1), m_{b_1}) \\
&= d(\sigma_{b_1}(h(p, (b_2, b_3, \dots), k)), \sigma_{b_1}(m_{b_1})) && (m_{b_1} \text{ is the fixed point of } \sigma_{b_1}) \\
&\leq c_{b_1} d(h(p, (b_2, b_3, \dots), k), m_{b_1}) && (\text{defn. of } c_{b_1}) \\
&\leq c_{\mathsf{max}}(d(h(p, (b_2, b_3, \dots), k), m_{b_2}) + d(m_{b_2}, m_{b_1})) && (\text{triangle inequality} + \text{defn. of } c_{\mathsf{max}}) \\
&\leq c_{\mathsf{max}}\Big(\big(c_{\mathsf{max}}^k \ell_p + \sum_{i=0}^{k-1} c_{\mathsf{max}}^i \ell_{\mathsf{max}} \big) + \ell_{\mathsf{max}} \Big) && (\text{induction hypothesis} + \text{defn. of } \ell_{\mathsf{max}})
\end{aligned}
$$

$$\leq c_{\mathsf{max}}^{k+1}\ell_p + \sum_{i=0}^{k-1} c_{\mathsf{max}}^{i+1}\ell_{\mathsf{max}} + c_{\mathsf{max}}^0\ell_{\mathsf{max}} \qquad\qquad (\text{since } c_{\mathsf{max}}\ell_{\mathsf{max}} < \ell_{\mathsf{max}} = c_{\mathsf{max}}^0\ell_{\mathsf{max}})$$

$$\leq c_{\mathsf{max}}^{k+1}\ell_p + \sum_{i=1}^{k} c_{\mathsf{max}}^i\ell_{\mathsf{max}} + c_{\mathsf{max}}^0\ell_{\mathsf{max}}$$

$$= c_{\mathsf{max}}^{k+1}\ell_p + \sum_{i=0}^{(k+1)-1} c_{\mathsf{max}}^i\ell_{\mathsf{max}}$$

This completes the induction. Let k be large enough so that $c_{\mathsf{max}}^k\ell_p < 1$. Then we see that for large enough k, $d(h(p,(b_i)_{i\in\mathbb{N}},k),m_{b_1}) \leq \ell^\star$, where

$$\ell^\star = 1 + \sum_{i=0}^{\infty} c_{\mathsf{max}}^i\ell_{\mathsf{max}} = 1 + \frac{\ell_{\mathsf{max}}}{1 - c_{\mathsf{max}}}$$

This is independent of p. We conclude that $d(\sigma_\omega((b_i)_{i\in\mathbb{N}}),m_{b_1}) \leq \ell^\star$. Now let us obtain the diameter of the Cantor set we have in mind. Let $c_{\mathsf{dm}} = 2\ell^\star + \ell_{\mathsf{max}}$. For two streams in A^ω, say $(b_1,b_2,\ldots)$ and $(b_1',b_2',\ldots)$, we have

$$d(\sigma_\omega(b_1,b_2,\ldots),\sigma_\omega(b_1',b_2',\ldots))$$
$$\leq d(\sigma_\omega(b_1,b_2,\ldots),m_{b_1}) + d(m_{b_1},m_{b_1'}) + d(m_{b_1'},\sigma_\omega(b_1',b_2',\ldots))$$
$$\leq \ell^\star + \ell_{\mathsf{max}} + \ell^\star$$
$$= c_{\mathsf{dm}}$$

With this estimate in hand, we can now let $c\colon A \to (0,1)$ be the constant function $c(a) = c_{\mathsf{max}}$ and consider the diameter-c_{dm} Cantor set (A^ω, d_c) equipped with the c-metric.

Now we show that $\sigma_\omega\colon A^\omega \to M$ is nonexpanding with respect to d_c. Consider $(a_i)_{i\in\mathbb{N}}$ and $(b_i)_{i\in\mathbb{N}}$ in A^ω. We would like to show that $d_c(\sigma_\omega((a_i)_{i\in\mathbb{N}}),\sigma_\omega((b_i)_{i\in\mathbb{N}})) \leq d_c((a_i)_{i\in\mathbb{N}},(b_i)_{i\in\mathbb{N}})$. If $a_i = b_i$ for all i, then $(a_i)_{i\in\mathbb{N}} = (b_i)_{i\in\mathbb{N}}$ and we have

$$d_c((a_i)_{i\in\mathbb{N}},(b_i)_{i\in\mathbb{N}}) = 0 = d_M(\sigma_\omega((a_i)_{i\in\mathbb{N}}),\sigma_\omega((b_i)_{i\in\mathbb{N}}))$$

Otherwise, let $n = \min\{i \in \mathbb{N} \mid a_i \neq b_i\} - 1$. Then $d_c((a_i)_{i\in\mathbb{N}},(b_i)_{i\in\mathbb{N}}) = c_{\mathsf{dm}} \prod_{i=1}^n c_{a_i}$ since $b_i = a_i$ for all $i \leq n$. On the other side, we have

$$d_M(\sigma_\omega((a_i)_{i\in\mathbb{N}}),\sigma_\omega((b_i)_{i\in\mathbb{N}}))$$
$$= d_M\big(\sigma_{a_1}\cdots\sigma_{a_n}\sigma_\omega((a_{i+n})_{i\in\mathbb{N}}),\sigma_{a_1}\cdots\sigma_{a_n}\sigma_\omega((b_{i+n})_{i\in\mathbb{N}})\big)$$
$$\leq \Big(\prod_{i=1}^n c_{a_i}\Big) d_M\big(\sigma_\omega((a_{i+n})_{i\in\mathbb{N}}),\sigma_\omega((b_{i+n})_{i\in\mathbb{N}})\big)$$
$$\leq \Big(\prod_{i=1}^n c_{a_i}\Big) c_{\mathsf{dm}}$$
$$= d_c((a_i)_{i\in\mathbb{N}},(b_i)_{i\in\mathbb{N}})$$

Thus, σ_ω is nonexpanding. Since $\sigma_\omega\colon A^\omega \to M$ is nonexpanding, it is continuous. Then, by item 1, $[\![x]\!]_{\alpha,\sigma} = \sigma_\omega(\mathsf{str}_\alpha(x))$ is the image of the non-empty compact set $\mathsf{str}_\alpha(x)$ under the continuous function σ_ω. It follows that $[\![x]\!]_{\alpha,\sigma}$ is non-empty and compact.

We now show item 3, that $[\![x]\!]_{\alpha,\sigma} : X \to \mathbf{K}(M)$ is a solution to (X, α). The idea is to construct a function

$$F \colon \mathbf{K}(M)^X \to \mathbf{K}(M)^X$$

then to show that fixed points of F are exactly solutions to (X, α), and finally to show that $[\![x]\!]_{\alpha,\sigma} : X \to \mathbf{K}(M)$ is a fixed point of F. We take

$$F(f)(x) = \bigcup_{x \xrightarrow{a} y} \sigma_a(f(y)) \tag{3.10}$$

Let us check that each set $F(f)(x)$ really is a compact subset of M. There are two reasons for this. First, each set $\sigma_a(K)$ is compact, being the image of a compact set under a continuous function. Second, the finite union of compact sets is compact.

It is clear that a solution to (X, α) is exactly a fixed point of F. Let us check next that the map $[\![x]\!]_{\alpha,\sigma}$ is a fixed point. That is, we must show that for all $x \in X$,

$$[\![x]\!]_{\alpha,\sigma} = \bigcup_{x \xrightarrow{a} y} \sigma_a([\![y]\!]_{\alpha,\sigma}) \tag{3.11}$$

We are going to drop the subscripts and write $[\![x]\!]$. Let $\sigma_\omega((a_i)_{i\in\mathbb{N}}) \in [\![x]\!]$. This stream $(a_i)_{i\in\mathbb{N}}$ is emitted by x. Fix a corresponding infinite path $x \xrightarrow{a_1} x_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} x_n \xrightarrow{a_{n+1}} \cdots$. This path begins with $x \xrightarrow{a_1} x_1$. Its tail shows that the stream $(a_{i+1})_{i\in\mathbb{N}}$ is emitted by x_1 and hence belongs to $\mathrm{str}_\alpha(x_1)$. Thus, $\sigma_\omega((a_{i+1})_{i\in\mathbb{N}}) \in [\![x_1]\!]$. It follows from the commutativity of the diagram in (3.4) that

$$\sigma_\omega((a_i)_{i\in\mathbb{N}}) = (\sigma_\omega \circ \sigma_{a_1})((a_{i+1})_{i\in\mathbb{N}}) = \sigma_{a_1}(\sigma_\omega((a_{i+1})_{i\in\mathbb{N}})) \in \{\sigma_a([\![y]\!]_{\alpha,\sigma}) \mid x \xrightarrow{a} y\}$$

As desired, $\sigma_\omega((a_i)_{i\in\mathbb{N}})$ belongs to the right-hand side of (3.11). This shows that $[\![x]\!]_{\alpha,\sigma} \subseteq \bigcup_{x \xrightarrow{a} y} \sigma_a([\![y]\!]_{\alpha,\sigma})$. The reverse inclusion is similar.

We now turn to item 4. Let us first assume that X is finite. The more general case of locally finite X is considered at the very end of this proof. It is here that we use the fact that the set $\mathbf{K}(M)^X$ of functions $f \colon X \to \mathbf{K}(M)$ is a complete metric space with respect to the product metric $d(f, g) = \max_{x \in X} d_H(f(x), g(x))$, where d_H is the Hausdorff metric. By the previous part, we need to verify that the function F from (3.10) is contractive. Then, by the Banach Fixed Point Theorem, F has just one fixed point. Let $f, g \colon X \to \mathbf{K}(M)$, and let $\ell = d(f, g) = \max_x d(f(x), g(x))$. By the definition of the Hausdorff metric we have the following for all $x \in X$: for all $z \in f(x)$ there is some $w \in g(x)$ with $d(x, w) \leq \ell$, and vice-versa. Let $c_{\mathsf{max}} < 1$ be the maximum of the contraction coefficients of all σ_a for $a \in A$.

We claim that $d(F(f), F(g)) \leq \ell c_{\mathsf{max}}$. Let $x \in X$, and let $z \in F(f)(x)$. Then there is some $(a, y) \in \alpha(x)$ such that $z \in \sigma_a(f(y))$. For z, there is some $w \in f(y)$ such that $z = \sigma_a(w)$. Let $v \in g(y)$ be such that $d(w, v) < \ell$. Then $\sigma_a(v) \in \sigma_a(g(y)) \subseteq F(g)(x)$. Write $z' = \sigma_a(v)$. Then we have

$$d(z, z') = d(\sigma_a(w), \sigma_a(v)) < c_{\mathsf{max}} d(w, v) < \ell c_{\mathsf{max}}$$

So, for every $z \in F(f)(x)$, we have some $z' \in F(g)(x)$ with $d(z, z') < \ell c_{\mathsf{max}}$. The converse assertion also holds. Thus, $d(F(f)(x), F(g)(x)) < \ell c_{\mathsf{max}}$. Since x is arbitrary, we obtain $d(F(f), F(g)) < \ell c_{\mathsf{max}}$.

At this point, we know that $[\![-]\!]_{\alpha,\sigma}$ is the only solution to (X, α) when X is finite. We also want to show that F has a unique solution when X is merely locally finite, but perhaps infinite. Fix $x \in X$, and also take any solution $\varphi \colon X \to \mathbf{K}(M)$. Let us write the finite

sub-LTS of X determined by x as $(\langle x \rangle, \alpha_0)$. The restriction $\varphi|_{\langle x \rangle}$ is a solution to $(\langle x \rangle, \alpha_0)$. By what we already know, $(\varphi|_{\langle x \rangle})(x) = [\![x]\!]_{\alpha_0, \sigma}$. Now

$$\varphi(x) = (\varphi|_{\langle x \rangle})(x) = [\![x]\!]_{\alpha_0, \sigma}$$

Since the right-hand side above does not involve φ, we see that φ is unique. $\qquad\square$

Remark 3.14. One step of the proof above was to show that for any given contraction operator interpretation $\sigma \colon A \to \mathrm{Con}(M)$ there is a map $c \colon A \to (0,1)$ such that $\sigma_\omega \colon A^\omega \to M$ is nonexpanding with respect to the c-metric d_c on A^ω. This implies that σ_ω is continuous with respect to d_c. But any two c-metrics induce the same topology on A^ω, so this actually tells us that σ_ω is continuous for any contraction operator interpretation σ. Of course, there is another way to arrive at the same fact, using an interchange of limits: given a sequence of streams $\{(a_i^{(j)})_{i \in \mathbb{N}}\}_{j \in \mathbb{N}}$ that converges to $(a_i)_{i \in \mathbb{N}}$, we know that for any $i \in \mathbb{N}$, there is an $l > 0$ such that for any $j > l$, $a_i^{(j)} = a_i$. For arbitrary $p \in M$, this implies that for any $i \in \mathbb{N}$, $\lim_{j \to \infty} \sigma_{a_1^{(j)}} \cdots \sigma_{a_i^{(j)}}(p) = \sigma_{a_1} \cdots \sigma_{a_i}(p)$. We already know that $\sigma_\omega((a_i^{(j)})_{i \in \mathbb{N}}) = \lim_{i \to \infty} \sigma_{a_1^{(j)}} \cdots \sigma_{a_i^{(j)}}(p)$ converges for each $j \in \mathbb{N}$. Thus, the double sequence $(\sigma_{a_1^{(j)}} \cdots \sigma_{a_i^{(j)}}(p))_{i,j \in \mathbb{N}}$ converges along every row and column. By the Moore-Osgood theorem, this allows us to exchange limits:

$$\begin{aligned}
\lim_{j \to \infty} \sigma_\omega((a_i^{(j)})_{i \in \mathbb{N}}) &= \lim_{j \to \infty} \lim_{i \to \infty} \sigma_{a_1^{(j)}} \cdots \sigma_{a_i^{(j)}}(p) \\
&= \lim_{i \to \infty} \lim_{j \to \infty} \sigma_{a_1^{(j)}} \cdots \sigma_{a_i^{(j)}}(p) \\
&= \lim_{i \to \infty} \sigma_{a_1} \cdots \sigma_{a_i}(p) \\
&= \sigma_\omega((a_i)_{i \in \mathbb{N}}) \\
&= \sigma_\omega\big(\lim_{j \to \infty} \{(a_i^{(j)})_{i \in \mathbb{N}}\}_{j \in \mathbb{N}} \big)
\end{aligned}$$

This provides a second argument showing that σ_ω is continuous.

Example 3.15. To see why some extra assumption is needed in item 4, consider the LTS $x_0 \xrightarrow{a} x_1 \xrightarrow{a} x_2 \xrightarrow{a} \cdots$. The smallest sub-LTS containing x_0 is the whole LTS, so this LTS is not locally finite. Let $M = \mathbb{R}$, and let $\sigma_a(x) = \frac{1}{2}x$. One solution is $[\![x_n]\!]_{\alpha, \sigma} = \{0\}$ for all n. Another solution is $\varphi(x_n) = \{2^n\}$. Indeed, for all $r \in \mathbb{R}$, we have a solution $\varphi(x_n) = \{r2^n\}$. What is going on in this example is that $\mathbf{K}(M)^X$ is not a metric space under our definitions: the singleton map is an isometric embedding of any space M into $\mathbf{K}(M)$. Thus $d(\{2^n\}, \{0\}) = 2^n$. So $\sup_n d(\{2^n\}, \{0\}) = \infty$. But all distances are finite in this paper, as this is one of the requirements of the Banach fixed point theorem.

The following consequence of Theorem 3.13 is an analog of Kleene's theorem (the correspondence between regular expressions and regular languages [Kle56]) for regular subfractals.

Corollary 3.16. *Fix a complete metric space M and a contraction operator interpretation $\sigma \colon A \to \mathrm{Con}(M)$. A set is a regular subfractal if and only if it is a component of the solution to a finite productive LTS.*

Proof. Let S be a regular subfractal as defined in Definition 3.5. So, there exists $e \in \mathsf{Term}$ such that $[\![e]\!]_{\gamma, \sigma} = S$. By Definition 3.10, for each $f \in \mathsf{Term}$, $[\![f]\!]_{\gamma, \sigma} = \bigcup_{f \xrightarrow{a} g} \sigma_a([\![g]\!]_{\gamma, \sigma})$.

Consider the finite productive LTS $\langle e \rangle$ generated by e. Let us write γ_0 for the restriction of γ to $\langle e \rangle$. Define $\varphi \colon \langle e \rangle \to \mathbf{K}(M)$ by $\varphi(g) = [\![g]\!]_{\gamma,\sigma}$. Then φ is a solution to $\langle e \rangle$. So, by Theorem 3.13, $\varphi = [\![-]\!]_{\gamma_0,\sigma}$. And so, $S = [\![e]\!]_{\gamma,\sigma} = \varphi(e) = [\![e]\!]_{\gamma_0,\sigma}$. Thus, S is a component of the solution to the finite productive LTS $\langle e \rangle$.

For the converse, let (X, α) be a finite productive LTS, and take a component of its solution, say $[\![x]\!]_{\alpha,\sigma}$. By Lemma 2.8, there is a term $e \in \mathsf{Term}$ such that $\mathrm{tr}_\gamma(e) = \mathrm{tr}_\alpha(x)$. By Lemma 2.3, $\mathrm{str}_\gamma(e) = \mathrm{str}_\alpha(x)$. By Definition 3.4, $[\![e]\!]_{\gamma,\sigma} = [\![x]\!]_{\alpha,\sigma}$. Thus, $[\![x]\!]_{\alpha,\sigma}$ is a regular subfractal. $\qquad\square$

Example 3.17. The following example of our setup is essential for the rest of the paper. Let $c \colon A \to (0,1)$ and let (A^ω, d_c) be a diameter-c_{dm} Cantor set equipped with the c-metric. For each $a \in A$, recall the a-*prefixing operator* $a \colon A^\omega \to A^\omega$ from (3.3). In the c-metric, $a(-)$ is a contraction operator with contraction coefficient c_a, so the map $\sigma \colon A \to \mathrm{Con}(A^\omega, d_c)$ defined $\sigma_a = a(-)$ is a contraction operator interpretation.

A fact which we shall need is that the map σ_ω in this example is the identity on A^ω. To this end, observe that

$$\sigma_\omega((a_1, a_2, \dots)) = \lim_{n \to \infty} \sigma_{a_1} \circ \cdots \circ \sigma_{a_n}(p) = \lim_{n \to \infty} a_1 a_2 \cdots a_n(p) = (a_1, a_2, \dots)$$

for any $p \in A^\omega$. Since (A^ω, d_c) is a compact metric space, it is in particular complete. Thus, Theorem 3.13 applies to this example. Since σ_ω is the identity on A^ω, $[\![x]\!]_{\alpha,\sigma}$ is simply $\mathrm{str}_\alpha(x)$ for a given productive LTS (X, α) and $x \in X$. It follows from Theorem 3.13 that $[\![x]\!]_{\alpha,\sigma} = \mathrm{str}_\alpha(x)$, hence $\mathrm{str}_\alpha \colon X \to A^\omega$ is the unique solution to (X, α) in A^ω.

Real-valued Cantor sets. Traditional real-valued Cantor sets (like the ternary Cantor set) are also able to faithfully represent streams. Here, by *traditional real-valued Cantor set*, we mean any self-similar set corresponding to a contraction operator interpretation $\sigma \colon A \to \mathrm{Con}(\mathbb{R})$ of the form $\sigma_a(p) = cp + r_a$ for some $c \in (0,1)$ and distinct $r_a \in \mathbb{R}$ for each $a \in A$. We will revisit this point in Section 4.

A map $f \colon \mathbb{R} \to \mathbb{R}$ of the form $f(p) = cp + r$ for some $c, r \in \mathbb{R}$ is called *affine*. With $|c| < 1$, this operator is a contraction with coefficient c. An *affine operator interpretation* is a contraction operator interpretation $\sigma \colon A \to \mathrm{Con}(\mathbb{R})$ where each σ_a is affine.

Proposition 3.18. *Let A be a finite set of actions. There exists an affine operator interpretation $\sigma \colon A \to \mathrm{Con}(\mathbb{R})$ such that $\sigma_\omega \colon A^\omega \to \mathbb{R}$ is injective.*

Proof. Let $A = \{a_1, \cdots a_n\}$ and write $c = 1/(2n - 1)$. Now consider the affine operator interpretation $\sigma \colon A \to \mathrm{Con}(\mathbb{R})$ given by

$$\sigma_{a_i}(p) = cp + 2c(i - 1)$$

for each $i \leq n$. Note that $\mathbf{S}_\sigma \subseteq [0,1]$. We are going to show that $\sigma_\omega \colon A^\omega \to \mathbb{R}$ is injective.

To this end, suppose that $d_c((b_i)_{i \in \mathbb{N}}, (b_i')_{i \in \mathbb{N}}) > 0$. Then there is an $i \in \mathbb{N}$ such that $b_i \neq b_i'$. Let $m = \min\{i \mid b_i \neq b_i'\} - 1$. We first handle the case $m = 0$, i.e., $b_1 \neq b_1'$. Let i and j be such that $b_1 = a_i$ and $b_1' = a_j$. Note that $i \neq j$, and indeed $|i - j| \geq 1$. Then for any $p, q \in \mathbb{R}$,

$$|\sigma_{b_1}(p) - \sigma_{b_1'}(q)| = |cp + 2c(i - 1) - cq - 2c(j - 1)| = c|(p - q) + 2(i - j)|$$

so that for any $p, q \in [0,1]$, $|(p - q) + 2(i - j)| > 0$. In particular, if $b_1 \neq b_1'$, then

$$|\sigma_\omega((b_i)_{i \in \mathbb{N}}) - \sigma_\omega((b_i')_{i \in \mathbb{N}})| = |\sigma_{b_1}(\sigma_\omega((b_{i+1})_{i \in \mathbb{N}})) - \sigma_{b_1'}(\sigma_\omega((b_{i+1}')_{i \in \mathbb{N}}))| > 0$$

When $m > 0$, we reduce to the first case:

$$
\begin{aligned}
&|\sigma_\omega((b_i)_{i\in\mathbb{N}}) - \sigma_\omega((b_i')_{i\in\mathbb{N}})| \\
&= |\sigma_{b_1}\cdots\sigma_{b_m}\sigma_\omega((b_{i+m})_{i\in\mathbb{N}}) - \sigma_{b_1}\cdots\sigma_{b_m}\sigma_\omega((b_{i+m}')_{i\in\mathbb{N}})| \\
&= c^m|\sigma_\omega((b_{i+m})_{i\in\mathbb{N}}) - \sigma_\omega((b_{i+m}')_{i\in\mathbb{N}})| \\
&> c^m 0 = 0
\end{aligned}
$$

since $b_{i+m} \neq b_{i+m}'$. It follows that σ_ω is injective.[4] □

4. Completeness

We have seen that finite productive LTSs (LTSs that only emit infinite streams) can be specified by process terms. We also introduced a family of fractal sets called regular subfractals, those subsets of self-similar sets obtained from the streams emitted by finite productive LTSs. An LTS itself is representative of a certain system of equations, and set-wise the system of equations is solved by the regular subfractals corresponding to it. Going from process terms to LTSs to regular subfractals, we see that a process term is representative of a sort of *uninterpreted fractal recipe*, which tells us how to obtain a regular subfractal from an interpretation of action symbols as contractions on a complete metric space.

Definition 4.1. Let (X, α) be a productive LTS. Given $x, y \in X$, we write $x \approx y$ and say that x and y are *fractal equivalent* (or that they are *equivalent fractal recipes*) if $[\![x]\!]_{\alpha,\sigma} = [\![y]\!]_{\alpha,\sigma}$ for every complete metric space M and every $\sigma : A \to \mathrm{Con}(M)$.

Theorem 4.2. *Let (X, α) be a finitely branching productive LTS and $x, y \in X$. Then $x \approx y$ if and only if $\mathrm{str}_\alpha(x) = \mathrm{str}_\alpha(y)$.*

Proof. Clearly, if $\mathrm{str}_\alpha(x) = \mathrm{str}_\alpha(y)$, then $x \approx y$ by Definition 3.4. Conversely, consider Example 3.17, i.e., for a map $c : A \to (0, 1)$ and diameter $c_{\mathsf{dm}} > 0$, a diameter-c_{dm} Cantor set equipped with the c-metric and the contraction operator interpretation given by prefixing. In Example 3.17, we argued that the unique solution to (X, α) with this contraction operator interpretation is the map $\mathrm{str}_\alpha : X \to A^\omega$. Thus, if $x \approx y$, then $\mathrm{str}_\alpha(x) = \mathrm{str}_\alpha(y)$. □

This gives way to a soundness/completeness theorem for our version of Rabinovich's logic with respect to its fractal semantics. Our proof relies on the axiomatization of trace equivalence that we saw in Theorem 2.10.

Theorem 4.3 (Soundness/Completeness). *Let $e, f \in \mathsf{Term}$. Then*

$$e \approx f \quad \text{if and only if} \quad e \equiv f$$

Proof. Let $e, f \in \mathsf{Term}$. Then since (Term, γ) is finitely branching and productive,

$$e \approx f \overset{(\text{Theorem } 4.2)}{\Longleftrightarrow} \mathrm{str}(e) = \mathrm{str}(f) \overset{(\text{Lemma } 2.3)}{\Longleftrightarrow} \mathrm{tr}(e) = \mathrm{tr}(f) \overset{(\text{Theorem } 2.10)}{\Longleftrightarrow} e \equiv f \qquad □$$

One interesting consequence of Theorems 4.3 and 3.18 is that one may restrict their attention to affine contraction operator interpretations on $\mathbb{R}$ and still obtain completeness.

---

[4] Another way to say all this is that the restrictions of σ_{a_i} and σ_{a_j} to $[0, 1]$ have disjoint image sets.

Corollary 4.4. *Let $e, f \in$ Term. Then $e \equiv f$ if and only if for every affine contraction operator interpretation $\sigma\colon A \to \mathrm{Con}(\mathbb{R})$, $[\![e]\!]_{\gamma,\sigma} = [\![f]\!]_{\gamma,\sigma}$.*

Proof. The forward direction is soundness. For the reverse direction, by Proposition 3.18, there is a contraction operator interpretation $\sigma\colon A \to \mathrm{Con}(\mathbb{R})$ such that $\sigma_\omega\colon A^\omega \to \mathbb{R}$ is injective. It follows that the setwise extension of σ_ω (by image sets) is injective. This implies that if $[\![e]\!]_{\gamma,\sigma} = [\![f]\!]_{\gamma,\sigma}$, then $\mathrm{str}_\gamma(e) = \mathrm{str}_\gamma(f)$, and therefore $e \approx f$. From Theorem 4.3, we obtain $e \equiv f$. $\qquad\square$

5. Subfractal Measures

In addition to showing the existence of self-similar sets and their correspondence with contraction operator interpretations (in Hutchinson's terminology, iterated function systems), Hutchinson also shows that every probability distribution on the contractions corresponds to a unique measure, called the *invariant measure*, that satisfies a certain recursive equation and whose support is the self-similar set. In this section, we replay the story up to this point, but with Hutchinson's invariant measure construction instead of the invariant (self-similar) set construction. We specify fractal measures using a probabilistic version of LTSs called *labelled Markov chains*, as well as a probabilistic version of Milner's specification language introduced by Stark and Smolka [SS00]. Similar to how fractal equivalence coincides with trace equivalence, *fractal measure equivalence* is equivalent to a probabilistic version of trace equivalence due to Kerstan and König [KK13].

Invariant measures. Recall that a *Borel probability measure* on a metric space M is a $[0,1]$-valued function ρ defined on the Borel subsets of M (the smallest σ-algebra containing the open balls of M) that is countably additive and assigns $\rho(\emptyset) = 0$ and $\rho(M) = 1$. A Borel probability measure ρ is *supported by a point $p \in M$* if for any open set U containing p, $\rho(U) > 0$. The *support* $\mathrm{supp}(\rho)$ of a Borel measure ρ is the set of points supporting ρ. A Borel probability measure is *compactly supported* if the set of points supporting it is a compact set. Note that the support of a Borel measure is always a topologically closed set, so a Borel measure is compactly supported if and only if its support is bounded. For a more detailed introduction to Borel measures, see for example [Fol99].

Hutchinson shows in [Hut81] that, given $\sigma\colon A \to \mathrm{Con}(M)$, each probability distribution $\theta\colon A \to [0,1]$ on A gives rise to a unique compactly supported Borel probability measure $\hat{\theta}_\sigma$ on M, called the *invariant measure*, that is supported by the self-similar set $\mathbf{S}_\sigma$ from (3.2) and such that for any Borel set $B \subseteq M$,

$$\hat{\theta}_\sigma(B) = \sum_{a \in A} \theta(a)\; \sigma_a^{\#}\hat{\theta}_\sigma(B) \tag{5.1}$$

Here and elsewhere, the *pushforward* $f^{\#}\rho$ of a Borel measure ρ with respect to a continuous (or more generally, Borel measurable) map f is the probability measure defined by $f^{\#}\rho(B) = \rho(f^{-1}(B))$ for any Borel subset B of M [Fol99]. The pushforwards $\sigma_a^{\#}\hat{\theta}_\sigma$ in (5.1) are well-defined because contractions are continuous maps. Equation 5.1 tells us that, similar to how self-similar sets are the union of their images over a set of contractions, the invariant measure $\hat{\theta}_\sigma$ is a convex combination of its pushforwards with respect to a set of contractions. We can use the first part of the Banach Fixed Point Theorem, Theorem 3.9 (1), to show that (5.1) is satisfied by at most one compactly supported Borel measure on M.

Fix a complete metric space M.

Definition 5.1. We write $\mathsf{Prob}(M)$ for the set of compactly supported Borel probability measures on M equipped with the *Kantorovich-Rubinstein metric*, defined by

$$d(\rho_1, \rho_2) = \sup\left\{ \int_M f \,\mathsf{d}\rho_1 - \int_M f \,\mathsf{d}\rho_2 \ \Big| \ f\colon M \to \mathbb{R} \text{ is nonexpanding}\right\}$$

for any $\rho_1, \rho_2 \in \mathsf{Prob}(M)$. Above, $\int_M f \,\mathsf{d}\rho_i$ is given by the Lebesgue integral.[5]

It is important to note that finiteness of $d(\rho_1, \rho_2)$ requires the assumption that ρ_1 and ρ_2 are compactly supported (equivalently, have bounded support).

Definition 5.2. Given a fixed probability distribution $\theta\colon A \to [0,1]$, the *Markov operator* $F\colon \mathsf{Prob}(M) \to \mathsf{Prob}(M)$ is defined by

$$F(\rho) = \sum_{a \in A} \theta(a) \ \sigma_a^{\#}\rho \tag{5.2}$$

for each $\rho \in \mathsf{Prob}(M)$.

A measure $\hat{\theta}_\sigma$ satisfies (5.1) if and only if it is a fixed point of F. Thus, we would like to use fixed-point techniques to show the existence and uniqueness of invariant measures. Indeed, Hutchinson proves in [Hut81, Theorem 4.4(1)] that F is a contraction in the Kantorovich-Rubinstein metric. Therefore, if F has a fixed-point, then Theorem 3.9(1) tells us it is unique, and so the invariant measure also would be unique.

However, it is not true for an arbitrary complete metric space M that $\mathsf{Prob}(M)$ is complete, so we cannot apply Theorem 3.9(2) to prove the existence of the invariant measure.[6] In the next two paragraphs, we describe an explicit construction of the invariant measure. A more general construction and verification of its stated properties can be found in Theorem 5.15.

The invariant measure can be concretely described as follows. Given $\sigma\colon A \to \mathrm{Con}(M)$, let $c(a) = c_a > 0$ be a contraction coefficient for σ_a for each $a \in A$, and consider the diameter-1 Cantor set equipped with the c-metric (A^ω, d_c). Now, given a probability distribution $\theta\colon A \to [0,1]$, we define $\hat{\theta}$, called the *stream measure specified by θ*, to be the unique Borel probability distribution on (A^ω, d_σ) such that

$$\hat{\theta}(B_w) = \prod_{i=1}^{n} \theta(a_i) \tag{5.3}$$

for any $w = a_1 \cdots a_n \in A^*$ (here, we take the empty product to be 1). As will be explained shortly, results from [KK13] tell us that (5.3) does indeed extend to a unique Borel probability measure on (A^ω, d_σ). This again relies on the fact that any two Borel measures that agree on all basic open sets must agree on all Borel sets. The next proposition tells us that the invariant measure is precisely $\hat{\theta}_\sigma = \sigma_\omega^{\#}\hat{\theta}$, where σ_ω is the map given by limits which we saw in Definition 3.2. We will prove a more general version of this result in Theorem 5.15.

---

[5]Note that we use the symbol d in the integral to represent the differential, so as not to clash with d as in *distance*.

[6]In his original paper [Hut81], Hutchinson fallaciously relies on the Banach fixed-point theorem to construct the invariant measure. This minor hiccup can be remedied by either restricting the space of measures in a suitable way [Kra06], by assuming M is compact, or by constructing the invariant measure explicitly. We opt for the latter approach.

Proposition 5.3. *Let M be a complete metric space, let $\sigma\colon A \to \mathrm{Con}(M)$, and let F be the Markov operator defined in (5.2). The unique fixed-point of F is the pushforward $\sigma_\omega^{\#}\hat{\theta}$, where $\sigma_\omega\colon A^\omega \to M$ is from Definition 3.2, and $\hat{\theta} \in \mathsf{Prob}(A^\omega)$ is defined in (5.3).*

In the sequel, we are going to view the specification θ of the invariant measure $\hat{\theta}_\sigma$ (the pushforward of the stream measure) as a probability distribution over the transitions in a one-state LTS. This construction generalizes to multiple-state transition systems by moving from probability distributions on A to *labelled Markov chains*, which are essentially LTSs where each state comes equipped with a probability distribution on its outgoing transitions. Again, the labels will be interpreted as contractions on a complete metric space.

Algebras, Coalgebras, and Corecursive Algebras. At this point we make a short digression into the general notions of *universal algebra* and *universal coalgebra*. A standard reference for universal algebra is [BS81], although we will be formulating its notions using category-theoretic language, in the style of [AT89]. Universal coalgebra is a category-theoretic framework for studying the many different types of transition systems appearing in computer science in a uniform way. A standard reference for universal coalgebra is [Rut00].

The purpose of this digression is to ease the formulation of *labelled Markov chains* and *coalgebra-to-algebra homomorphisms* that will play a significant role in the rest of the paper. Labelled Markov chains are a probabilistic analogue of labelled transition systems. While labelled transition systems can mostly be studied without the categorical language of universal coalgebra, the language of coalgebra greatly improves ease-of-use for probabilistic systems [Sok11]. Following our digression, we will return to the topic of subfractal measures.

Definition 5.4. Let $H\colon \mathcal{C} \to \mathcal{C}$ be an endofunctor on any category. An H-*algebra* is a pair (C, γ) consisting of an object C of $\mathcal{C}$ and a morphism $\gamma\colon HC \to C$. An H-*algebra homomorphism* $h\colon (C, \gamma_C) \to (D, \gamma_D)$ is a morphism $h\colon C \to D$ such that $\gamma_D \circ H(h) = h \circ \gamma_C$.

An H-*coalgebra* is a pair (X, β) consisting of an object X of $\mathcal{C}$ and a morphism $\beta\colon X \to HX$. An H-*coalgebra homomorphism* $h\colon (X, \beta_X) \to (Y, \beta_Y)$ is a morphism $h\colon X \to Y$ such that $\beta_Y \circ h = H(h) \circ \beta_X$.

In this paper, we are mainly interested in two endofunctors on the category Set of sets and functions. First, $FX = \mathcal{P}_*(A \times X)$ where $\mathcal{P}_*$ is the non-empty finite power set functor and A is our set of actions (a fixed finite set in all of our work). Given $f\colon X \to Y$, $F(f)\colon \mathcal{P}_*(A \times X) \to \mathcal{P}_*(A \times Y)$ is given by $F(f)(U) = \{(a, f(x)) \mid (a, x) \in U\}$. The second functor of interest is $GX = \mathcal{D}(A \times X)$. Here, $\mathcal{D}(X)$ is the set of finitely supported probability distributions on X, i.e., functions $\theta\colon X \to [0,1]$ such that $\theta(x) = 0$ for all but finitely many $x \in X$, and $\sum_{x \in X} \theta(x) = 1$. This time, $G(f)(\theta)(y) = \sum\{\theta(x) \mid f(x) = y\}$ for all $y \in Y$.

Example 5.5. Due to Lemma 2.3(2), productive and finitely branching LTSs are exactly the coalgebras of $\mathcal{P}_*$, the non-empty finite power set functor.

Later in the paper, we will see algebras and coalgebras of $\mathcal{P}_*(A \times (-))$ and $\mathcal{D}(A \times (-))$, and in due course we will name them as shown below.

Set Functor	Name of Algebra	Name of Coalgebra
$FX = \mathcal{P}_*(A \times X)$	labelled transition algebra	productive, finitely branching labelled transition system
$FX = \mathcal{D}(A \times X)$	labelled Markov algebra	labelled Markov chain

Aside from the specific examples of labelled transition algebras and labelled Markov algebras that are of direct concern to us in this paper, there are two motivating examples that may help the reader. One such example is the labelled transition algebra of infinite (unordered) trees decorated by labels from A. There, the algebra structure turns a set $\{(a_1, t_1), \ldots, (a_n, t_n)\}$ into the tree whose immediate subtrees are $t_1, \ldots, t_n$ and outgoing root-to-child relations are labelled with $a_1, \ldots, a_n$ respectively. A potentially helpful example of a labelled Markov algebra is similar to the infinite trees from before, but whose parent-to-child relationships carry (nonzero) probabilities in addition to labels.

Coalgebras for $\mathcal{P}_*(A \times (-))$ and $\mathcal{D}(A \times (-))$ that are of particular interest to us are the ones that are finite around each state.

Definition 5.6. Let $H \colon \mathsf{Set} \to \mathsf{Set}$ be an endofunctor and (X, β) an H-coalgebra. A *subcoalgebra* of (X, β) is an H-coalgebra (U, β_U) such that $U \subseteq X$ and the inclusion map $U \hookrightarrow X$ is a coalgebra homomorphism. We say that (X, β) is *finite* if X is finite, and *locally finite* if every state $x \in X$ is included in a finite subcoalgebra of (X, β).

We also have the notion of a *finitely corecursive algebra* (Definition 5.7), which is a weaker version of the corecursive algebras first introduced by Capretta et al [CUV06]. Finitely corecursive algebras are also weaker than *iterative algebras*, and these are weaker than the better-known notion of *completely iterative algebras* [BÉ93]. Connections between completely iterative algebras and fractals date as far back as [MM06].

Definition 5.7. Let $H \colon \mathsf{Set} \to \mathsf{Set}$ be an endofunctor. An algebra $\gamma \colon HC \to C$ is *finitely corecursive* if for every locally finite H-coalgebra (X, β), there is a unique *coalgebra-to-algebra homomorphism* $\beta^\dagger \colon (X, \beta) \to (C, \gamma)$, i.e., a map $\beta^\dagger \colon X \to C$ such that $\beta^\dagger = \gamma \circ H(\beta^\dagger) \circ \beta$:

$$\begin{array}{ccc} X & \xrightarrow{\ \beta^\dagger\ } & C \\ {\scriptstyle \beta}\downarrow & & \uparrow{\scriptstyle \gamma} \\ HX & \xrightarrow[\ H(\beta^\dagger)\ ]{} & HC \end{array}$$

The map $\beta^\dagger$ is also called *the solution of β in the algebra (C, γ)*.

Lemma 5.8. *Let $\gamma \colon HC \to C$ be a finitely corecursive algebra for the endofunctor H. Let (X, β_X) and (Y, β_Y) be locally finite coalgebras, and let $\varphi \colon X \to Y$ be a coalgebra morphism. Then $\beta_X^\dagger = \beta_Y^\dagger \circ \varphi$.*

Proof. The proof is indicated in the diagram below.

$$\begin{array}{ccccc} X & \xrightarrow{\ \varphi\ } & Y & \xrightarrow{\ \beta_Y^\dagger\ } & C \\ {\scriptstyle \beta_X}\downarrow & & {\scriptstyle \beta_Y}\downarrow & & \uparrow{\scriptstyle \gamma} \\ HX & \xrightarrow[\ H(\varphi)\ ]{} & HY & \xrightarrow[\ H(\beta_Y^\dagger)\ ]{} & HC \end{array}$$

The commutativity of the square on the left expresses that φ is a coalgebra morphism, and that the square on the right commutes expresses that $\beta_Y^\dagger$ is a solution to (Y, β_Y) in C. Since both squares commute, so does the overall outside. This shows that $\beta_Y^\dagger \circ \varphi$ is a solution to (X, β) in C. By uniqueness of solutions, $\beta_Y^\dagger \circ \varphi = \beta_X^\dagger$. $\qquad\square$

The following lemma allows one to check finite corecursiveness by only considering finite coalgebras.

Lemma 5.9. *Let $H\colon \mathsf{Set} \to \mathsf{Set}$ be an endofunctor and let $\gamma\colon HC \to C$ be an algebra. Then (C, γ) is a finitely corecursive algebra if and only if for any finite H-coalgebra (X, β), there is a unique coalgebra-to-algebra homomorphism $(X, \beta) \to (C, \gamma)$.*

Proof. The forward direction follows from the fact that finite coalgebras are locally finite.

For the reverse direction, let (X, β) be a locally finite coalgebra. Every state of (X, β) is contained in a finite subcoalgebra and by assumption, finite subcoalgebras admit unique coalgebra-to-algebra homomorphisms into (C, γ). We start with the claim that there can be at most one coalgebra-to-algebra homomorphism $(X, \beta) \to (C, \gamma)$.

Suppose that $\varphi, \psi\colon (X, \beta) \to (C, \gamma)$ are coalgebra-to-algebra homomorphisms and let $x \in X$. We aim to show that $\varphi(x) = \psi(x)$. Since (X, β) is locally finite, there is a finite subcoalgebra (U, β_U) of (X, β) such that $x \in U$. Let $\iota_U\colon U \hookrightarrow X$ be the inclusion map of U into X, and recall that (by definition) $\iota_U\colon (U, \beta_U) \to (X, \beta)$ is a coalgebra homomorphism. The composition of a coalgebra homomorphism with a coalgebra-to-algebra homomorphism is a coalgebra-to-algebra homomorphism by Lemma 5.8, so in particular this holds for $\varphi \circ \iota_U$ and $\psi \circ \iota_U$.

$$
\begin{array}{ccc}
U & \xrightarrow{\ \iota_U\ } X & \overset{\varphi}{\underset{\psi}{\rightrightarrows}} C \\
\beta_U \downarrow & \quad \downarrow \beta & \quad \uparrow \gamma \\
HU & \xrightarrow[\ H(\iota_U)\ ]{} HX & \overset{H(\varphi)}{\underset{H(\psi)}{\rightrightarrows}} HC
\end{array}
$$

We have assumed that (U, β_U) is finite, so there is at most one coalgebra-to-algebra homomorphism $(U, \beta_U) \to (C, \gamma)$. It follows that $\varphi \circ \iota_U = \psi \circ \iota_U$, and in particular,

$$\varphi(x) = \varphi \circ \iota_U(x) = \psi \circ \iota_U(x) = \psi(x)$$

This shows that there is at most one coalgebra-to-algebra homomorphism $(X, \beta) \to (C, \gamma)$.

It now suffices to construct a coalgebra-to-algebra homomorphism $(X, \beta) \to (C, \gamma)$. For each finite subcoalgebra (U, β_U) of (X, β), let $\varphi_U\colon (U, \beta_U) \to (C, \gamma)$ be the unique coalgebra-to-algebra homomorphism. We define the relation

$$\varphi = \{(x, y) \mid x \in X \text{ and there is a finite subcoalgebra } (U, \beta_U) \text{ with } y = \beta_U(x)\}$$

We need to show that φ is a function and that it is a coalgebra-to-algebra homomorphism.

To see that it is a function, suppose $x \in U \cap V$ for two finite subcoalgebras (U, β_U) and (V, β_V), and let $(x, y), (x, y') \in \varphi$ such that $y = \varphi_U(x)$ and $y' = \varphi_V(x)$. The intersection of two subcoalgebras is also a subcoalgebra [GS02, Theorem 3.1], so if we set $W = U \cap V$, W is also the state space of a finite subcoalgebra (W, β_W). Now, because the composition of a coalgebra-to-algebra homomorphism with a coalgebra homomorphism is always a coalgebra-to-algebra homomorphism (again using Lemma 5.8), φ_W, $\varphi_U \circ \iota_W$, and $\varphi_V \circ \iota_W$ are all coalgebra-to-algebra homomorphisms $(W, \beta_W) \to (C, \gamma)$. By uniqueness of coalgebra-to-algebra homomorphisms for finite coalgebras, $\varphi_U \circ \iota_W = \varphi_W = \varphi_V \circ \iota_W$. It follows that

$$y = \varphi_U(x) = \varphi_U \circ \iota_W(x) = \varphi_V \circ \iota_W(x) = \varphi_V(x) = y'$$

To see that φ is a coalgebra-to-algebra homomorphism, let $x \in U \subseteq X$ and calculate

$$\varphi(x) = \varphi_U(x)$$
$$= \gamma \circ H(\varphi_U) \circ \beta_U(x)$$
$$= \gamma \circ H(\varphi) \circ H(\iota_U) \circ \beta_U(x)$$
$$= \gamma \circ H(\varphi) \circ \beta \circ \iota_U(x)$$
$$= \gamma \circ H(\varphi) \circ \beta(x) \qquad \square$$

An extended example: regular subfractals via finite corecursive algebras. In this next part of the paper, recall that for a complete metric space M and a finite set X, the set M^X obtains a complete metric space structure from the product metric, $d^X(f,g) = \max_{x \in X} d(f(x), g(x))$. Recall also that $\mathbf{K}(M)$ is the space of non-empty compact subsets of M using the Hausdorff metric in (3.6).

Theorem 5.10. *Let M be a complete metric space, and let $\sigma \colon A \to \mathrm{Con}(M)$. For the* Set *endofunctor $FX = \mathcal{P}_*(A \times X)$, the algebra $k \colon \mathcal{P}_*(A \times \mathbf{K}(M)) \to \mathbf{K}(M)$ given by*

$$k(C) = \bigcup\{\sigma_a(K) \mid (a, K) \in C\}. \tag{5.4}$$

is finitely corecursive.[7] *Moreover, if an F-coalgebra (X, β) is locally finite, its solution in the algebra $\mathbf{K}(M)$ is $[\![-]\!]_{\beta,\sigma} \colon X \to \mathbf{K}(M)$.*

We saw in Example 5.5 that an F-coalgebra structure is the same data as a productive and finitely-branching LTS. For this reason, we generally identify the two notions and refer to an F-coalgebra as an LTS and vice versa.

Proof. To see the first statement, by Lemma 5.9, it suffices to show that finite F-coalgebras admit unique coalgebra-to-algebra homomorphisms into $\mathbf{K}(M)$. Let (X, β) be a finite LTS. Recall from Theorem 3.13 that finite productive LTSs have unique solutions in the sense of Definition 3.10, i.e., there is a unique map $\varphi \colon X \to \mathbf{K}(M)$ such that

$$\varphi(x) = \bigcup_{x \xrightarrow{a} y} \sigma_a(\varphi(y))$$

for any $x \in X$ (this is the equation from Definition 3.10). Chasing the coalgebra-to-algebra homomorphism square around its underside, one finds that an arbitrary map $\varphi \colon X \to \mathbf{K}(M)$ is a coalgebra-to-algebra homomorphism if and only if for all $x \in X$, the equation on the right below holds:

$$\varphi(x) = k \circ \mathcal{P}_*(1_A \times \varphi) \circ \beta(x)$$
$$= k(\{(a, \varphi(y)) \mid x \xrightarrow{a} y\})$$
$$= \bigcup_{x \xrightarrow{a} y} \sigma_a(\varphi(y))$$

So coalgebra-to-algebra maps are the same thing as solutions in the sense of Definition 3.10. Hence, by Theorem 3.13, every finite F-coalgebra does indeed admit a unique coalgebra-to-algebra homomorphism into $(\mathbf{K}(M), k)$. Once again, we use Lemma 5.9 to conclude that $(\mathbf{K}(M), k)$ is finitely corecursive.

Moreover, by our characterization of coalgebra-to-algebra maps as solutions to the corresponding productive LTS (X, β) above, by Theorem 3.13(4), $[\![-]\!]_{\alpha,\sigma} : X \to \mathbf{K}(M)$ is the unique coalgebra-to-algebra homomorphism. $\hspace{1cm}\square$

Theorem 5.10 tells us that there is a direct connection between the regular subfractals of Section 3 and finitely corecursive algebras. In the next section, we turn to the probabilistic analog of labelled transition systems and their measure-theoretic semantics, where we will see another example of a finitely corecursive algebra.

Labelled Markov Chains. Labelled Markov chains generalize our specification of invariant measures to multi-state transition systems with branching given by finitely supported probability distributions.

Definition 5.11. A *labelled Markov chain* (LMC) is a coalgebra for the functor $\mathcal{D}(A \times -)$. That is, it is a pair (X, β) consisting of a set X of *states* and a *transition function* $\beta \colon X \to \mathcal{D}(A \times X)$. We write $x \xrightarrow{r|a}_{\beta} y$ if $\beta(x)(a, y) = r$, often dropping the symbol β if it is clear form context. A *homomorphism* of LMCs $h \colon (X, \beta_X) \to (Y, \beta_Y)$ is thus a map $h \colon X \to Y$ such that $\mathcal{D}(1_A \times h) \circ \beta_X = \beta_Y \circ h$ where 1_A is the identity map on A.

As we have already seen, given a complete metric space M and a contraction operator interpretation $\sigma \colon A \to \mathrm{Con}(M)$, every state x of a productive LTS (X, α) with labels in A corresponds to a regular subfractal $[\![x]\!]_{\alpha,\sigma}$ of the set $\mathbf{S}_\sigma$ from (3.2). This regular subfractal is the continuous image of the set $\mathrm{str}_\alpha(x)$ under the map $\sigma_\omega \colon (A^\omega, d_\sigma) \to M$. The set $[\![x]\!]_{\alpha,\sigma}$ is characterized by its satisfaction of the equations representing the LTS (X, α).

Every LMC (X, β) has an *underlying* LTS $(X, \bar{\beta})$, where $\bar{\beta}(x) = \{(a, y) \mid \beta(x)(a, y) > 0\}$. Note that since $\beta(x)$ is a probability distribution, our underlying LTS is automatically productive since at least one transition must have positive probability. For each $x \in X$, we are going to define a probability measure $\hat{\beta}_\sigma(x)$ on $\mathbf{S}_\sigma$ whose support is $[\![x]\!]_{\bar{\beta},\sigma}$, and that satisfies a recursive system of equations represented by the LMC (X, β). Roughly, $\hat{\beta}_\sigma(x)$ is the pushforward of a certain Borel probability measure $\hat{\beta}(x)$ on A^ω.

Definition 5.12. Suppose we are given an LMC (X, β) and a state $x \in X$. We follow Kerstan and König [KK13] and define the *trace measure* $\hat{\beta}(x) \in \mathsf{Prob}(A^\omega)$ by giving its value on a basic open set B_w. Given $w = a_1 \cdots a_n \in A^*$, we define

$$\hat{\beta}(x)(B_w) = \sum \{r_1 \cdots r_n \mid x \xrightarrow{r_1|a_1} x_1 \longrightarrow \cdots \xrightarrow{r_n|a_n} x_n \text{ for some } x_1, x_2, \ldots, x_n\} \quad (5.5)$$

In particular, $\hat{\beta}(x)(B_\epsilon) = \hat{\beta}(x)(A^\omega) = 1$.

We use the following convenient proposition, stated as Theorem 3.5 in Goy and Rot [GR18], which tells us that (5.5) defines a unique Borel probability measure on A^ω.

Proposition 5.13. *Let $j \colon A^* \to [0, 1]$ satisfy $j(w) = \sum_{a \in A} j(wa)$ for any $w \in A^*$ and $j(\epsilon) = 1$, where ϵ is the empty word. Then there is a unique compactly supported Borel probability measure $\rho \in \mathsf{Prob}(A^\omega)$ such that for any $w \in A^*$, $\rho(B_w) = j(w)$.*

A simple calculation using (5.5) shows that $\hat{\beta}(x)(B_w) = \sum_{a \in A} \hat{\beta}(x)(B_{wa})$ for any LMC (X, β). So, if we take $j(w) = \hat{\beta}(x)(B_w)$, we immediately have $j(\epsilon) = 1$ and $j(w) = \sum_{a \in A} j(wa)$ like in Proposition 5.13. Applying the proposition, there is a unique Borel probability measure $\hat{\beta}(x)$ on A^ω such that (5.5) holds for any basic open set B_w.

Definition 5.14. Let (X, β) be an LMC, and $\sigma \colon A \to \mathrm{Con}(M)$ be a contraction operator interpretation in a complete metric space. For each $x \in X$, the *regular subfractal measure* corresponding to x is $\hat{\beta}_\sigma(x) = \sigma_\omega^\# \hat{\beta}(x)$, i.e., $\hat{\beta}_\sigma \colon X \to \mathsf{Prob}(M)$ is defined by $\hat{\beta}_\sigma = \sigma_\omega^\# \circ \hat{\beta}$.

Systems of Probabilistic Equations. The next theorem is the main result in this section. Like Theorem 5.10, it gives a finitely corecursive algebra which may be used to derive some of the basic results on fractals. We would like to prove the two the same way, by using an argument which calls on the Banach fixed point theorem. However, as we noted above, for a complete space M, $\mathsf{Prob}(M)$ is not complete in general, so we need a different way.

Theorem 5.15. *Let M be a complete metric space, and let $\sigma \colon A \to \mathrm{Con}(M)$. The algebra $p : \mathcal{D}(A \times \mathsf{Prob}(M)) \to \mathsf{Prob}(M)$ given by*

$$p(r_1(a_1, \theta_1) + \cdots + r_n(a_n, \theta_n))(B) = \sum_{i=1}^{n} r_i \sigma_{a_i}^\# \theta_i(B) \tag{5.6}$$

is a finitely corecursive algebra for the Set endofunctor $FX = \mathcal{D}(A \times X)$. More specifically, given a locally finite labelled Markov chain (X, β), its solution $\beta^\dagger$ is $\hat{\beta}_\sigma \colon X \to \mathsf{Prob}(M)$ (see Definition 5.14). Also, the support of $\hat{\beta}_\sigma(x)$ is $[\![x]\!]_{\bar{\beta},\sigma}$.

Proof. Again, we appeal to Lemma 5.9. Fix a finite coalgebra (X, β). We aim to show that there is a unique coalgebra-to-algebra homomorphism $(X, \beta) \to (\mathsf{Prob}(M), p)$. Let us first characterize coalgebra-to-algebra maps among all maps $\varphi \colon X \to \mathsf{Prob}(M)$ in the following way: φ is a coalgebra-to-algebra map (i.e., the diagram on the left below commutes) if and only if for any $x \in X$ and any Borel set $B \subseteq M$, the equations on the right below hold:

$$\varphi(x)(B) = p \circ \mathcal{D}(1_A \times \varphi) \circ \beta(x)(B)$$

$$X \xrightarrow{\varphi} \mathsf{Prob}(M) \qquad\qquad = p\Big(\sum_{x \xrightarrow{r|a} y} r(a, \varphi(y))\Big)(B)$$

$$\beta \downarrow \qquad\qquad \uparrow p$$

$$\mathcal{D}(A \times X) \xrightarrow{\mathcal{D}(1_A \times \varphi)} \mathcal{D}(A \times \mathsf{Prob}(M)) \qquad = \sum_{x \xrightarrow{r|a} y} r \sigma_a^\# (\varphi(y))(B)$$

We are going to show that $\hat{\beta}_\sigma \colon X \to \mathsf{Prob}(M)$ is the unique map with this property. This can be done in two steps.

(i) First, we show that the map $F_\beta \colon \mathsf{Prob}(M)^X \to \mathsf{Prob}(M)^X$ defined by

$$F_\beta(h)(x)(B) = p \circ \mathcal{D}(1_A \times h) \circ \beta(x)(B) = \sum_{x \xrightarrow{r|a} y} r \sigma_a^\# (h(y))(B)$$

is a contraction mapping on the metric space $(\mathsf{Prob}(M)^X, d^X)$, where the metric is the product metric, $d^X(f, g) = \max\{d(f(y), g(y)) \mid y \in X\}$, and

(ii) that $\hat{\beta}_\sigma$ is a fixed-point of this mapping.

For (i), note that indeed, $F_\beta(h)(x)$ is a compactly supported Borel probability measure, because it is a finite convex combination of pushforwards of compactly supported Borel probability measures. For the rest our proof of (i), we reuse the proof technique of [Hut81, Theorem 4.4(1)]. For each $a \in A$, let c_a be a nonzero contraction coefficient for σ_a. Let $c = \max\{c_a \mid a \in A\}$, so that $0 < c < 1$ and for any $a \in A$, $c_a/c \leq 1$. For any nonexpanding $f \colon M \to \mathbb{R}$ and any $y \in X$,

$$
\int_M f \ \mathsf{d}(\sigma_a^\# h(y)) = c \int_M \frac{1}{c} f \ \mathsf{d}(\sigma_a^\# h(y))
$$

$$
= c \int_{\sigma_a^{-1}(M)} \frac{1}{c} f \circ \sigma_a \ \mathsf{d}h(y)
$$

$$
= c \int_M \frac{1}{c} f \circ \sigma_a \ \mathsf{d}h(y)
$$

The second equality uses the change-of-variables formula, which in its general form states that $\int_U f(x) \ \mathsf{d}(g^\# \mu) = \int_{g^{-1}(U)} f \circ g \ \mathsf{d}\mu$ [BR07, Theorem 3.6.1]. The last equality holds since $\sigma_a^{-1}(M) = M$. Since $0 < c_a/c \leq 1$, $\frac{1}{c} f \circ \sigma_a$ is a nonexpanding map into $\mathbb{R}$. Taking d to be the Kantorovich-Rubenstein metric, this implies that for any $\rho_1, \rho_2 \in \mathsf{Prob}(M)$,

$$
\int_M \frac{1}{c} f \circ \sigma_a \ \mathsf{d}\rho_1 - \int_M \frac{1}{c} f \circ \sigma_a \ \mathsf{d}\rho_2 \leq d(\rho_1, \rho_2) \tag{5.7}
$$

Putting these observations together, we obtain the following for any $g, h \colon X \to \mathsf{Prob}(M)$. If $f \colon M \to \mathbb{R}$ is a nonexpanding map, then

$$
\int_M f \ \mathsf{d}F_\beta(g)(x) - \int_M f \ \mathsf{d}F_\beta(h)(x)
$$

$$
= \int_M \sum_{x \xrightarrow{r|a} y} rf \ \mathsf{d}(\sigma_a^\# g(y)) - \int_M \sum_{x \xrightarrow{r|a} y} rf \ \mathsf{d}(\sigma_a^\# h(y)) \qquad \text{(definition of } F)
$$

$$
= \sum_{x \xrightarrow{r|a} y} r \Big(\int_M f \ \mathsf{d}(\sigma_a^\# g(y)) - \int_M f \ \mathsf{d}(\sigma_a^\# h(y)) \Big) \qquad \text{(linearity of } \int)
$$

$$
= \sum_{x \xrightarrow{r|a} y} rc \Big(\int_M \frac{1}{c} f \circ \sigma_a \ \mathsf{d}g(y) - \int_M \frac{1}{c} f \circ \sigma_a \ \mathsf{d}h(y) \Big) \qquad \text{(linearity of } \int)
$$

$$
\leq \sum_{x \xrightarrow{r|a} y} rc \ d(g(y), h(y)) \tag{5.7}
$$

$$
\leq c \ \max\{d(g(y), h(y)) \mid y \in X \text{ and } x \xrightarrow{r|a} y\} \qquad (\textstyle\sum r = 1)
$$

$$
\leq c \ d^X(g, h) \qquad \text{(def. of } d^X)
$$

Since $d^X(F_\beta(g), F_\beta(h)) = \max\{d(F_\beta(g)(x), F_\beta(h)(x)) \mid x \in X\}$ and $d(F_\beta(g)(x), F_\beta(h)(x))$ is the supremum over nonexpanding $f \colon M \to \mathbb{R}$ in the first expression in the calculation above, this shows that F_β is a contraction mapping in the product metric on $\mathsf{Prob}(M)^X$.

To see item (ii), we begin by showing that for any $x \in X$ and Borel set $B \subseteq A^\omega$,

$$\hat{\beta}(x)(B) = \sum_{x \xrightarrow{r|a} y} r\sigma_a^\#(\beta(y))(B) \qquad (5.8)$$

By the Identity and Extension Theorems for σ-finite premeasures [KK13], it suffices to show (5.8) for each basic open set $B = B_w$, with $w \in A^*$. We proceed by induction on the length of w. We begin with the base case $w = \epsilon$. Let us use $a(-)$ again for the prefixing map $a(a_1, a_2, \dots) = (a, a_1, a_2, \dots)$ on A^ω. Since $a^{-1}(A^\omega) = A^\omega$ and $\sum \{r \mid x \xrightarrow{r|a} y\} = 1$,

$$\hat{\beta}(x)(B_\epsilon) = \hat{\beta}(x)(A^\omega) = 1 = \sum_{x \xrightarrow{r|a} y} r\hat{\beta}(y)(a^{-1}(A^\omega)) = \sum_{x \xrightarrow{r|a} y} ra^\# \hat{\beta}(y)(B_\epsilon)$$

Assume our result for $w = a_1 \cdots a_n$, and let $a \in A$, and consider $aw = aa_1 \cdots a_n$. Then

$$\begin{aligned}
\hat{\beta}(x)(B_{aw}) &= \sum \{rr_1 \cdots r_n \mid x \xrightarrow{r|a} y \xrightarrow{r_1|a_1} y_1 \longrightarrow \cdots \longrightarrow x_n\} \\
&= \sum_{x \xrightarrow{r|a} y} r \sum \{r_1 \cdots r_n \mid y \xrightarrow{r_1|a_1} y_1 \longrightarrow \cdots \longrightarrow x_n\} \\
&= \sum_{x \xrightarrow{r|a} y} r\hat{\beta}(y)(B_w) &\text{(induction hypothesis)} \\
&= \sum_{x \xrightarrow{r|a} y} r\hat{\beta}(y)(a^{-1}(B_{aw})) &(a^{-1}(B_{aw}) = B_w) \\
&= \sum_{x \xrightarrow{r|a} y} ra^\#(\hat{\beta}(y))(B_{aw})
\end{aligned}$$

This shows (5.8). Finally, we will use (5.8) to prove (ii). First consider a measure $\rho \in \mathsf{Prob}(A^\omega)$ and a Borel set $B \subseteq A^\omega$. Since $\sigma_a \circ \sigma_\omega = \sigma_\omega \circ a(-)$ (this is (3.4)), we have

$$\begin{aligned}
\sigma_a^\# \sigma_\omega^\# \rho(B) &= \rho(\sigma_\omega^{-1} \circ \sigma_a^{-1}(B)) \\
&= \rho((\sigma_a \circ \sigma_\omega)^{-1}(B)) \\
&= \rho((\sigma_\omega \circ a(-))^{-1}(B)) \\
&= \rho(a(-)^{-1} \circ \sigma_\omega^{-1}(B)) \\
&= a^\# \rho(\sigma_\omega^{-1}(B))
\end{aligned}$$

For any $x \in X$ and Borel set $B \subseteq M$, if we take $\rho = \hat{\beta}(y)$ below, then from the above calculation we obtain

$$\begin{aligned}
F_\beta(\hat{\beta}_\sigma)(x)(B) &= \sum_{x \xrightarrow{r|a} y} r\sigma_a^\# \sigma_\omega^\# \hat{\beta}(y)(B) &\text{(def. of } F_\beta, \hat{\beta}_\sigma) \\
&= \sum_{x \xrightarrow{r|a} y} ra^\# \hat{\beta}(y)(\sigma_\omega^{-1}(B)) \\
&= \hat{\beta}(x)(\sigma_\omega^{-1}(B)) &(5.8)
\end{aligned}$$

$$= \sigma_\omega^\# \hat{\beta}(x)(B)$$
$$= \hat{\beta}_\sigma(x)(B) \qquad\qquad\qquad\qquad (\text{def. of } \hat{\beta}_\sigma)$$

At this point, we return to (i) and
refitem-f-fixed-point from near the start of this proof. We have verified that $\hat{\beta}_\sigma$ is a fixed point of F_β, and we also know that F_β is contractive and consequently has at most one fixed point. So, $\hat{\beta}_\sigma$ is the unique fixed point. It is therefore the unique coalgebra-to-algebra homomorphism.

For the last statement, let us first verify that the support of $\hat{\beta}(x)$ is precisely $\mathrm{str}_{\bar{\beta}}(x)$. To see that $\mathrm{str}_{\bar{\beta}}(x) \subseteq \mathrm{supp}(\hat{\beta}(x))$, let $(a_i)_{i\in\mathbb{N}} \in \mathrm{str}_{\bar{\beta}}(x)$. Then there is a path $x \xrightarrow{r_1|a_1} x_1 \xrightarrow{r_2|a_2} \cdots$ in (X,β). It suffices to see that for any $w \in A^*$ such that $(a_i)_{i\in\mathbb{N}} \in B_w$, $\hat{\beta}(x)(B_w) > 0$. But, if $(a_i)_{i\in\mathbb{N}} \in B_w$, then $a_1 \cdots a_n = w$, so $\hat{\beta}(x)(B_w) = \prod_{i=1}^n r_i > 0$. Hence, $(a_i)_{i\in\mathbb{N}} \in \mathrm{supp}(\hat{\beta}(x))$. This shows $\mathrm{str}_{\bar{\beta}}(x) \subseteq \mathrm{supp}(\hat{\beta}(x))$.

For the reverse inclusion, given $(a_i)_{i\in\mathbb{N}} \in \mathrm{supp}(\hat{\beta}(x))$, we know that every B_w containing $(a_i)_{i\in\mathbb{N}}$ has positive trace measure. Remember that $(a_i)_{i\in\mathbb{N}} \in B_w$ if and only if $a_1 \cdots a_n = w$. But $\hat{\beta}(x)(B_w) > 0$ if and only if there is a path $x \xrightarrow{r_1|a_1} x_1 \xrightarrow{r_2|a_2} \cdots$ in (X,β) with $\hat{\beta}(x)(B_w) = \prod_{i=1}^n r_i$. Since w was arbitrary, we have just seen that every finite initial segment of $(a_i)_{i\in\mathbb{N}}$ corresponds to a path in (X,β). By Lemma 2.3, $(a_i)_{i\in\mathbb{N}} \in \mathrm{str}_{\bar{\beta}}(x)$. Thus, $\mathrm{str}_{\bar{\beta}}(x) = \mathrm{supp}(\hat{\beta}(x))$.

Now let us use the previous paragraph to prove the last statement of the theorem. We use the following well-known fact from measure theory, that for any Borel measure ρ and continuous map f between metric spaces, $\mathrm{supp}(f^\#\rho) = \overline{f(\mathrm{supp}(\rho))}$, where $\overline{(-)}$ denotes topological closure. In our case, we have $f = \sigma_\omega$ and $\rho = \hat{\beta}$, which gives

$$\mathrm{supp}(\hat{\beta}_\sigma(x)) = \mathrm{supp}(\sigma_\omega^\# \hat{\beta}(x))$$
$$= \overline{\sigma_\omega(\mathrm{supp}(\hat{\beta})(x))} \qquad\qquad (\text{fact above})$$
$$= \overline{\sigma_\omega(\mathrm{str}_{\bar{\beta}}(x))} \qquad\qquad (\mathrm{str}_{\bar{\beta}}(x) = \mathrm{supp}(\hat{\beta}(x)))$$
$$= \overline{[\![x]\!]_{\bar{\beta},\sigma}} \qquad\qquad (\text{def. of } [\![-]\!]_{\bar{\beta},\sigma})$$
$$= [\![x]\!]_{\bar{\beta},\sigma} \qquad\qquad ([\![x]\!]_{\bar{\beta},\sigma} \text{ is closed by Theorem 3.13})$$

$\square$

An aside: chaos games for regular subfractals and fractal measures. Every probability distribution θ on our set A of actions determines a probability measure $\hat{\theta}$ on A^ω. Now suppose we are also given a contraction operator interpretation $\sigma\colon A \to \mathrm{Con}(M)$ and a fixed $p \in M$. Each infinite sequence $(a_1, a_2, \ldots)$ gives us an infinite sequence of points $p_n \in M$ defined recursively by setting $p_0 = p$, and $p_{n+1} = \sigma_{a_n}(p_n)$. We are interested in the case that the starting point p belongs to the self-similar set $\mathbf{S}_\sigma$ determined by σ (see Equation 3.2). In this case, each point p_n also belongs to $\mathbf{S}_\sigma$. So we obtain a map $\hat{p}\colon A^\omega \to \mathbf{S}_\sigma^\omega$. This map $\hat{p}$ turns out to be measurable.

Before we mention some specific events which are of interest, let us recast the situation under discussion in more dynamic terms. Again given the initial point p, we imagine generating the points p_n by a random process. The point $p_0 = p$ is given, and from p_n we determine p_{n+1} by randomly choosing a from A with probability $\theta(a)$ and then setting

$p_{n+1} = \sigma_a(p_n)$. All of our choices are independent, and the past has no effect on the present. This is called the *random iteration algorithm* (or *chaos game*), and it is a standard topic in the fractals literature (see e.g. [Bar88, Chapter 9]).

In essence, the chaos game selects a stream $(a_1, a_2, \dots) \in A^\omega$ by repeatedly sampling θ and considers its image $\hat{p}(a_1, a_2, \dots)$. As we have seen, the probability that the stream which it selects lies in a given Borel set $B \subseteq A^\omega$ is $\hat{\theta}(B)$. For the purposes of the next few paragraphs, we will call a Borel subset of A^ω an *event* and refer to the *probability of an event* B when we mean $\hat{\theta}(B)$.

Among the notable results about the chaos game are the following. Firstly, if $\theta(a) > 0$ for all $a \in A$, then for any point $p \in \mathbf{S}_\sigma$, the event in which its corresponding run of the chaos game $\{p_0, p_1, p_2, \dots\}$ is dense in $\mathbf{S}_\sigma$ has probability 1 [Fal86]. Additionally, for a given Borel set $U \subseteq M$ and starting point $p \in M$ (not necessarily in $\mathbf{S}_\sigma$), the event in which the following equation holds also has probability 1 with respect to its corresponding run of the chaos game [DS84, pg. 7] (also [Elt87]).

$$\hat{\theta}_\sigma(U) = \lim_{n \to \infty} \frac{1}{n} |U \cap \{p_1, \cdots, p_n\}| \qquad (5.9)$$

In words, the value of $\hat{\theta}_\sigma(U)$ is the limit of the proportion of the first n points in the infinite sequence $(p_i)_{i \in \mathbb{N}}$ that happen to belong to U.

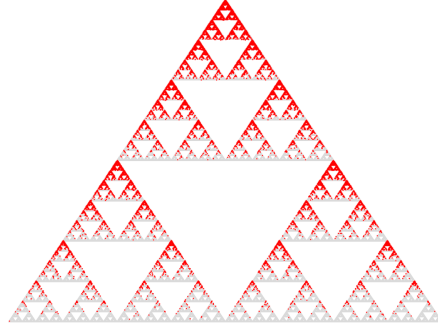

Figure 7: One run of the chaos game on the Sierpinski gasket from Figure 1 with probability distribution given by $\theta(a) = 0.8$, $\theta(b) = 0.1$, and $\theta(c) = 0.1$. Above, the points generated by the chaos game are drawn in red. In this run, of the 200,000 points generated, 160,081 points appeared in the top third of the gasket, i.e., ≈ 0.8 of the total.

Example 5.16. Let A and σ be as in Figure 1. Consider the distribution $\theta\colon A \to [0, 1]$ given by $\theta(a) = 0.8$, $\theta(b) = 0.1$, and $\theta(c) = 0.1$, and let $\hat{\theta}_\sigma$ be the invariant measure specified by θ. Running the chaos game with starting point $p = (0, 0)$ we obtained the graph in Figure 7. Since 80% of the distribution is on a, approximately 80% of the points generated by the chaos game should be in the top third of the Sierpiński gasket.

The right-hand side of (5.9) was obtained by generating a sequence of points using a probability distribution on A and an initial point p of the underlying complete metric space. For regular submeasures, which are generated by a productive labelled Markov chain instead of a probability distribution on A, experiments indicate that a similar game can be played.

Let $\sigma \colon A \to \mathrm{Con}(M)$ be a contraction operator interpretation and let (X, β) be a productive labelled Markov chain. Fix $x \in X$ and let $K = [\![x]\!]_{\beta, \sigma}$. We define the *multistate chaos game* for x, σ, and β with initial point $p \in [\![x]\!]_{\beta, \sigma}$ to be the following stochastic procedure: generate a path $x = x_1 \xrightarrow{r_1 | a_1} x_2 \xrightarrow{r_2 | a_2} \cdots$ by sampling the distribution $\beta(x_i)$ at each step. Now set $q_1 = p$ and $q_{i+1} = \sigma_{a_i}(q_i)$ for each $i \in \mathbb{N}$. This defines a sequence of points $(q_i)_{i \in \mathbb{N}}$ in M, usually not all of which are in $[\![x]\!]_{\beta, \sigma}$. Notice that $q_i \in [\![x_i]\!]_\beta$ for each $i \in \mathbb{N}$. The outcome of the multistate chaos game is the subsequence $(p_{i_k})_{k \in \mathbb{N}}$ of $(q_i)_{i \in \mathbb{N}}$, i.e., with $i_1 < i_2 < i_3 < \cdots$, consisting of all the points p_{i_k} such that at index i_k the state x_{i_k} is $x_{i_k} = x$.

Experimental results so far indicate that the event in which the set $\{p_{i_k}\}_{k \in \mathbb{N}}$ is dense in the regular subfractal $[\![x]\!]_{\beta, \sigma}$ has probability 1. Additionally, we suspect that an equation analogous to that of (5.9) also holds for multistate chaos games: for a Borel set $U \subseteq M$,

$$\hat{\beta}_\sigma(x)(U) = \lim_{n \to \infty} \frac{1}{n} |U \cap \{p_{i_1}, p_{i_2}, \cdots\}| \tag{5.10}$$

where β and $(p_{i_k})_{k \in \mathbb{N}}$ were defined above. To the authors' knowledge, (5.10) has not yet been verified or disproven in the literature. We will pursue this issue in further work.

Example 5.17. Consider the labelled Markov chain (X, β) displayed in Figure 8. One run of the multistate chaos game starting at state x and the point $p = (0,0)$ generated the sequence of points in the plot on the right in the same figure. The plot suggests that the bottom-left third, i.e., $\sigma_\omega(B_b)$, carries more weight than the rest of the plot: a raw count of the points in the bottom-left third of the regular subfractal reveals that 100,120 out of the 200,000 appear in that third. Thus, the proportion of the points generated that land in $\sigma_\omega(B_b)$ is very close to the actual value, $\hat{\beta}_\sigma(\sigma_\omega(B_b)) = \hat{\beta}(B_b) = 0.5$. Further experiments have given similar results.
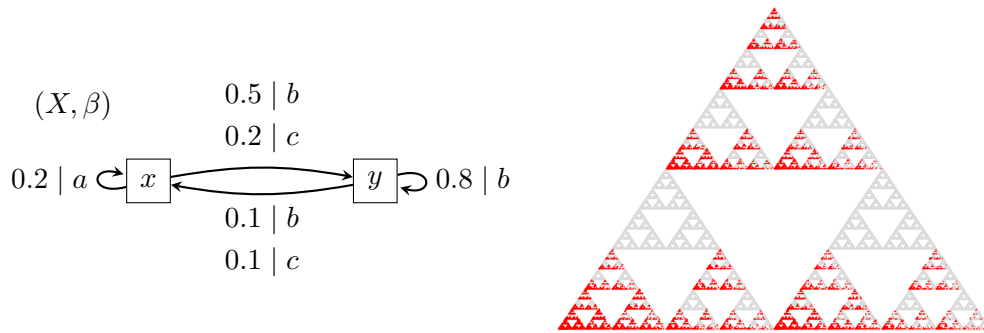

Figure 8: The sequence of points generated by 200,000 iterations of the multistate chaos game for (X, β), starting from the state x and the point p. Above, the points generated by the chaos game are drawn in red. In this run, 100,120 out of the 200,000 points generated appear in the bottom-left third of the regular subfractal, i.e., ≈ 0.5 of the total.

6. Probabilistic Process Terms

Next, we introduce a syntax for specifying LMCs. Our specification language is essentially the *productive* fragment of Stark and Smolka's process calculus [SS00], meaning that the μ-expressions do not involve deadlock and all variables are guarded (see Definition 2.5).

Definition 6.1. The set of *probabilistic μ-expressions* is given by the grammar

$$v \mid ae \mid e_1 \oplus_r e_2 \mid \mu v\ e$$

Here $r \in (0,1)$, and otherwise we make the same stipulations as in Definition 2.5. The set of *probabilistic process terms* PTerm consists of the closed and guarded probabilistic μ-expressions.

Instead of *languages* of streams, the analog of trace semantics appropriate for probabilistic process terms is a measure-theoretic semantics consisting of trace measures introduced in the previous section (Definition 5.12).

We turn next to the semantics. For use there, and in other definitions, we need a preliminary definition.

Definition 6.2. The *unravelling number of e*, denoted $\#_{un}(e)$, is defined by the following recursion on guarded (not necessarily closed) terms: for any $a \in A$, e, e_1, e_2,

$$\#_{un}(ae) = 0 \qquad \#_{un}(e_1 \oplus_r e_2) = \max_{i \in \{1,2\}} \{\#_{un}(e_i)\} \qquad \#_{un}(\mu v\ e) = 1 + \#_{un}(e)$$

Suppose that e were displayed as a syntax tree. Then $\#_{un}(e)$ would be the maximum number of μx nodes, with the maximum taken along all paths from the root e to some internal node labelled with an action prefix a.

For example, $\#_{un}(\mu v\ (av \oplus_{\frac{1}{2}} \mu w\ (bw))) = 1 + \max\{\#_{un}(av), \#_{un}(\mu w\ (bw))\} = 2$.

Proposition 6.3. *For all guarded terms e,*
(1) *whenever $e[f/v]$ is defined, $\#_{un}(e) = \#_{un}(e[f/v])$.*
(2) *Assume that $\mu v\ e$ is guarded and closed. Then $\#_{un}(\mu v\ e) = 1 + \#_{un}(e[\mu v\ e/v])$.*

Proof. The first part is proven by induction on e. The base case for variables is trivial, since variables are not guarded. The only induction step worth mentioning is that of a guarded term $\mu w\ e$, with w a variable different from v, and such that $(\mu w\ e)[f/v]$ is defined. In this case, e also is guarded, and the induction hypothesis applies to it. Also w has no free occurrences in f, and therefore $e[f/v]$ is well defined. We have

$$\begin{aligned}
\#_{un}(\mu w\ e) &= 1 + \#_{un}(e) \\
&= 1 + \#_{un}(e[f/v]) &\text{(induction hypothesis)} \\
&= \#_{un}(\mu w\ (e[f/v])) \\
&= \#_{un}((\mu w\ e)[f/v])
\end{aligned}$$

In the second part, assume that $\mu v\ e$ is guarded and closed. Then e is guarded (but not necessarily closed). We use the first part, taking f to be $\mu v\ e$. Now $e[\mu v\ e/v]$ is well defined, since v is not free in $\mu v\ e$. And $\#_{un}(\mu v\ e) = 1 + \#_{un}(e) = 1 + \#_{un}(e[\mu v\ e/v])$. $\qquad\square$

The important point in Proposition 6.3 is the last one, since it enables us to define functions t on PTerm by recursion on $\prec$ in such a way that $t(\mu v\ e)$ only depends on $t(e[\mu v\ e/v])$.

$$\delta(ae)(b,f) = \begin{cases} 1 & f = e \text{ and } b = a \\ 0 & \text{otherwise} \end{cases}$$

$$\delta(e_1 \oplus_r e_2)(b,f) = r\delta(e_1)(b,f) + (1-r)\delta(e_2)(b,f)$$

$$\delta(\mu v\ e)(b,f) = \delta(e[\mu v\ e/v])(b,f)$$

Figure 9: The LMC structure (PTerm, δ). Above, $a, b \in A$, $r \in (0,1)$, and $e, e_i, f \in \mathsf{PTerm}$.

Definition 6.4. For each term e, we write $\#_\oplus(e)$ for the number of $\oplus_r$-operators in e. We define a function $\mathsf{rht} : \mathsf{PTerm} \to \omega \times \omega$ by[8]

$$\mathsf{rht}(e) = (\#_{un}(e), \#_\oplus(e))$$

Let $\prec$ be the strict lexicographic order on $\omega \times \omega$, and recall that this is a strict well-order.

Here is the point of these definitions. For $i = 1, 2$, $\mathsf{rht}(e_i) \prec \mathsf{rht}(e_1 \oplus_r e_2)$. And $\mathsf{rht}(e[\mu v\ e/v]) \prec \mathsf{rht}(\mu v\ e)$.

Definition 6.5. We define the LMC (PTerm, δ) in Figure 9 and call it the *syntactic LMC*. The *trace measure semantics* $\mathsf{trm}(e)$ of a probabilistic process term e is defined to be $\mathsf{trm}(e) = \hat{\delta}(e)$ (see Definition 5.12).

Two probabilistic process terms e and f are *trace measure equivalent* if they have the same trace measure semantics. Given $\sigma \colon A \to \mathrm{Con}(M)$, the *subfractal measure semantics of* $e \in \mathsf{PTerm}$ *corresponding to* σ is $\hat{\delta}_\sigma(e)$ (see Definition 5.14).

The next result gives a recursive method for computing the trace measure semantics of a process term. This recursive method will be useful in Section 7.

Lemma 6.6. *For any $w \in A^*$, $a \in A$, $e, e_i \in \mathsf{PTerm}$, and $r \in (0,1)$, $\mathsf{trm}(e)(A^\omega) = 1$ and*

$$\mathsf{trm}(ae)(B_w) = \begin{cases} \mathsf{trm}(e)(B_u) & w = au \\ 0 & \text{otherwise} \end{cases}$$

$$\mathsf{trm}(e_1 \oplus_r e_2)(B_w) = r\,\mathsf{trm}(e_1)(B_w) + (1-r)\,\mathsf{trm}(e_2)(B_w)$$

$$\mathsf{trm}(\mu v\ e)(B_w) = \mathsf{trm}(e[\mu v\ e/v])(B_w)$$

Proof. For each probabilistic process term e, define $j(e) \colon A^* \to [0,1]$ inductively by setting $j(e)(\epsilon) = 1$ for all $e \in \mathsf{PTerm}$, and for $a \in A$, $r \in [0,1]$, $e, e_1, e_2 \in \mathsf{PTerm}$,

$$j(ae)(w) = \begin{cases} j(e)(u) & w = au \\ 0 & \text{otherwise} \end{cases} \qquad \begin{aligned} j(e_1 \oplus_r e_2)(w) &= rj(e_1)(w) + (1-r)j(e_2)(w) \\ j(\mu v\ e)(w) &= j(e[\mu v\ e/v])(w) \end{aligned}$$

We are going to show that $j(e)(w) = \mathsf{trm}(e)(B_w)$ for every $e \in \mathsf{PTerm}$ by induction on w, and then by a subinduction on $\mathsf{rht}(e)$ (see Definition 6.4). Since $\mathsf{trm}(e)$ already satisfies $\mathsf{trm}(e)(B_w) = \sum_{a \in A} \mathsf{trm}(e)(B_{wa})$ and $\mathsf{trm}(e)(B_\epsilon) = \mathsf{trm}(e)(A^\omega) = 1$, we get $j(e)(w) = \sum_{a \in A} j(e)(wa)$ and $j(e)(\epsilon) = 1$, allowing us to apply the uniqueness part of Proposition 5.13.

Here is the base case of induction on w: For any $e \in \mathsf{PTerm}$, $j(e)(\epsilon) = 1$ and $\mathsf{trm}(e)(B_\epsilon) = \mathsf{trm}(e)(A^\omega) = 1$.

---

[8]The notation rht stands for "r-height".

In the induction step, we consider the word $w = bu$. We assume that $j(e)(u) = \mathrm{trm}(e)(B_u)$ for every $e \in \mathsf{PTerm}$ and we show the same thing for bu. We argue by induction on $\mathsf{rht}(e)$.

- Our first case is for $\mathsf{rht}(e) = (0,0)$, i.e., for a process term ae. If $b = a$, we recall that $w = bu$ and then calculate:

$$
\begin{aligned}
j(ae)(bu) &= j(e)(u) \\
&= \mathrm{trm}(e)(u) && \text{(top level induction, on } w) \\
&= 1\,\mathrm{trm}(e)(u) \\
&= \delta(ae)(a,e)\,\mathrm{trm}(e)(u) \\
&= \mathrm{trm}(ae)(B_{au}) && (\mathrm{trm} = \hat{\delta},\ (5.8))
\end{aligned}
$$

If $b \neq a$, then $j(ae)(bu) = 0$ and $\mathrm{trm}(ae)(B_{bu}) = 0$, since no string starting with a is in B_{bu}. Note that we have shown the claim for all terms of the form ae.

Now let $f \in \mathsf{PTerm}$ where $(0,0) \prec \mathsf{rht}(f)$ and assume that for all e with $\mathsf{rht}(e) \prec \mathsf{rht}(f)$, that $j(e)(w) = \mathrm{trm}(e)(B_w)$.

- We next consider a term $e_1 \oplus_r e_2$. Note that $\mathsf{rht}(e_i)$ is strictly less than $\mathsf{rht}(e_1 \oplus_r e_2)$ for $i = 1, 2$, so we have assumed that $j(e_i)(w) = \mathrm{trm}(e_i)(B_w)$ for $i \in \{1, 2\}$. Then

$$
\begin{aligned}
&j(e_1 \oplus_r e_2)(w) \\
&= r j(e_1)(w) + (1-r) j(e_2)(w) \\
&= r\,\mathrm{trm}(e_1)(B_w) + (1-r)\,\mathrm{trm}(e_2)(B_w) && \text{(ind. hyp., } \mathsf{rht}(e_1), \mathsf{rht}(e_2) \prec \mathsf{rht}(e_1 \oplus_r e_2)) \\
&= r \sum_{g \in \mathsf{PTerm}} \delta(e_1)(b,g)\,\mathrm{trm}(g)(B_u) \\
&\qquad\quad + (1-r) \sum_{g \in \mathsf{PTerm}} \delta(e_2)(b,g)\,\mathrm{trm}(g)(B_u) && (w = bu),\ (5.8) \\
&= \sum_{g \in \mathsf{PTerm}} (r\delta(e_1)(b,g) + (1-r)\delta(e_2)(b,g))\,\mathrm{trm}(g)(B_u) \\
&= \sum_{g \in \mathsf{PTerm}} \delta(e_1 \oplus_r e_2)(b,g)\,\mathrm{trm}(g)(B_u) && (\text{def. } \delta) \\
&= \mathrm{trm}(e_1 \oplus_r e_2)(B_w)
\end{aligned}
$$

- Finally we consider a term $\mu v\ e$. Again, since $\mathsf{rht}(e[\mu v\ e/v]) \prec \mathsf{rht}(\mu v\ e)$, we assume that we have $j(e[\mu v\ e/v])(w) = \mathrm{trm}(e[\mu v\ e/v])(B_w)$. So we have

$$
\begin{aligned}
j(\mu v\ e)(w) &= j(e[\mu v\ e/v])(w) \\
&= \mathrm{trm}(e[\mu v\ e/v])(B_w) && \text{(ind. hyp. } \mathsf{rht}(e[\mu v\ e/v]) \prec \mathsf{rht}(\mu v\ e)) \\
&= \sum_{g \in \mathsf{PTerm}} \delta(e[\mu v\ e/v])(b,g)\ \mathrm{trm}(g)(B_u) && (w = bu,\ (5.8)) \\
&= \sum_{g \in \mathsf{PTerm}} \delta(\mu v\ e)(b,g)\,\mathrm{trm}(g)(B_u) && (\text{def. of } \delta) \\
&= \mathrm{trm}(\mu v\ e)(B_w) && (5.8)
\end{aligned}
$$

This completes the inductive step on w, hence the overall proof. $\qquad\square$

Two process terms have the same trace measure semantics iff they have the same semantics in any contraction operator interpretation. This is an easy consequence of our definitions. We have a similar result for probabilistic process terms.

Theorem 6.7. *Let $e, f \in \mathsf{PTerm}$. Then $\mathrm{trm}(e) = \mathrm{trm}(f)$ if and only if for any contraction operator interpretation $\sigma \colon A \to \mathrm{Con}(M)$, $\hat{\delta}_\sigma(e) = \hat{\delta}_\sigma(f)$.*

Proof. Let $e, f \in \mathsf{PTerm}$. Suppose $\mathrm{trm}(e) = \mathrm{trm}(f)$, and let $\sigma \colon A \to \mathrm{Con}(M)$ be any contraction operator interpretation. Then

$$\hat{\delta}_\sigma(e) = \sigma_\omega^\# \hat{\delta}(e) = \sigma_\omega^\# \hat{\delta}(f) = \hat{\delta}_\sigma(f)$$

where the first and third equalities are by definition. As for the second equality, we assumed that $\mathrm{trm}(e) = \mathrm{trm}(f)$, and so by Definition 6.5, $\hat{\delta}(e) = \hat{\delta}(f)$. Hence, $\sigma_\omega^\# \hat{\delta}(e) = \sigma_\omega^\# \hat{\delta}(f)$.

Conversely, suppose that $\hat{\delta}_\sigma(e) = \hat{\delta}_\sigma(f)$ for every contraction operator interpretation σ. Then this specifically holds for the Cantor set $(A^\omega, d_{\frac{1}{2}})$ with the 1/2-metric. We also have the prefixing operator interpretation $\sigma \colon A \to \mathrm{Con}(A^\omega)$ given by $\sigma_a = a(-)$. Indeed, each $a(-)$ has a contraction coefficient of 1/2. It suffices to show that $\mathrm{trm} = \hat{\delta}_\sigma$, since this would mean that $\mathrm{trm}(e) = \hat{\delta}_\sigma(e) = \hat{\delta}_\sigma(f) = \mathrm{trm}(f)$. As we saw in Example 3.17, σ_ω in this setting is the identity map. So are its inverse and pushforward, therefore, and so

$$\hat{\delta}_\sigma(e)(B_w) = \sigma_\omega^\# \hat{\delta}(e)(B_w) = \hat{\delta}(e)(B_w) = \mathrm{trm}(e)(B_w)$$

for any $e \in \mathsf{PTerm}$ and $w \in A^*$. It follows that $\mathrm{trm} = \hat{\delta}_\sigma$, as requested. $\qquad\square$

7. Axioms for Subfractal Measure Equivalence

In this section, we propose an inference system for deriving equations between trace measure equivalent (and therefore, subfractal measure equivalent) probabilistic process terms. Our inference system can be found in Figure 11. We show that it is *sound* with respect to trace measure semantics in this section, in the sense that if $e = f$ can be proven from the axioms, then $\mathrm{trm}(e) = \mathrm{trm}(f)$. We also establish *completeness* with respect to trace measure semantics, the converse of soundness. The completeness theorem relies on a recent result due to Cîrstea, Sokolova, Silva, and the authors [CMN$^+$25].

Similar to how Rabinovich's axiomatization of trace semantics for process terms [Rab93] adds a left-distributivity axiom to Milner's axioms for bisimilarity [Mil84], the inference system in Figure 11 adds a left-distributivity axiom (D) to Stark and Smolka's inference system [SS00] axiomatizing so-called *probabilistic bisimilarity*.[9] Due to the intimate connection between our axioms and bisimilarity of probabilistic systems, we discuss probabilistic bisimilarity next.

---

[9]The same approach to axiomatizing finite trace semantics for probabilistic systems was taken by Silva and Sokolova in [SS11].

Probabilistic Bisimilarity and Trace Measure Equivalence. Fix an LMC (X, β).

Definition 7.1. A *bisimulation equivalence* on an LMC (X, β) is an equivalence relation $R \subseteq X \times X$ such that for any $(x, y) \in R$, $a \in A$, and equivalence class $S \in X/R$,

$$\sum_{z \in S} \beta(x)(a, z) = \sum_{z \in S} \beta(y)(a, z)$$

For states $x, y \in X$, if $(x, y) \in R$ for some bisimulation equivalence R on (X, β), then we write $x \leftrightarrow y$ and say x and y are *bisimilar*. The relation $\leftrightarrow \subseteq X \times X$ is called *probabilistic bisimilarity*.

The following lemma gives a well-known characterization of probabilistic bisimilarity that will play a role in subsequent proofs.

Lemma 7.2. *Let (X, β) be an LMC. Given $x, y \in X$, x and y are bisimilar if and only if there is an LMC (Y, β_Y) and a homomorphism $h \colon (X, \beta) \to (Y, \beta_Y)$ such that $h(x) = h(y)$.*[10]

Proof. This is standard in the coalgebra literature. See, for example, [Sok11, Corollary 2.3]. □

While both bisimilarity and trace measure equivalence can be seen as observational equivalences for LMCs, probabilistic bisimilarity is a particularly granular notion of equivalence. For example, it is not true that for any LMC (X, β), $\hat{\beta} \colon X \to \mathsf{Prob}(A^\omega)$ is a homomorphism of LMCs: if $\hat{\beta}$ were a homomorphism of LMCs, then states which are trace measure equivalent would always be bisimilar. As we can see from Figure 10, however, there are examples of trace measure equivalent states that are not bisimilar.

Consequently, the set $\mathsf{Prob}(A^\omega)$ of Borel probability measures on A^ω therefore does not carry a coalgebraic structure that is universal among LMCs. On the other hand, it does carry an *algebra* structure that is universal among the class of LMCs relevant to us in this paper.
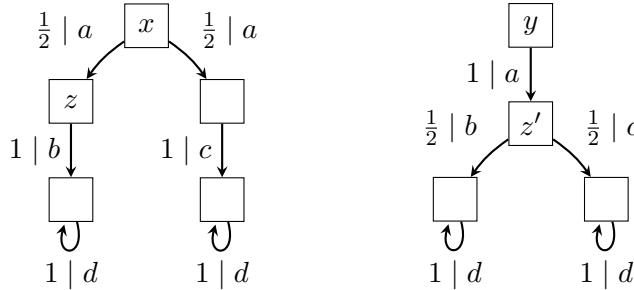

Figure 10: A standard example of trace equivalent states x, y that are not bisimilar [LS91]. If x and y were bisimilar, then there would be a bisimulation equivalence that relates them. Any such bisimulation equivalence must also relate z and z', which are not bisimilar.

---

[10]The latter condition is often called *behavioural equivalence* in the literature. See, for eg., [RBB$^+$17, Sok11].

Corollary 7.3. *Define the labelled Markov algebra* $(\mathsf{Prob}(A^\omega), \alpha)$ *by*

$$\alpha(\theta)(B) = \sum_{(a,\rho)\in\mathrm{supp}(\theta)} \theta(a,\rho)\ a^\# \rho(B)$$

Then $(\mathsf{Prob}(A^\omega), \alpha)$ *is a finitely corecursive algebra.*[11] *Furthermore, given a locally finite LMC* (X, β), *the corecursive map induced by* β *is* $\hat{\beta}$, *given by Equation 5.5.*

Proof. The first part follows from Theorem 5.15, taking M to be the Cantor set. We also take each $\sigma_a : A^\omega \to A^\omega$ to be prefixing by a; we have also written this map as $a(-)$. For these choices, the map $\sigma_\omega : A^\omega \to A^\omega$ is the identity (see Example 3.17). As Theorem 5.15 shows, the induced corecursive map is $\hat{\beta}_\sigma$. Recall that $\hat{\beta}_\sigma$ is the pushforward $\sigma_\omega^\# \hat{\beta}$. Since $\sigma_\omega = \mathsf{id}$, we see that $\hat{\beta}_\sigma = \hat{\beta}$. □

Theorem 7.4. *Let* (X, β_X) *be an LMC,* $x, x' \in X$. *If* $x \leftrightarrow x'$, *then* $\mathrm{trm}(x) = \mathrm{trm}(x')$.

Proof. Let (Y, β_Y) and $h : (X, \beta_X) \to (Y, \beta_Y)$ be as in Definition 7.1. By Corollary 7.3, the maps $\widehat{\beta_X} : X \to \mathsf{Prob}(A^\omega)$ and $\widehat{\beta_Y} : Y \to \mathsf{Prob}(A^\omega)$ are the corecursive maps induced by β_X and β_Y. So by Lemma 5.8, $\widehat{\beta_X} = \widehat{\beta_Y} \circ h$. Then

$$\mathrm{trm}(x) = \widehat{\beta_X}(x) = \widehat{\beta_Y} \circ h(x) = \widehat{\beta_Y} \circ h(x') = \widehat{\beta_X}(x') = \mathrm{trm}(x'). \qquad \square$$

Axiomatization. We are ready to present an inference system, which can be found in Figure 11, and show that it is sound with respect to the subfractal measure semantics of probabilistic process terms.

Definition 7.5. Given $e, f \in \mathsf{PTerm}$, write $e \equiv f$ and say that e and f are *provably equivalent* if the equation $e \equiv f$ can be derived from the inference rules in Figure 11.

(I) $\quad\quad e \oplus_r e \equiv e$	$(\alpha) \quad \mu w\ e \equiv \mu v\ e[v/w]$
(C) $\quad\quad e_1 \oplus_r e_2 \equiv e_2 \oplus_{1-r} e_1$	$(\mathsf{Cong}) \quad \dfrac{e_1 \equiv f_1 \ \cdots \ e_n \equiv f_n}{g[\vec{e}/\vec{v}] \equiv g[\vec{f}/\vec{v}]}$
(A) $\quad (e_1 \oplus_r e_2) \oplus_s e_3 \equiv e_1 \oplus_{rs} (e_2 \oplus_{\frac{s(1-r)}{1-rs}} e_3)$	
(D) $\quad\quad a(e_1 \oplus_r e_2) \equiv ae_1 \oplus_r ae_2$	$(\mathsf{RSP}) \quad \dfrac{g \equiv e[g/v]}{g \equiv \mu v\ e}$
(F) $\quad\quad \mu v\ e \equiv e[\mu v\ e/v]$	

Figure 11: Axioms for subfractal measure equivalence (equational logic axioms not depicted). Here, $e, g, e_i, f_i \in \mathsf{PTerm}$ for all i. In (Cong), g has precisely the free variables $v_1, \ldots, v_n$, and no variable that appears free in f_i is bound in g for any i. In (α), v does not appear free in e.

Theorem 7.6 (Soundness). *For any* $e, f \in \mathsf{PTerm}$, *if* $e \equiv f$, *then for any complete metric space* M *and any* $\sigma : A \to \mathrm{Con}(M)$, $\hat{\delta}_\sigma(e) = \hat{\delta}_\sigma(f)$.

---

[11]Recall footnote 7. Since (A^ω, d_c) is bounded (for any choice of c-metric d_c), $(\mathsf{Prob}(A^\omega), \alpha)$ is in fact corecursive, as opposed to just finitely corecursive. Again, we will not need this stronger result.

Proof. Stark and Smolka [SS00] show that all the inference rules in Figure 11 except for (D) are sound with respect to bisimilarity in (PTerm, δ).

Therefore, by Theorem 7.4, all the inference rules in Figure 11 except for (D) are sound with respect to trace measure equivalence. It therefore suffices to show that (D) is sound with respect to trace measure equivalence.

Let $e_1, e_2 \in \mathsf{PTerm}$, $a \in A$, and $r \in (0, 1)$. Then for each $w \in A^*$,

$$\mathrm{trm}(a(e_1 \oplus_r e_2))(B_w) = \begin{cases} \mathrm{trm}(e_1 \oplus_r e_2)(B_u) & w = au \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} r\,\mathrm{trm}(e_1)(B_u) + (1-r)\,\mathrm{trm}(e_2)(B_u) & w = au \\ 0 & \text{otherwise} \end{cases}$$

$$= r\,\mathrm{trm}(ae_1)(B_w) + (1-r)\,\mathrm{trm}(ae_2)(B_w)$$

$$= \mathrm{trm}(ae_1 \oplus_r ae_2)(B_w)$$

Note that the second equality used Lemma 6.6. Thus the trace measures associated to $a(e_1 \oplus_r e_2)$ and $ae_1 \oplus_r ae_2$ agree on basic open sets. By the Identity theorem [KK13, Corollary 2.5], $\mathrm{trm}(a(e_1 \oplus_r e_2))$ and $\mathrm{trm}(ae_1 \oplus_r ae_2)$ agree on on all Borel sets.   □

Theorem 7.7 (Completeness). *[CMN$^+$25, Theorem 3.6] The logical system in Figure 11 provides a complete axiomatization of trace measure semantics. That is, for any $e, f \in \mathsf{PTerm}$, if for any complete metric space M and any $\sigma \colon A \to \mathrm{Con}(M)$ we have $\hat{\delta}_\sigma(e) = \hat{\delta}_\sigma(f)$, then $e \equiv f$.*

We use Theorem 7.7 to prove another completeness theorem, the probabilistic analog of Corollary 4.4. We need the lemma below.

Lemma 7.8. *Let $f \colon A^\omega \to \mathbb{R}$ be injective and continuous. Write $f^\# \colon \mathsf{Prob}(A^\omega) \to \mathsf{Prob}(\mathbb{R})$ for the pushforward map. Then $f^\#$ is injective.*

Proof. Let $\rho_1, \rho_2 \in \mathsf{Prob}(A^\omega)$, and suppose that $f^\#(\rho_1) = f^\#(\rho_2)$. For any subset $B \subseteq A^\omega$, $f^{-1} \circ f(B) = B$, because f is injective. Furthermore, by the Lusin-Suslin theorem [Kec95, Theorem 15.1], $f(B)$ is Borel if B is Borel. This allows for the following calculation: for any Borel set $B \subseteq A^\omega$,

$$\begin{aligned} \rho_1(B) &= \rho_1(f^{-1} \circ f(B)) & (f \text{ injective}) \\ &= f_\# \rho_1(f(B)) & (f(B) \text{ Borel, def. of } f^\#) \\ &= f_\# \rho_2(f(B)) & (\text{by assumption}) \\ &= \rho_2(f^{-1} \circ f(B)) & (\text{def. of } f^\#) \\ &= \rho_2(B) & (f \text{ injective}) \end{aligned}$$

Hence, $\rho_1 = \rho_2$.   □

Theorem 7.9. *The logical system in Figure 11 provides a complete axiomatization of trace measure equivalence for affine contraction operator interpretations, i.e., for any $e, f \in \mathsf{PTerm}$, $e \equiv f$ if and only if for any affine contraction interpretation $\sigma \colon A \to \mathrm{Con}(\mathbb{R})$, $\hat{\delta}_\sigma(e) = \hat{\delta}_\sigma(f)$.*

Proof. Recall from Remark 3.14 that σ_ω is continuous for any contraction operator interpretation σ. In Proposition 3.18, we constructed a contraction operator interpretation $\sigma \colon A \to \mathrm{Con}(\mathbb{R})$ such that $\sigma_\omega \colon A^\omega \to \mathbb{R}$ is injective. This provides us with an injective continuous map $\sigma_\omega \colon \mathsf{Prob}(A^\omega) \to \mathsf{Prob}(\mathbb{R})$.

Given $e, f \in \mathsf{PTerm}$ such that $e \equiv f$, $\hat{\delta}_\sigma(e) = \hat{\delta}_\sigma(f)$ follows from soundness. For the reverse direction, suppose $\hat{\delta}_\sigma(e) = \hat{\delta}_\sigma(f)$. Then $\sigma_\omega^\# \hat{\delta}(e) = \sigma_\omega^\# \hat{\delta}(f)$. From Lemma 7.8 we see that $\sigma_\omega^\#$ is injective because σ_ω is continuous and injective. It follows that $\hat{\delta}(e) = \hat{\delta}(f)$. Finally, by Theorem 7.7, from $\hat{\delta}(e) = \hat{\delta}(f)$ we know that $e \equiv f$.   □

So we see that the subfractal measure axioms and rules in Figure 11 are complete for affine real-valued contraction operator interpretations.

8. Related Work

This paper is part of a larger effort of examining topics in continuous mathematics from the standpoint of coalgebra and theoretical computer science. The topic itself is quite old, and originates perhaps with Pavlovic and Escardó's paper "Calculus in Coinductive Form" [PE98]. Another early contribution is Pavlovic and Pratt [PP02]. These papers proposed viewing some structures in continuous mathematics—the real numbers, for example, and power series expansions—in terms of final coalgebras and streams. The next stage in this line of work was a set of papers specifically about fractal sets and final coalgebras. For example, Leinster [Lei11] offered a very general theory of self-similarity that used categorical modules in connection with the kind of gluing that is prominent in constructions of self-similar sets. In a different direction, papers like [BMRR14] showed that for some very simple fractals (such as the Sierpiński gasket treated here), the final coalgebras were Cauchy completions of the initial algebras.

Other generalizations of iterated function systems. Many generalizations of Hutchinson's self-similar sets have appeared in the literature. We mention two closely related generalizations below.

Graph IFSs. The generalization that most closely resembles our own is that of an attractor for a *directed-graph iterated function system* (or *graph IFS*) [MW88]. An LTS paired with a contraction operator interpretation on a complete metric space is equivalent data to that of a graph IFS, and equivalent statements to Theorem 3.13 can be found for example in [MW88, Edg90, MM10]. As opposed to the regular subfractal corresponding to one state, as we have studied above, the geometric object studied in the graph IFSs literature (the *attractor* for the graph IFS) is typically the union of the regular subfractals corresponding to all the states (in our terminology), and geometric properties such as Hausdorff dimension and connectivity are emphasized [MW88, Fal86, Edg90, EM92, Boo11]. We have taken this work in a slightly different direction by presenting a coalgebraic perspective on graph IFSs, seeing each state of a labelled transition system as a "recipe" for constructing fractal sets on its own. We have also allowed the interpretations of the labels to vary to obtain a semantics for process terms.

In a certain sense, regular subfractals are a more general class of objects than attractors of graph IFSs. Call a language of streams $L \subseteq A^\omega$ a *graph IFS language* if there is a finite LTS (X, α) such that $L = \bigcup_{x \in X} \mathrm{str}_\alpha(x)$. Given a contraction operator interpretation $\sigma \colon A \to \mathrm{Con}(M)$ on a complete metric space M, the fractal generated by a graph IFS consisting of (X, α) and σ is precisely the image of L under σ_ω. Every graph IFS language is the stream language emitted by a state in a finite LTS, but not conversely.

Lemma 8.1. *Let L be a graph IFS language. There is a finite LTS (X, α) and a state $x \in X$ such that $L = \mathrm{str}_\alpha(x)$.*

Proof. Let (Y, α_Y) be a finite LTS such that $L = \bigcup_{y \in Y} \mathrm{str}_{\alpha_Y}(y)$. Define $X = Y + \{x_0\}$, and

$$\alpha(z) = \begin{cases} \alpha_Y(z) & z \in Y \\ \bigcup_{y \in Y} \alpha_Y(y) & z = x_0 \end{cases}$$

In other words, $x_0 \xrightarrow{a} y$ if and only if there is a $y' \in Y$ such that $y' \xrightarrow{a} y$. Then

$$\mathrm{str}_\alpha(x_0) = \bigcup_{x_0 \xrightarrow{a} y} a \; \mathrm{str}_{\alpha_Y}(y) = \bigcup_{y' \xrightarrow{a} y} a \; \mathrm{str}_{\alpha_Y}(y) = \bigcup_{y \in Y} \mathrm{str}_{\alpha_Y}(y) = L \qquad \square$$

To show the converse of the above lemma is false, consider the language $\{(a, b, b, b, \dots)\}$ with one stream in it. This is not a graph IFS language for a simple reason: a graph IFS language includes the streams emitted by every state of an LTS, and therefore is closed under deleting initial segments. For example, if a state x emits $(a, b, b, b, \dots)$, then it must have an outoing transition $x \xrightarrow{a} y$ such that y emits $(b, b, b, \dots)$. It follows that regular stream languages, and by extension regular subfractals, properly generalize graph IFS languages.

Generalized IFSs. Another generalization of self-similar sets is Mihail and Miculescu's notion of attractor for a *generalized iterated function system* [MM10]. A generalized IFS is essentially one of Hutchinson's IFSs but with multi-arity contractions $\sigma_a \colon M^n \to M$. Attractors for generalized IFSs have been shown to be a proper generalization of self-similar sets [Str15], as well as limits of attractors for infinite IFSs [Oli23]. A common generalization of graph IFSs and generalized IFSs might be achieved by considering coalgebras of the form $X \to \mathcal{P}(\coprod_{n \in \mathbb{N}} A_n \times X^n)$ and interpreting each $a \in A_n$ as an n-ary contraction. We suspect that a similar story to the one we have outlined in this paper is possible for this common generalization.

Process algebra. The process terms we use to specify labelled transition systems and labelled Markov chains are fragments of known specification languages. Milner used process terms to specify LTSs in [Mil84], and we have repurposed his small-step semantics here. Stark and Smolka use probabilistic process terms to specify labelled Markov chains (in our terminology) in [SS00], and we have used them for the same purpose. Both of these papers also include complete axiomatizations of bisimilarity, and we have also repurposed their axioms.

However, fractal semantics is strictly coarser than bisimilarity, and in particular, fractal equivalence of process terms is trace equivalence. Rabinovich added a single axiom to Milner's axiomatization to obtain a sound and complete axiomatization of trace equivalence of μ-expressions [Rab93]. We used this axiom in our proof of Theorem 4.3. In contrast, the axiomatization of trace equivalence for probabilistic processes is only well-understood for *finite* traces, see Silva and Sokolova's [SS11], which our probabilistic process terms do not exhibit. We use the trace semantics of Kerstan and König [KK13] because it takes into account infinite traces. Infinite trace semantics of probabilistic systems has yet to see a complete axiomatization in the literature, although a complete axiomatization of the total variation distance for these systems has been obtained in the form of a so-called *quantitative equational theory* [BBLM18].

Other types of syntax. In this paper, we used the specification language of μ-terms as our basic syntax. As it happens, there are two other flavors of syntax that we could have employed. These are *iteration theories* [BÉ93], and terms in the Formal Language of Recursion FLR, especially its FLR_0 fragment. The three flavors of syntax for fixed point terms are compared in a number of papers: In [HMM+98], it was shown that there is an equivalence of categories between FLR_0 structures and iteration theories, and Bloom and Ésik make a similar connection between iteration theories and the μ-calculus in [BE94]. Again, these results describe general matters of equivalence, but it is not completely clear that for a specific space or class of spaces that they are equally powerful or equally convenient specification languages. We feel this matter deserves some investigation.

Equivalence under hypotheses. A specification language fairly close to iteration theories was used by Milius and Moss to reason about fractal constructions in [MM09] under the guise of *interpreted solutions* to recursive program schemes [MM06]. Moreover, [MM09] contains important examples of reasoning about the equality of fractal sets under assumptions about the contractions. Based on the general negative results on reasoning from hypotheses in the logic of recursion [HMM+98], we would not expect a completeness theorem for fractal equivalence under hypotheses. However, we do expect to find sound logical systems which account for interesting phenomena in the area.

9. A QUESTION ABOUT REGULAR SUBFRACTALS

Before we end this paper, we would like to pose a question regarding the relationship between regular subfractals and self-similar sets.

Certain regular subfractals that have been generated by LTSs with multiple states happen to coincide with self-similar sets using a different alphabet of action symbols and under a different contraction operator interpretation. For example, the twisted Sierpiński gasket in Figure 2 is the self-similar set generated by the iterated function system consisting of the compositions $\sigma_a, \sigma_b\sigma_b, \sigma_b\sigma_c, \sigma_c\sigma_b$, and $\sigma_c\sigma_c$.

Question 9.1. Is every regular subfractal a self-similar set? In other words, are there regular subfractals which can only be generated by a multi-state LTS?

Example 9.2. To illustrate the subtlety of this question, consider the following LTS.
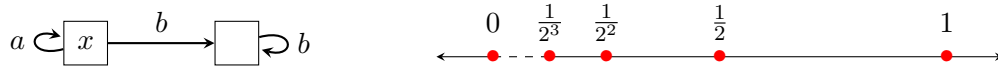
The state x emits $(a, a, \ldots)$ (an infinite stream of a's) and $(a, \ldots, a, b, b, \ldots)$, a stream with some finite number (possibly 0) of a's followed by an infinite stream of b's. Now let $M = \mathbb{R}$ with Euclidean distance and consider the contraction operator interpretation $\sigma_a(r) = \frac{1}{2}r$ and $\sigma_b(r) = \frac{1}{2}r + \frac{1}{2}$. Let $K = \{0\} \cup \{\frac{1}{2^n} | n \geq 0\}$ (depicted above). Then K is the component of the solution at x. This example is interesting because unlike the Twisted Sierpiński gasket in Figure 2, there is no obvious finite set of compositions σ_a and σ_b such that K is the self-similar set generated by that iterated function system.

There is an LTS (X, α) with X a singleton set $\{x\}$, and a contraction operator interpretation whose solution is K. We take the set of action labels underlying X to be $B = \{f, g, h\}$ and use the contraction operator interpretation $\sigma_f(r) = 0$, $\sigma_g(r) = 1$ and $\sigma_h(r) = \frac{1}{2}r$. It is easy to verify that $K = \bigcup_{i \in \{f,g,h\}} \sigma_i(K)$.

But we claim that K cannot be obtained using a single-state LTS and *the same contractions* $\sigma_a(r) = \frac{1}{2}r$ and $\sigma_b(r) = \frac{1}{2}r + \frac{1}{2}$, or using *any (finite) compositions* of σ_a and σ_b. Indeed, suppose there were such a finite collection $\sigma_1, \ldots, \sigma_n$ consisting of (finite) compositions of σ_a and σ_b such that $K = \bigcup_{i=1}^n \sigma_i(K)$. Since $1 \in K$, we must be using the stream $(b, b, b, \ldots)$ (since if there is an a at position n, the limit corresponding to the stream would be $\leq 1 - \frac{1}{2^n} < 1$), so some σ_i must consist of a composition of σ_b some number $m \geq 1$ of times with itself. Similarly, the only way to obtain 0 is with $(a, a, a, \ldots)$, so there must be some σ_j which is a composition of σ_a some number of times $p \geq 1$ with itself. But then $\lim_{n \to \infty} \sigma_i \circ \sigma_j \circ \sigma_i^n(r) = 1 - (\frac{2^p - 1}{2^{m+p}}) > \frac{1}{2}$, since $m, p \geq 1$. That point must be in the subset of $\mathbb{R}$ generated by this LTS. However, it is not in K, since $\frac{1}{2} < 1 - (\frac{2^p - 1}{2^{m+p}}) < 1$. More generally, we cannot obtain K using a single-state LTS even if we allowed finite sums of compositions of σ_a and σ_b.

Once again, it is possible to find a single state LTS whose corresponding subset of $\mathbb{R}$ is K, but to do this we needed to *change the alphabet and also the contractions*. Perhaps un-coincidentally, the constant operators are exactly the limits of the two contractions from the original interpretation. Our question is whether this can always be done.

Towards understanding Question 9.1, it may be useful to focus on a specific complete metric space M, such as $M = K = \{\frac{1}{2^n} \mid n \geq 0\} \cup \{0\}$ above, and give a characterization of all contractions on M. This would allow us to characterize its corresponding regular subfractals and self-similar sets.

Boore and Falconer have answered a restricted version of Question 9.1 in [BF13]. There, they present a graph IFS whose attractor is not the self-similar set generated by any iterated function system consisting of *similitudes*.[12] As we saw in Lemma 8.1, every graph IFS generates a regular subfractal, so the example of Boore and Falconer is indeed a regular subfractal that is not the self-similar set generated by an IFS consisting of similitudes. In other words, this is a regular subfractal which cannot be generated with a single state LTS whose contraction operators are similitudes. But not every contraction is a similitude, so the example given in [BF13] does not fully answer Question 9.1. It is unknown if their example could be generated by a single state LTS with contractions that are not similitudes.

10. Conclusion

This paper connects fractals to trace semantics, a topic originating in process algebra. This connection is our main contribution, because it opens up a line of communication between two very different areas of study. The study of fractals is a well-developed area, and like most of mathematics it is pursued without a special-purpose specification language. When we viewed process terms as recipes for fractals, we provided a specification language that was not present in the fractals literature. Of course, one also needs a contraction operator interpretation to actually define a fractal, but the separation of syntax (the process terms) and semantics (the fractals obtained using contraction operator interpretations of the syntax) is something that comes from the tradition of logic and theoretical computer science. Similarly, the use of a logical system and the emphasis on soundness and completeness is a new contribution here.

All of the above opens questions about fractals and their specifications. Our most concrete question was posed in Section 9. A more logic oriented question is inspired by

---

[12]A similitude is a contraction f such that for some $c \in [0, 1)$, $d(f(p), g(q)) = cd(p, q)$ for all $p, q \in M$.

Corollary 4.4, which states that the axioms in Figure 5 are complete for a specific contraction operator interpretation corresponding to the real-valued Cantor set. That is, no additional equations are imposed by the Cantor set interpretation. We would also like to know if we can obtain completeness theorems for other self-similar sets by allowing for extra equations in the axiomatization. For example, several additional equations between process terms are sound for the contraction operator interpretation corresponding to the Sierpinski gasket in Figure 1, including

$$a \; \mu v \; (bv) = b \; \mu v \; (av) \qquad b \; \mu v \; (cv) = c \; \mu v \; (bv) \qquad c \; \mu v \; (av) = a \; \mu v \; (cv)$$

We would like to know if adding these equations gives a complete axiomatization of the Sierpinski gasket interpretation. Lastly, and most speculatively, since LTSs (and other automata) appear so frequently in decision procedures from process algebra and verification, we would like to know if our semantics perspective on fractals can provide new complexity results in fractal geometry. We hope we have initiated a line of research where questions and answers come from both the analytic side and from theoretical computer science.

References

[AT89]   Jirí Adámek and Vera Trnková. *Automata and Algebras in Categories*. Springer Dordrecht, 1989.

[Bae05]   Jos C. M. Baeten. A brief history of process algebra. *Theor. Comput. Sci.*, 335(2-3):131–146, 2005. `doi:10.1016/j.tcs.2004.07.036`.

[Ban22]   Stefan Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta Mathematicae*, 3:133–181, 1922.

[Bar88]   Michael Barnsley. *Fractals everywhere*. Academic Press, Inc., Boston, MA, 1988.

[BBLM18]   Giorgio Bacci, Giovanni Bacci, Kim G. Larsen, and Radu Mardare. Complete axiomatization for the total variation distance of markov chains. *Electronic Notes in Theoretical Computer Science*, 336:27–39, 2018. The Thirty-third Conference on the Mathematical Foundations of Programming Semantics (MFPS XXXIII). `doi:10.1016/j.entcs.2018.03.014`.

[BÉ93]   Stephen L. Bloom and Zoltán Ésik. *Iteration Theories - The Equational Logic of Iterative Processes*. EATCS Monographs on Theoretical Computer Science. Springer, 1993. `doi:10.1007/978-3-642-78034-9`.

[BE94]   Stephen L. Bloom and Zoltán Ésik. Solving polynomial fixed point equations. In *Mathematical Foundations of Computer Science 1994 (Košice, 1994)*, volume 841 of *Lecture Notes in Comput. Sci.*, pages 52–67. Springer, Berlin, 1994.

[BF13]   G. C. Boore and K. J. Falconer. Attractors of directed graph IFSs that are not standard IFS attractors and their Hausdorff measure. *Mathematical Proceedings of the Cambridge Philosophical Society*, 154(2):325–349, 2013. `doi:10.1017/S0305004112000576`.

[BMRR14]   Prasit Bhattacharya, Lawrence S. Moss, Jayampathy Ratnayake, and Robert Rose. Fractal sets as final coalgebras obtained by completing an initial algebra. In Franck van Breugel, Elham Kashefi, Catuscia Palamidessi, and Jan Rutten, editors, *Horizons of the Mind. A Tribute to Prakash Panangaden - Essays Dedicated to Prakash Panangaden on the Occasion of His 60th Birthday*, volume 8464 of *Lecture Notes in Computer Science*, pages 146–167. Springer, 2014.

[Boo11] Graeme C. Boore. *Directed Graph Iterated Function Systems*. PhD thesis, University of St. Andrews, 2011.

[BPS01] Jan A. Bergstra, Alban Ponse, and Scott A. Smolka, editors. *Handbook of Process Algebra*. North-Holland / Elsevier, 2001. `doi:10.1016/b978-0-444-82830-9.x5017-6`.

[BR07] Vladimir Igorevich Bogachev and Maria Aparecida Soares Ruas. *Measure Theory*, volume 1. Springer, 2007. `doi:10.1007/978-3-540-34514-5`.

[BS81] S. Burris and H. P. Sankappanavar. *A Course in Universal Algebra*. Springer New York, NY, 1981.

[CMN+25] Corina Cîrstea, Lawrence S. Moss, Victoria Noquez, Todd Schmid, Alexandra Silva, and Ana Sokolova. A Complete Inference System for Probabilistic Infinite Trace Equivalence. In Jörg Endrullis and Sylvain Schmitz, editors, *33rd EACSL Annual Conference on Computer Science Logic (CSL 2025)*, volume 326 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 30:1–30:23, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/LIPIcs.CSL.2025.30`.

[CUV06] Venanzio Capretta, Tarmo Uustalu, and Varmo Vene. Recursive coalgebras from comonads. *Inform. Comput.*, 204:437–468, 2006.

[DS84] Percy Diaconis and Mehrdad Shahshahani. Products of random matrices and computer image generation. Technical report, Stanford University, 1984.

[Edg90] Gerald A. Edgar. *Measure, Topology, and Fractal Geometry*. Springer New York, NY, 1 edition, 1990. `doi:10.1007/978-1-4757-4134-6`.

[Elt87] John H. Elton. An erogodic theorem for iterated maps. *Ergodic Theory and Dynamical Systems*, 7(4), 1987. `doi:10.1017/S0143385700004168`.

[EM92] G. A. Edgar and R. Daniel Mauldin. Multifractal decompositions of digraph recursive fractals. *Proceedings of the London Mathematical Society*, s3-65(3):604–628, 1992. `doi:10.1112/plms/s3-65.3.604`.

[Fal86] Kenneth J. Falconer. *The Geometry of Fractal Sets*. Cambridge Tracts in Mathematics. Cambridge University Press, 1986. `doi:10.1017/CBO9780511623738`.

[Fol99] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Wiley, 2nd edition, 1999.

[GR18] Alexandre Goy and Jurriaan Rot. (In)finite trace equivalence of probabilistic transition systems. In Corina Cîrstea, editor, *Coalgebraic Methods in Computer Science - 14th IFIP WG 1.3 International Workshop, CMCS 2018, Colocated with ETAPS 2018, Thessaloniki, Greece, April 14-15, 2018, Revised Selected Papers*, volume 11202 of *Lecture Notes in Computer Science*, pages 100–121. Springer, 2018. `doi:10.1007/978-3-030-00389-0\_7`.

[Gro15] Wilh Groß. Grundzüge der mengenlehre. *Monatshefte für Mathematik und Physik*, 26(1):A34–A35, 1915. `doi:10.1007/BF01999507`.

[GS02] H. Peter Gumm and Tobias Schröder. Coalgebras of bounded type. *Math. Struct. Comput. Sci.*, 12(5):565–578, 2002. `doi:10.1017/S0960129501003590`.

[HMM+98] A. J. C. Hurkens, Monica McArthur, Yiannis N. Moschovakis, Lawrence S. Moss, and Glen T. Whitney. The logic of recursive equations. *J. Symbolic Logic*, 63(2):451–478, 1998.

[Hoa78] C. A. R. Hoare. Communicating sequential processes. *Commun. ACM*, 21(8):666–677, 1978. `doi:10.1145/359576.359585`.

[Hut81] John E. Hutchinson. Fractals and self similarity. *Indiana University Mathematics Journal*, 30(5):713–747, 1981.

[Kec95] A. Kechris. *Classical Descriptive Set Theory*. Graduate Texts in Mathematics. Springer New York, 1995.

[KK13] Henning Kerstan and Barbara König. Coalgebraic trace semantics for continuous probabilistic transition systems. *Logical Methods in Computer Science*, Volume 9, Issue 4, Dec 2013. `doi:10.2168/LMCS-9(4:16)2013`.

[Kle56] S. C. Kleene. Representation of events in nerve nets and finite automata. In Claude Shannon and John McCarthy, editors, *Automata Studies*, pages 3–41. Princeton University Press, Princeton, NJ, 1956.

[Kle02] Stephen Cole Kleene. *Mathematical Logic*. Courier Corporation, 2002.

[Kra06] A. S. Kravchenko. Completeness of the space of separable measures in the kantorovich-rubenstein metric. *Siberian Mathematical Journal*, 47(1), 2006.

[LAS78]   J. Arthur Seebach Lynn Arthur Steen. *Counterexamples in Topology*. Springer New York, NY, 1978. `doi:10.1007/978-1-4612-6290-9`.

[Lei11]   Tom Leinster. A general theory of self-similarity. *Adv. Math.*, 226(4):2935–3017, 2011.

[LS91]   Kim Guldstrand Larsen and Arne Skou. Bisimulation through probabilistic testing. *Inf. Comput.*, 94(1):1–28, 1991. `doi:10.1016/0890-5401(91)90030-6`.

[Man77]   Benoit B. Mandelbrot. *Fractals: Form, Chance, and Dimension*. Mathematics Series. W. H. Freeman, 1977.

[Mil78]   Robin Milner. Synthesis of communicating behaviour. In Józef Winkowski, editor, *Mathematical Foundations of Computer Science 1978, Proceedings, 7th Symposium, Zakopane, Poland, September 4-8, 1978*, volume 64 of *Lecture Notes in Computer Science*, pages 71–83. Springer, 1978. `doi:10.1007/3-540-08921-7\_57`.

[Mil84]   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`.

[MM06]   Stefan Milius and Lawrence S. Moss. The category-theoretic solution of recursive program schemes. *Theor. Comput. Sci.*, 366(1-2):3–59, 2006. `doi:10.1016/j.tcs.2006.07.002`.

[MM09]   Stefan Milius and Lawrence S. Moss. Equational properties of recursive program scheme solutions. *Cah. Topol. Géom. Différ. Catég.*, 50(1):23–66, 2009.

[MM10]   Alexandru Mihail and Radu Miculescu. Generalized IFSs on noncompact spaces. *Fixed Point Theory and Applications*, 1(584215), 2010. `doi:10.1155/2010/584215`.

[MW88]   R Daniel Mauldin and S. C. Williams. Hausdorff dimension in graph directed constructions. *Transactions of the American Mathematical Society*, 309(2):811–829, 1988. `doi:10.1090/S0002-9947-1988-0961615-4`.

[Oli23]   Elismar R. Oliveira. The Hutchinson-Barnsley theory for generalized iterated function systems by means of infinite iterated function systems, 2023. `arXiv:2204.00373`.

[PE98]   Dusko Pavlovic and Martín Hötzel Escardó. Calculus in coinductive form. In *Thirteenth Annual IEEE Symposium on Logic in Computer Science, Indianapolis, Indiana, USA, June 21-24, 1998*, pages 408–417. IEEE Computer Society, 1998.

[PP02]   Dusko Pavlovic and Vaughan R. Pratt. The continuum as a final coalgebra. *Theor. Comput. Sci.*, 280(1-2):105–122, 2002.

[Rab93]   Alexander Moshe Rabinovich. A complete axiomatisation for trace congruence of finite state behaviors. In Stephen D. Brookes, Michael G. Main, Austin Melton, Michael W. Mislove, and David A. Schmidt, editors, *Mathematical Foundations of Programming Semantics, 9th International Conference, New Orleans, LA, USA, April 7-10, 1993, Proceedings*, volume 802 of *Lecture Notes in Computer Science*, pages 530–543. Springer, 1993. `doi:10.1007/3-540-58027-1\_25`.

[RBB+17]   Jurriaan Rot, Filippo Bonchi, Marcello M. Bonsangue, Damien Pous, Jan Rutten, and Alexandra Silva. Enhanced coalgebraic bisimulation. *Math. Struct. Comput. Sci.*, 27(7):1236–1264, 2017. `doi:10.1017/S0960129515000523`.

[Rut00]   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`.

[Sok11]   Ana Sokolova. Probabilistic systems coalgebraically: A survey. *Theor. Comput. Sci.*, 412(38):5095–5110, 2011. `doi:10.1016/J.TCS.2011.05.008`.

[SS00]   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.

[SS11]   Alexandra Silva and Ana Sokolova. Sound and complete axiomatization of trace semantics for probabilistic systems. In Michael W. Mislove and Joël Ouaknine, editors, *Twenty-seventh Conference on the Mathematical Foundations of Programming Semantics, MFPS 2011, Pittsburgh, PA, USA, May 25-28, 2011*, volume 276 of *Electronic Notes in Theoretical Computer Science*, pages 291–311. Elsevier, 2011. `doi:10.1016/j.entcs.2011.09.027`.

[Str15]   Filip Strobin. Attractors of generalized IFSs that are not attractors of IFSs. *Journal of Mathematical Analysis and Applications*, 422(1):99–108, 2015. `doi:10.1016/j.jmaa.2014.08.029`.