Edinburgh Research Explorer

Fixed-Points for Quantitative Equational Logics

Fixed-Points for Quantitative Equational Logics

Radu Mardare
Department of Computer Science
University of Aalborg

Prakash Panangaden*
School of Computer Science
McGill University

Gordon Plotkin
School of Informatics
University of Edinburgh

1$^{\text{st}}$ July 2021

Abstract

We develop a fixed-point extension of quantitative equational logic and give semantics in one-bounded complete quantitative algebras. Unlike previous related work about fixed-points in metric spaces, we are working with the notion of approximate equality rather than exact equality. The result is a novel theory of fixed points which can not only provide solutions to the traditional fixed-point equations but we can also define the rate of convergence to the fixed point. We show that such a theory is the quantitative analogue of a Conway theory and also of an iteration theory; and it reflects the metric coinduction principle. We study the Bellman equation for a Markov decision process as an illustrative example.

1 Introduction

Quantitative equational logic was introduced in [MPP16, MPP17] as a way of generalizing the standard concept of equational logic to encompass the concept of approximate equality. Essentially, it allows one to use a logical framework to perform metric reasoning. The present work is an extension of that formalism to reason about fixed points of functions. Fixed point theory is the mathematical way to understand recursion and iteration [SB, Bak71] and was extensively studied in a partial order setting based ultimately on Kleene's fixed point theorem [Kle52] or some other related fixed-point theorem like the Knaester-Tarski theorem. In this paper we develop the metric version of fixed point theory based on the Banach fixed point theorem, which says that contractive functions on a bounded complete metric space have *unique* fixed points.

We follow the categorical axiomatization of fixed-point theories by Simpson and Plotkin [SP00], which focusses on the *Conway theories* developed independently

---

by Bloom and Esik [BE93] and by Hasegawa [Has99]. We develop an axiomatization that satisfies quantitative analogues of their formulations. We are also able to leverage the completeness proof from [MPP16] to obtain a completeness result in our case. We also give an axiomatization of fixed-point operators and show how one can reason about convergence and convergence rates. We study the relation to a metric coinduction principle due to Kozen [Koz06, Koz07]: our axiomatization is the metric analogue of *Park induction* and the Kozen coinduction principle is the quantitative version of *Scott induction*, see [EB95] for a comprehensive presentation of these. Finally we develop an extended example: the Bellman equation for Markov Decision Processes [Put94] which plays a central role in reinforcement learning [SB98].

We summarize very briefly the formalism introduced in [MPP16, MPP17]. The equality symbol = is annotated by a (small) real number ε so that one can write *approximate equality* statements of the form: $s =_\varepsilon t$, where s, t are terms of some theory. Intuitively, one thinks of this as meaning that s and t are "within ε" of each other. The rules of quantitative equational logic are analogous to the rules for ordinary equational logic except for an infinitary "continuity" rule that allows one to infer $s =_\varepsilon t$ from $s =_{\varepsilon_i} t$ where the ε_i converge to ε from above. One can then introduce quantitative algebras which are algebras that have metric structure and in which all the operations are nonexpansive. A completeness theorem is established and it is shown that free algebras can be defined and one can relate theories to monads on suitable categories of metric spaces. One of the main examples given in [MPP16] is related to spaces of probability distributions with the Kantorovich metric.

The authors of [MPP16] have used *extended* metrics: metrics that can take on infinite values. We have used 1-bounded metrics in this paper instead. From the topological point of view these are the same: by using the standard transformation $d'(x, y) = d(x, y)/(1 + d(x, y))$ one can transform the extended metric d into a 1-bounded metric *with the same topology*. Interestingly, under this transformation a contractive function in the 1-bounded sense becomes a function that moves all points into the same connected component in the extended metric sense.

There is a comprehensive study of iteration theories [BE93] which develops a variety of examples including metric fixed point theories. We will comment on this and other interesting related work [GS18, Koz06] at the end of this paper. For now we remark that other treatments of metric fixed-point theories are based on the traditional notion of equality and hence do not allow quantitative reasoning about convergence. There are a number of examples from [MPP16], such as barycentric algebras, that cannot be done without the quantitative setting. We also have new examples such as the combination of probabilistic choice and nondeterminism.

In order to carry out our program we are forced to keep track not just of the fact that functions are contractive but *exactly how contractive they are* and, furthermore, we need to track this information for each input to the function. So the traditional notion of arity needs to be enriched with quantitative infor-

2

mation that we call *Banach patterns*. The details are, in some places, intricate but the intuition will be, we hope, clear. We have not seen any related work that keeps track of this kind of quantitative information.

2 Notation

In what follows we will often manipulate tuples of real numbers. These encode the contractiveness information that we need in order to be able to define fixed points, and are useful for managing sets of variables in complex terms.

If $\overline{\alpha} = \langle \alpha_1, .., \alpha_n \rangle$, $\overline{\beta} = \langle \beta_1, .., \beta_m \rangle$ are tuples for $n \geq 1$ and $i \leq n$, let $|\overline{\alpha}| = n$ and we use the following notations
$\overline{\alpha} \setminus i = \langle \alpha_1, ..\alpha_{i-1}, \alpha_{i+1}, ..\alpha_n \rangle$,
for $x \in \mathbb{R}$, $\overline{\alpha}[x/i] = \langle \alpha_1, ..\alpha_{i-1}, x, \alpha_{i+1}, ..\alpha_n \rangle$ and
$\overline{\alpha}[\overline{\beta}/i] = \langle \alpha_1, ..\alpha_{i-1}, \beta_1 ..\beta_m, \alpha_{i+1}, ..\alpha_n \rangle$.
If we have a tuple $\overline{\alpha}$, we denote its i-th component by α_i.

Let $\mathbb{U}_n$ denote the set of all tuples $\overline{\alpha} = \langle \alpha_1 ..\alpha_n \rangle \in [0, 1]^n$ s.t. $\displaystyle\sum_{1 \leq i \leq n} \alpha_i \leq 1$. And let $\mathbb{U} = \displaystyle\bigcup_{i \geq 0} \mathbb{U}_i$.

For arbitrary $\overline{\alpha}, \overline{\alpha}^1, \ldots, \overline{\alpha}^n, \langle \lambda_1, .., \lambda_n \rangle \in \mathbb{U}_n$ and $r \leq 1$, we define the following operations:

1. Scalar multiplication. $r\overline{\alpha} = \langle r\alpha_1, \ldots, r\alpha_n \rangle$

2. Subconvex sum. $\sum_i \lambda_i \overline{\alpha}^i = \langle \sum_i \lambda_i \alpha_1^i, .., \sum_i \lambda_n \alpha_n^i \rangle$

3. Contraction. For $i < j$, $\overline{\alpha}[i < j] = (\overline{\alpha} \setminus j)[\alpha_i + \alpha_j/i]$.

4. Iteration. For $i \leq n$ s.t. $\alpha_i < 1$, $\mu i.\overline{\alpha} = \frac{1}{1-\alpha_i}(\overline{\alpha} \setminus i)$.

2.1 Banach patterns

In what follows we introduce the concept of *Banach pattern* that will be used to characterize nonexpansive functions on metric spaces. Recall that if (A, d^A) and (B, d^B) are metric spaces, then $f : (A, d^A)^n \to (B, d^B)$ is a nonexpansive function if for arbitrary $\langle a_1, .., a_n \rangle, \langle b_1, .., b_n \rangle \in A^n$,

$$d^B(f(a_1..a_n), f(b_1..b_n)) \leq \max_{i \leq n} d^A(a_i, b_i).$$

Definition 2.1 *Let* $f : (A, d^A)^n \to (B, d^B)$ *be a function between two metric spaces.* f *admits Banach patterns if there exists a set* $\theta \subseteq_{fin} \mathbb{U}_n$ *such that for any* $\langle a_1..a_n \rangle, \langle b_1..b_n \rangle \in A^n$,

$$d^B(f(a_1..a_n), f(b_1..b_n)) \leq \max_{\overline{\alpha} \in \theta} \sum_{i \leq n} \alpha_i d^A(a_i, b_i).$$

In this case, θ *is a **Banach pattern** for* f, *and we write* $f : n : \theta$.

Example 2.2 *Let (M, d) be a 1-bounded metric space and $\Delta(M, d)$ the space of Borel probability distributions on (M, d) metrized with the Kantorovich metric $K^d : \Delta(M, d)^2 \to [0, 1]$.*

Consider, for $\epsilon \in [0, 1]$, the barycentric operation on $\Delta(M, d)$, $+_\epsilon : \Delta(M, d)^2 \to \Delta(M, d)$ defined for arbitrary $\mu, \nu \in \Delta(M, d)$ by

$$\mu +_\epsilon \nu = \epsilon\mu + (1 - \epsilon)\nu.$$

In [MPP16] it has been demonstrated that for arbitrary $\mu, \mu', \nu, \nu' \in \Delta(M, d)$,

$$K^d(\mu +_\epsilon \mu', \nu +_\epsilon \nu') \le \epsilon K^d(\mu, \mu') + (1 - \epsilon) K^d(\nu, \nu'),$$

hence, $+_\epsilon$ has Banach pattern the singleton $\{\langle \epsilon, 1 - \epsilon \rangle\}$.

Example 2.3 *For another example where the pattern is not a singleton we consider the non-deterministic choice function on Δ, $\oplus : \Delta(M, d)^2 \to H(\Delta(M, d))$, where for a metric space X, HX denotes the space of compact subsets equipped with the Hausdorff metric. The function $\oplus$ is nonexpansive in the Hausdorff metric defined for K^d, [MPP16]. Being nonexpansive in this sense, this function satisfies for arbitrary $\mu, \mu', \nu, \nu' \in \Delta(M, d)$,*

$$K^d(\mu \oplus \mu', \nu \oplus \nu') \le \max\{K^d(\mu, \nu), K^d(\mu', \nu')\}.$$

In this case the Banach pattern is not a singleton, but we have $\oplus : 2 : \{\langle 0, 1 \rangle, \langle 1, 0 \rangle\}$.

Example 2.4 *For a third example, we consider, the function*

$$f : \Delta(M, d)^3 \to \Delta(M, d)$$

defined, for arbitrary $\mu, \nu, \eta \in \Delta(M, d)$ by

$$f(\mu, \nu, \eta) = (\mu +_\epsilon \nu) \oplus \eta,$$

for some $\epsilon \le 1$. We note that for arbitrary $\mu, \nu, \eta, \mu', \nu', \eta' \in \Delta(M, d)$,

$$K^d(f(\mu, \nu, \eta), f(\mu', \nu', \eta'))$$

$$\le \max\{\epsilon K^d(\mu, \mu') + (1 - \epsilon) K^d(\nu, \nu'), K^d(\eta, \eta')\},$$

and in this case we have $f : 3 : \{\langle \epsilon, 1 - \epsilon, 0 \rangle, \langle 0, 0, 1 \rangle\}$.

Observe that a function $f : (A, d^A)^n \to (B, d^B)$ is nonexpansive iff it admiths Banach patterns. Indeed, if f is nonexpansive, then

$$\{\langle 1, 0, ..0 \rangle, \langle 0, 1, 0.., 0 \rangle, .., \langle 0, .., 0, 1 \rangle\} \subseteq U_n$$

is a Banach pattern for it, the one encoding exactly the nonexpansiveness property. And reverse, if f admits a Banach pattern $\theta \subseteq_{fin} \mathbb{U}_n$, then nonexpansivess derives from

$$\max_{\overline{\alpha} \in \theta} \sum_{i \le n} \alpha_i d^A(a_i, b_i) \le \max_{i \le n} d^A(a_i, b_i).$$

4

However, often a Banach pattern brings more information about the nonexpansiveness of a function.

We will add Banach patterns to the algebraic signatures over the category of metric spaces when we will define quantitative algebras with fixed points.

It is useful to define some operations on patterns, in addition to the set theoretic operations. Let $\theta, \theta^1..\theta^n \subseteq \mathbb{U}_n$, $\lambda \leq 1$ and $\langle \lambda_1, .., \lambda_n \rangle \in \mathbb{U}_n$.

1. Scalar multiplication. $\lambda\theta = \{\lambda\overline{\alpha} \mid \overline{\alpha} \in \theta\} \subseteq \mathbb{U}_n$.

2. Subconvex sum. $\displaystyle\sum_{i \leq n} \lambda_i\theta^i = \{\sum_{i \leq n} \lambda_i\overline{\alpha}^i \mid \overline{\alpha}^i \in \theta_i\}$.

3. Contraction. $\theta[i < j] = \{\overline{\alpha}[i < j] \mid \overline{\alpha} \in \theta\} \subseteq \mathbb{U}_{n-1}$.

4. Composition. For $\zeta_1..\zeta_n \in \mathbb{U}_m$,

$$\theta \circ \langle \zeta_1..\zeta_n \rangle = \{\sum_{i \leq n} \alpha_i\overline{\beta}^i \mid \overline{\alpha} \in \theta, \overline{\beta}^i \in \zeta_i\} \subseteq \mathbb{U}_m.$$

5. Fixed point. If for all $\overline{\alpha} \in \theta$, $\alpha_i < 1$, let

$$\mu i.\theta = \{\mu i.\overline{\alpha} \mid \overline{\alpha} \in \theta\} \subseteq \mathbb{U}_{n-1}.$$

Whenever θ satisfies $[\forall\overline{\alpha} \in \theta, \alpha_i < 1]$, we say that θ is *i-contractive* and denote this by $\theta \triangleright i$.

We also generalize the notation we introduced for tuples and for $\theta \subseteq \mathbb{U}_n$, $\zeta \subseteq \mathbb{U}_m$ and $i \leq n$, let
$\theta \setminus i = \{\overline{\alpha} \setminus i \mid \overline{\alpha} \in \theta\} \subseteq \mathbb{U}_{n-1}$ and
$\theta[\zeta/i] = \{\overline{\alpha}[\overline{\beta}/i] \mid \overline{\alpha} \in \theta, \overline{\beta} \in \zeta\} \subseteq \mathbb{U}_{n+m-1}$.

3 Quantitative Equational Reasoning

In this section we recall the main concepts of quantitative equational reasoning and quantitative algebras [MPP16].

3.1 Quantitative Equational Theory

We start with a *signature* Ω, which is a set of function symbols of finite arity (constants have arity 0). We write $f : n \in \Omega$ for a function f of arity $n \geq 0$.

Given a set X, let $\hat{\Omega}X$ be the Ω-*algebra generated by* X, i.e., the set of all terms constructed on top of X by using the functions in Ω. Note that this set comes already equipped with the structure of an Ω-algebra.

For a set X of *variables*, one defines *quantitative equations*[1] over $\hat{\Omega}X$, which

---

[1]In [MPP16] quantitative equations are defined for $\epsilon \in \mathbb{Q}_+$. We chose to avoid this restriction here in order to get a simpler axiomatization. However, all these developments work properly if we restrict to rational indices.

have the form $t =_\epsilon s$ for $t, s \in \hat{\Omega}X$ and $\epsilon \in \mathbb{R}_+$. We use $\mathcal{E}(\hat{\Omega}X)$ to denote the set of quantitative equations on $\hat{\Omega}X$.

Let $\mathcal{J}(\hat{\Omega}X)$ be the class of *quantitative judgements* on $\hat{\Omega}X$, which are constructions of the form

$$\{s_i =_{\epsilon_i} t_i \mid i \in I\} \vdash s =_\epsilon t,$$

where I is a countable (possible empty) index set, $s_i, t_i, s, t \in \hat{\Omega}X$ and $\epsilon_i, \epsilon \in \mathbb{R}_+$ for all $i \in I$.

If $\Gamma \vdash \phi \in \mathcal{J}(\hat{\Omega}X)$, where $\Gamma \subseteq \mathcal{E}(\hat{\Omega}X)$ and $\phi \in \mathcal{E}(\hat{\Omega}X)$, we refer to the elements of Γ as the *hypotheses* and to ϕ as the *conclusion* of the quantitative judgement.

Definition 3.1 (Quantitative Equational Theory) *Given a signature Ω and a set X of variables, the* deductive closure *of a set $\mathcal{U}$ of quantitative judgements on $\hat{\Omega}X$ is the smallest set $\overline{\mathcal{U}}$ of quantitative judgements on $\hat{\Omega}X$ such that $\mathcal{U} \subseteq \overline{\mathcal{U}}$, and for arbitrary $t, s \in \hat{\Omega}X$, $\epsilon, \epsilon' \in \mathbb{R}_+$, $f : |I| \in \Omega$, $\Gamma, \Theta \subseteq \mathcal{E}(\hat{\Omega}X)$ and $\overline{s} = (s_i)_{i \in I}, \overline{t} = (t_i)_{i \in I} \subseteq \hat{\Omega}X$ and any substitution σ*

$$
\begin{aligned}
\textbf{(Refl)} \quad & \vdash t =_0 t \in \overline{\mathcal{U}}, \\
\textbf{(Symm)} \quad & \{t =_\epsilon s\} \vdash s =_\epsilon t \in \overline{\mathcal{U}}, \\
\textbf{(Triang)} \quad & \{t =_\epsilon u, u =_{\epsilon'} s\} \vdash t =_{\epsilon+\epsilon'} s \in \overline{\mathcal{U}}, \\
\textbf{(Max)} \quad & \{t =_\epsilon s\} \vdash t =_{\epsilon+\epsilon'} s \in \overline{\mathcal{U}}, \text{ for all } \epsilon' > 0, \\
\textbf{(NExp)} \quad & \{t_i =_\epsilon s_i \mid i \in I\} \vdash f(\overline{t}) =_\epsilon f(\overline{s}) \in \overline{\mathcal{U}};
\end{aligned}
$$

and $\overline{\mathcal{U}}$ is closed under the following rules

$$
\begin{aligned}
\textbf{(Cont)} \quad & \frac{\Gamma \vdash s =_{\epsilon'} t \text{ for all } \epsilon' > \epsilon}{\Gamma \vdash s =_\epsilon t}, \\[2mm]
\textbf{(Subst)} \quad & \frac{\Gamma \vdash t =_\epsilon s}{\sigma(\Gamma) \vdash \sigma(t) =_\epsilon \sigma(t)}, \\[2mm]
\textbf{(Assumpt)} \quad & \frac{t =_\epsilon s \in \Gamma}{\Gamma \vdash t =_\epsilon s}, \\[2mm]
\textbf{(Cut)} \quad & \frac{\Theta \vdash t =_\epsilon s, \ \Gamma \vdash \Theta}{\Gamma \vdash t =_\epsilon s}.
\end{aligned}
$$

where $\Gamma \vdash \Theta$ means that $\Gamma \vdash \phi$ for all $\phi \in \Theta$. A quantitative equational theory *of signature Ω over X is a set $\mathcal{U}$ of quantitative judgements on $\hat{\Omega}X$ such that*

$$\mathcal{U} = \overline{\mathcal{U}}.$$

Definition 3.2 (Quantitative Algebra) *Given a signature Ω, a* quantitative algebra *over Ω is a tuple $\mathcal{A} = (A, \Omega^{\mathcal{A}}, d^{\mathcal{A}})$, where $(A, \Omega^{\mathcal{A}})$ is an algebra of signature Ω, (A, d) is a metric space and any $f : |I| \in \Omega$ is nonexpansive.*

A *homomorphism of quantitative algebras* is a *non-expansive* Ω-homomorphism (of Ω-algebras).

Given a quantitative algebra $\mathcal{A} = (A, \Omega^{\mathcal{A}}, d^{\mathcal{A}})$ of signature Ω and a set X of variables, an *assignment* on $\mathcal{A}$ is a function $\alpha : X \rightarrow A$. It can be canonically

extended to a homomorphism of Ω-algebras $\alpha : \hat{\Omega}X \to A$ by defining, for any $f : |I| \in \Omega$ and any $(t_i)_{i \in I} \subseteq \hat{\Omega}X$,

$$\alpha(f((t_i)_{i \in I})) = f^{\mathcal{A}}((\alpha(t_i))_{i \in I}).$$

We denote by $\Omega[X|\mathcal{A}]$ the set of assignments on $\mathcal{A}$.

Definition 3.3 (Satisfaction) *Let $\mathcal{A} = (A, \Omega^{\mathcal{A}}, d^{\mathcal{A}})$ be an Ω-quantitative algebra and $\{s_i =_{\epsilon_i} t_i \mid i \in I\} \vdash s =_{\epsilon} t$ a quantitative judgement on $\hat{\Omega}X$. $\mathcal{A}$* satisfies *this quantitative judgement under the assignment $\alpha \in \Omega[X|\mathcal{A}]$, written*

$$\{s_i =_{\epsilon_i} t_i \mid i \in I\} \models_{\mathcal{A},\alpha} s =_{\epsilon} t,$$

if $[\forall i \in I, d^{\mathcal{A}}(\alpha(t_i), \alpha(s_i)) \leq \epsilon_i]$ implies $d^{\mathcal{A}}(\alpha(s), \alpha(t)) \leq \epsilon$.
We say $\mathcal{A}$ satisfies the quantitative judgement, *or it is a* model *of the quantitative judgement, written*

$$\{s_i =_{\epsilon_i} t_i \mid i \in I\} \models_{\mathcal{A}} s =_{\epsilon} t,$$

if

$$\forall \alpha \in \Omega[X|\mathcal{A}], \ \{s_i =_{\epsilon_i} t_i \mid i \in I\} \models_{\mathcal{A},\alpha} s =_{\epsilon} t.$$

Similarly, for a set of quantitative judgements (or a quantitative equational theory) $\mathcal{U}$, we say that $\mathcal{A}$ is a model of $\mathcal{U}$ if $\mathcal{A}$ satisfies every element of $\mathcal{U}$; for simplifying notation we denote this by $\mathcal{A} \models \mathcal{U}$. Let $\mathbf{QA}(\mathcal{U})$ denote the set of models of $\mathcal{U}$. If $\mathfrak{M}$ is a set of Ω-quantitative algebras and $\Gamma \vdash \phi \in \mathcal{J}(\hat{\Omega}X)$, we write $\Gamma \models_{\mathfrak{M}} \phi$ whenever $\Gamma \models_{\mathcal{A}} \phi$ for all $\mathcal{A} \in \mathfrak{M}$. In [MPP16] the following completeness result is established.

Theorem 3.4 (Completeness) *Given a quantitative equational theory $\mathcal{U}$ over $\hat{\Omega}X$,*

$$\Gamma \models_{\mathbf{QA}(\mathcal{U})} \phi \quad \text{iff} \quad \Gamma \vdash \phi \in \mathcal{U}.$$

3.2 Limits in quantitative theories

Although not explicit in [MPP16], quantitative equational theories have built in the mechanism for equational reasoning about convergence: this will be useful to us.

Definition 3.5 *In general, given a quantitative equational theory $\mathcal{U}$ over $\hat{\Omega}X$, we say that a sequence $(s_i)_{i \geq 1} \subseteq \hat{\Omega}X$ is* convergent *in $\mathcal{U}$ if there exists $s \in \hat{\Omega}X$ such that*

$$\forall \epsilon > 0 \ \exists k \ \forall i \geq k, \ \vdash s_i =_{\epsilon} s \in \mathcal{U}.$$

We say that s is a limit *of the sequence $(s_i)_{i \geq 1}$.*

It is easy to prove the following using (Triang), (Symm) and (Cont).

Proposition 3.6 *Let $\mathcal{U}$ be a quantitative equational theory over $\hat{\Omega}X$. If the sequence $(s_i)_{i \geq 1} \subseteq \hat{\Omega}X$ is convergent in $\mathcal{U}$ and it has both $s \in \hat{\Omega}X$ and $t \in \hat{\Omega}X$ as limits, then*

$$\vdash s =_0 t \in \mathcal{U}.$$

This motivates us to denote the limit of the sequence $(s_i)_{i\geq1}$ by $\lim_i s_i$.

We can construct convergent sequences of terms by applying non-expansive functions to convergent sequences.

Lemma 3.7 *Let $\mathcal{U}$ be a quantitative equational theory over $\hat{\Omega}X$, $f:n, g:m \in \hat{\Omega}X$ and $(s_k)_{k\geq1} \subseteq \hat{\Omega}X$ be a convergent sequence in $\mathcal{U}$. Then for $\overline{x} \subseteq X^n$ and $\overline{y} \in X^m$,*
(1). $(g(\overline{y}[s_k/j]))_{k\geq1}$ is a convergent sequence in $\mathcal{U}$ and

$$\vdash \lim_k g(\overline{y}[s_k/j]) =_0 g(\overline{y}[\lim_k s_k/j]) \in \mathcal{U}.$$

(2). $(f(\overline{x}[g(\overline{y}[s_k/j])/i]))_{k\geq1}$ is convergent in $\mathcal{U}$ and

$$\vdash \lim_k f(\overline{x}[g(\overline{y}[s_k/j])/i]) =_0 f(\overline{x}[\lim_k g(\overline{y}[s_k/j])/i]) \in \mathcal{U}.$$

Proof. (1). Let $s = \lim_k s_k$. Hence, $\forall \epsilon > 0 \ \exists p \ \forall k \geq p$,
$\vdash s_k =_\epsilon s \in \mathcal{U}$. Applying (NExp), we get that
$\forall \epsilon > 0 \ \exists p \ \forall i \geq p, \vdash g(\overline{y}[s_k/j]) =_\epsilon g(\overline{y}[s/i]) \in \mathcal{U}$, i.e.,
$\vdash \lim_k g(\overline{y}[s_k/j]) =_0 g(\overline{y}[\lim_k s_k/j]) \in \mathcal{U}$.
(2). After observing that $f(\overline{x}[g(\overline{y})/i])$ is nonexpansive, we conclude, as above, that $(f(\overline{x}[g(\overline{y}[s_k/j])/i]))_{k\geq1}$ is convergent in $\mathcal{U}$. Next, we apply (1) and prove that
$\vdash \lim_k f(\overline{x}[g(\overline{y}[s_k/j])/i]) =_0 f(\overline{x}[g(\overline{y}[s/j])/i]) \in \mathcal{U}$ and
$\vdash f(\overline{x}[\lim_k g(\overline{y}[s_k/j])/i]) =_0 f(\overline{x}[g(\overline{y}[s/j])/i]) \in \mathcal{U}$.
(Triang) concludes the proof. ∎

These are easy proofs, the point of including them is to show that standard facts about the continuity of nonexpansive functions can be stated and proved within the framework of quantitative equational logic.

4   Banach Quantitative Theories

In this section we identify a particular class of quantitative equational theories that we will call Banach theories. Later we will see that the Banach theories are the ones for which we can define fixed-point operators.

A *Banach signature* Ω is a signature that assigns to each function symbol f an arity $n \in \mathbb{N}$ and a Banach pattern $\theta \in \mathbb{U}_n$; we write $f:n:\theta$. In particular, for constants $c \in \Omega$, we have that $c:0:\{\langle 0 \rangle\} \in \Omega$.

We extend the concept of Banach pattern from the elements of a Banach signature Ω to all the terms of $\hat{\Omega}X$ by defining, for arbitrary $f(\overline{x}) \in \hat{\Omega}X$ with $\overline{x} \in X^n$ and $f:n:\theta$; and any $g_1(\overline{y}), \ldots, g_n(\overline{y}) \in \hat{\Omega}X$ with $\overline{y} \in X^m$ and $g_i:m:\zeta_i$ for $i \leq n$, the following Banach patterns for contraction and term composition.

1. If $i < j \leq n$ and $h(\overline{x} \setminus j) = f(\overline{x}[x_i/j])$, then

$$h:n-1:\theta[i<j].$$

2. If $h(\overline{y}) = f(g_1(\overline{y}), .., g_n(\overline{y}))$, then

$$h : m : \theta \circ \langle \zeta_1..\zeta_n \rangle.$$

With this definition, we will write $t : n : \theta \in \hat{\Omega}X$ to describe any n-ary term with Banach pattern θ that can be defined in $\hat{\Omega}X$. If, in addition $\theta \rhd i$, we write

$$f : n : \theta \rhd i \in \hat{\Omega}X.$$

The reader might usefully think of these definitions first in the case where the Banach patterns are all singletons, in which case these formulas can be seen as a quantitative analogue of how composition would be defined in operads (multicategories).

Definition 4.1 (Banach closure) *Consider a quantitative equational theory* $\mathcal{U}$ *over a set X of variables and a Banach signature Ω. The* Banach closure *of $\mathcal{U}$ is the smallest quantitative equational theory $\mathcal{U}^B$ that contains $\mathcal{U}$ together with the axiom*

$$\textbf{(1-bound)} \quad \vdash x =_1 y \,,$$

and it is closed under the following rule stated for arbitrary $\epsilon_i \leq 0$ for $i \leq n$.

$$\textbf{(Banach)} \quad \frac{f : n : \theta \in \Omega}{\{x_i =_{\epsilon_i} y_i \mid i \leq n\} \vdash f(x_1 \ldots x_n) =_\delta f(y_1 \ldots y_n)} \,,$$

where $\delta = \max\limits_{\overline{\alpha} \in \theta} \sum\limits_{i \leq n} \alpha_i \epsilon_i$.

Definition 4.2 (Banach theory) *A quantitative equational theory $\mathcal{U}$ over $\hat{\Omega}X$ is a* Banach theory *if Ω is a Banach signature and*

$$\mathcal{U} = \mathcal{U}^B.$$

The following two results guarantee that the way we defined the patterns for composition and contraction respects the Banach rule.

Lemma 4.3 *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega}X$ and $t : n : \theta \in \hat{\Omega}X$. Then, for arbitrary $\epsilon_i \geq 0$ for $i \leq n$,*

$$\{x_i =_{\epsilon_i} y_i \mid i \leq n\} \vdash t(x_1 \ldots x_n) =_\delta t(y_1 \ldots y_n) \in \mathcal{U},$$

where $\delta = \max\limits_{\overline{\alpha} \in \theta} \sum\limits_{i \leq n} \alpha_i \epsilon_i$.

Corollary 4.4 *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega}X$ and $\mathcal{A} = (A, \Omega, d) \in \mathbf{QA}(\mathcal{U})$. Then, for any term $t : n : \theta \in \hat{\Omega}X$ and any $\overline{a}, \overline{b} \in A^n$,*

$$d(t^{\mathcal{A}}(\overline{a}), t^{\mathcal{A}}(\overline{b})) \leq \max\limits_{\overline{\alpha} \in \theta} \sum\limits_{i \leq n} \alpha_i d(a_i, b_i).$$

5   Quantitative Fixed-Point Judgements

In this section we show how one can add fixed-point operators, which are essentially second-order constructions, to any Banach theory.

Definition 5.1 *Let Ω be a Banach signature and X a set of variables. The fixed-point extension of $\hat{\Omega}X$ is the set*

$$\hat{\Omega}^{\mu}X = \bigcup_{i \geq 0} \Omega_i,$$

where Ω_i is defined inductively on $i \geq 0$ as follows:
$\Omega_0 = \hat{\Omega}X,$
$\Omega_{k+1} = \{\mu i.f : (n-1) : \mu i.\theta \mid f : n : \theta \triangleright i \in \Omega_k\}.$

Let $\mathcal{J}(\hat{\Omega}^{\mu}X)$ be the set of judgements on $\hat{\Omega}^{\mu}X$, i.e., judgements involving quantitative equations between terms in $\hat{\Omega}^{\mu}X$. In this way we can speak of quantitative equational theories over $\hat{\Omega}^{\mu}X$, respecting the requirements of Definition 3.1.

Definition 5.2 (Fixed-point extension of Banach theory) *Given a Banach theory $\mathcal{U}$ over $\hat{\Omega}X$, its fixed-point extension $\mathcal{U}^{\mu}$ is the smallest quantitative equational theory over $\hat{\Omega}^{\mu}X$ that contains $\mathcal{U}$ and it is closed under the fixed-point approximation rule (Approx) stated below for arbitrary $t, u \in \hat{\Omega}^{\mu}X$, $\overline{s} \in (\hat{\Omega}^{\mu}X)^n$, and $\epsilon \geq 0$.*

$$\textbf{(Approx)} \quad \frac{t : n : \theta \triangleright i \in \hat{\Omega}^{\mu}X}{u =_{\epsilon} t(\overline{s}[u/i]) \vdash u =_{\frac{\epsilon}{1-a}} (\mu i.t)(\overline{s} \setminus i)} ,$$

where $a = \max\{\alpha_i \mid \overline{\alpha} \in \theta\}$.

Note that since $t : n : \theta \triangleright i$, $a < 1$.

When we take a fixed point, the resulting function of the remaining arguments may not permit further fixed point operations to be performed. The Banach patterns allows us to track exactly when we can and cannot take further fixed points.

Notation. To simplify the presentation in what follows, it is useful to adopt a syntactic convention that will allow us to focus on certain variables in terms with many variables, while treating the rest of them as parameters. If $f(\overline{x}) \in \hat{\Omega}^{\mu}X$ is a function of arity n with free variables $\overline{x} = \langle x_1..x_n \rangle$, and we need to focus on its i-th variable x_i, we write $f(\!(x_i)\!)$. For instance if $s \in \hat{\Omega}^{\mu}X$, $f(\!(s)\!)$ denotes $f(\overline{x}[s/i])$. Similarly, if the focus is on two variables, say x_i, x_j for $i < j \leq n$, we write $f(\!(x_i, x_j)\!)$. We will use this notation in what follows any time there is no danger of confusion. It will allow us to avoid carrying extra variables around in the syntax.

Given a Banach signature Ω and a set X of variables, the concept of iteration of a function on its i-th variables, $i \leq n$, can be introduced for an arbitrary function $f : n \in \hat{\Omega}^{\mu}X$. Let $\overline{x} = \langle x_1 \ldots x_n \rangle \in X^n$ and $s \in \hat{\Omega}^{\mu}X$. We define:

$$[f]_i^1(\overline{x}[s/i]) = f(\overline{x}[s/i]),$$
$$[f]_i^{k+1}(\overline{x}[s/i]) = f(\overline{x}[[f]_i^k(\overline{x}[s/i])/i]).$$

With the previous notation, we can denote the k-th iteration on the i-th variable of f on s by $[f]_i^k(\!(s)\!)$.

We conclude this section with two results regarding fixed-point quantitative theories. The theorem below encodes, in terms of quantitative equational logic, the fact that in a Banach theory (we will see later that they are interpreted in 1-bounded complete metric spaces) the sequence of iterations of a function f on its i-th contractive variable, where the function is contractive, is a Cauchy sequence that has as limit $\mu i.f$. Moreover, and here is the novelty that quantitative setting provides, we can monitor "the speed" of the convergence, and this provides us a powerful tool for building approximation theories.

Theorem 5.3 (Banach) *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega}X$ and $f : n : \theta \rhd i \in \hat{\Omega}^\mu X$. Let $a = \max\{\alpha_i \mid \overline{\alpha} \in \theta\}$. We focus on the i-th variable of f, denoted $f(\!(x_i)\!)$. Then,*
(1). $y =_\epsilon z \vdash [f]_i^k(\!(y)\!) =_{\epsilon a^k} [f]_i^k(\!(z)\!) \in \mathcal{U}^\mu$;
(2). $y =_\epsilon f(\!(y)\!) \vdash [f]_i^k(\!(y)\!) =_{\epsilon a^k \frac{1-a^l}{1-a}} [f]_i^{k+l}(\!(y)\!) \in \mathcal{U}^\mu$;
and for any $s \in \hat{\Omega}^\mu X$ and any $\overline{t} \in (\hat{\Omega}^\mu X)^n$,
(3). $\forall \epsilon > 0 \; \exists k \; \forall m \quad \vdash [f]_i^k(\overline{t}[s/i]) =_\epsilon [f]_i^{k+m}(\overline{t}[s/i]) \in \mathcal{U}^\mu$;
(4). $\forall \epsilon > 0 \; \exists k \; \forall m \quad \vdash [f]_i^{k+m}(\overline{t}[s/i]) =_\epsilon \mu i.f(\overline{t} \setminus i) \in \mathcal{U}^\mu$.

Proof. A consequence of the (Banach) rule is that
$y =_\epsilon z \vdash f(\!(y)\!) =_{\epsilon a} f(\!(z)\!) \in \mathcal{U}^\mu$. We apply this repeatedly to get (1) and use (Triang) to get (2).
To prove (3), we start from $\vdash y =_1 f(\!(y)\!) \in \mathcal{U}^\mu$ which we get from (1-bound) and apply (2) observing that since $a < 1$, $a^k \frac{1-a^l}{1-a}$ can be made arbitrarly small for any l by choosing a sufficiently large k.
For (4) we first observe that from (3) we get that

$$\forall \epsilon > 0, \; \exists k \; \forall m, \quad \vdash [f]_i^{k+m}(\overline{t}[s/i]) =_{\epsilon(1-a)} [f]_i^{k+m+1}(\overline{t}[s/i]) \in \mathcal{U}^\mu.$$

We use this in (Approx) instantiated with $u = [f]_i^{k+m}(\overline{t}[s/i])$ and $\epsilon = \frac{\epsilon}{1-a}$ and we get (4). ■

We can talk about convergent sequences in $\hat{\Omega}^\mu X$, in the same way that we discussed them in quantitative algebras. Note that the previous theorem provides an important limit argument: namely the fixed point is obtained as the limit of the iterates. This is, of course, how the Banach fixed-point theorem is supposed to work. These results show how Banach-style reasoning is internalized in quantitative logic.

Corollary 5.4 *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega}X$ and $f : n : \theta \rhd i \in \hat{\Omega}^\mu X$. Then for any $\overline{t} \in (\hat{\Omega}^\mu X)^n$ and any $s \in \hat{\Omega}^\mu X$, $([f]_i^k(\overline{t}[s/i]))_{k \geq 1}$ is a convergent sequence in $\mathcal{U}^\mu$ and moreover,*

$$\vdash \lim_k [f]_i^k(\overline{t}[s/i]) =_0 \mu i.f(\overline{t} \setminus i) \in \mathcal{U}^\mu.$$

11

The next theorem shows that $\mu i.f$ is indeed the unique parametric fixed point of f in its i-th variable.

Theorem 5.5 (Parametric fixed-point) *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega} X$ and $f : n : \theta \triangleright i \in \hat{\Omega}^\mu X$. Then, for any $s \in \hat{\Omega}^\mu X$ and any $\bar{t} \in (\hat{\Omega}^\mu X)^n$,*
(1). $\vdash \mu i.f(\bar{t} \setminus i) =_0 f(\bar{t}[\mu i.f(\bar{t} \setminus i)/i]) \in \mathcal{U}^\mu$;
(2). $s =_0 f(\bar{t}[s/i]) \vdash s =_0 \mu i.f(\bar{t} \setminus i) \in \mathcal{U}^\mu$.

Proof. Let $a = \max\{\alpha_i \mid \overline{\alpha} \in \theta\}$.
(1). From Theorem 5.3 (4), $\forall \epsilon > 0 \ \exists k \ \forall m$,
$\vdash [f]_i^{k+m-1}(\bar{t}[s/i]) =_{\frac{\epsilon}{2a}} \mu i.f(\bar{t} \setminus i) \in \mathcal{U}^\mu$. And applying Theorem 5.3 (1) to it we get
$\vdash [f]_i^{k+m}(\bar{t}[s/i]) =_{\frac{\epsilon}{2}} f(\bar{t}[\mu i.f(\bar{t} \setminus i)/i]) \in \mathcal{U}^\mu$.
On the other hand, Theoren 5.3 (4) also guarantees that
$\vdash [f]_i^{k+m}(\bar{t}[s/i]) =_{\frac{\epsilon}{2}} \mu i.f(\bar{t} \setminus i) \in \mathcal{U}^\mu$.
We apply (Triang) to the previous two equations and get that for any $\epsilon > 0$,
$\vdash \mu i.f(\bar{t} \setminus i) =_\epsilon f(\bar{t}[\mu i.f(\bar{t} \setminus i)/i]) \in \mathcal{U}^\mu$. Now (Cont) concludes the proof.
(2). Now we instantiate (Approx) with $u = s$ and $\epsilon = 0$. ∎

5.1 Semantics of fixed-point judgements

The fixed-point quantitative theories will be interpreted on quantitative algebras over 1-bounded complete metric spaces.

Let Ω be a Banach signature and $\mathbf{C}(\Omega)$ the category of Ω-quantitative algebras over 1-bounded complete metric spaces. If $\mathcal{U}$ is a quantitative equational theory over $\hat{\Omega} X$, let $\mathbf{C}(\mathcal{U})$ denote the class of models of $\mathcal{U}$ in $\mathbf{C}(\Omega)$.

Let $\mathcal{A} = (A, \Omega, d) \in \mathbf{C}(\Omega)$, $f : A^n \to A$, $a \in A$ and $i \leq n$. We define the sequence of iterations of f on a for its i-th variable, which is the family of functions $[f]_i^k : A^{n-1} \to A$ for $k \in \mathbb{N}$, inductively as follows, where $\overline{x} = \langle x_1..x_n \rangle$ is a sequence of variables

$$[f]_i^1(\overline{x} \setminus i) = f(\overline{x}[a/i]),$$

$$[f]_i^{k+1}(\overline{x} \setminus i) = f(\overline{x}[[f]_i^k(\overline{x} \setminus i)/i]).$$

We know from Banach's fixed-point theorem [Ban22] that if f is contractive in its i-th variable, then the sequence $([f]_i^k)$ is Cauchy and has a unique limit, which can be achieved by iterating f on any element of A. Let $f_i^* : A^{n-1} \to A$ denote this limit; this is a function of the remaining $n - 1$ paramemeters and gives the fixed point in the iterated position.

We will use this fact to interpret any fixed-point term in $\mathcal{A}$. Suppose that $t : n : \theta \triangleright i \in \hat{\Omega}^\mu X$. Then, $t^\mathcal{A} : A^n \to A$ is $\max\{\alpha_i \mid \overline{\alpha} \in \theta\}$-contractive in its i-th variable. Hence, applying Banach Theorem we have that there exists

$$[t^\mathcal{A}]_i^* : A_{n-1} \to A.$$

We use this to interpret $\mu i.t$ in $\mathcal{A}$ by defining

$$(\mu i.t)^\mathcal{A} = [t^\mathcal{A}]_i^*.$$

In this way, all the terms in $\hat{\Omega}^\mu X$ can be interpreted in $\mathcal{A}$. And this allows us to interpret any quantitative equation and any quantitative judgement by canonically extending the usual definition as follows.

Given an assignment $\iota \in \Omega[X|\mathcal{A}]$ and a tuple $\overline{z} = \langle z_1..z_n \rangle \in X^n$, let $\iota(\overline{z}) = \langle \iota(z_1)..\iota(z_n) \rangle$. With this notation, we extend ι canonically, from $\hat{\Omega}X$ to $\hat{\Omega}^\mu X$, by letting for any $\mu i.t(\overline{z} \setminus i) \in \hat{\Omega}^\mu X$,

$$\iota(\mu i.t(\overline{z} \setminus i)) = [t^{\mathcal{A}}]_i^* (\overline{\iota(z \setminus i)}).$$

Definition 5.6 (Satisfaction for fixed-point judgements) *Let Ω be a Banach signature and $\mathcal{A} \in \mathbf{C}(\Omega)$. Let*

$$\{s_i =_{\epsilon_i} t_i \mid i \in I\} \vdash s =_\epsilon t \in \mathcal{J}(\hat{\Omega}^\mu X).$$

Then, for $\iota \in \Omega[X|\mathcal{A}]$, we write

$$\{s_i =_{\epsilon_i} t_i \mid i \in I\} \models_{\mathcal{A},\iota} s =_\epsilon t,$$

if $[\forall i \in I, \ d^{\mathcal{A}}(\iota(t_i), \iota(s_i)) \leq \epsilon_i]$ implies $d^{\mathcal{A}}(\iota(s), \iota(t)) \leq \epsilon$. Similarly, for any $\Gamma \vdash \phi \in \mathcal{J}(\hat{\Omega}^\mu X)$,

$$\Gamma \models_{\mathcal{A}} \phi \ \text{iff} \ \forall \iota \in \Omega[X|\mathcal{A}], \ \Gamma \models_{\mathcal{A},\iota} \phi$$

and for a set $\mathfrak{M} \subseteq \mathbf{C}(\Omega)$,

$$\Gamma \models_{\mathfrak{M}} \phi \ \text{iff} \ \forall \mathcal{A} \in \mathfrak{M}, \ \Gamma \models_{\mathcal{A}} \phi.$$

The next theorem states that for a Banach theory the set of its models coincides with the set of models of its fixed-point extension.

Theorem 5.7 *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega}X$ and $\mathcal{U}^\mu$ its fixed-point extension. Then, for any $\mathcal{A} \in \mathbf{C}(\Omega)$,*

$$\mathcal{A} \models \mathcal{U} \ \text{iff} \ \mathcal{A} \models \mathcal{U}^\mu.$$

Proof. The right-to-left implication follows from $\mathcal{U} \subseteq \mathcal{U}^\mu$. We prove the left-to-right implication as follows. It is sufficient to demonstrate that any $\mathcal{A} \in \mathbf{C}(\Omega)$ satisfies (Approx).
Let $f : n : \theta \triangleright i \in \hat{\Omega}^\mu X$ and $a = \max\{\alpha_i \mid \overline{\alpha} \in \theta\}$. We need to prove that for any $\overline{t} \in (\hat{\Omega}^\mu X)^n$, any $s \in \hat{\Omega}^\mu X$ and any $\epsilon \geq 0$, $s =_\epsilon f(\overline{t}[s/i]) \models_{\mathcal{A}} s =_{\frac{\epsilon}{1-a}} \mu i.f(\overline{t} \setminus i)$.
Let $\iota \in \Omega[X|\mathcal{A}]$ and let $\sigma = \iota(s)$ and $\overline{\tau} = \iota(\overline{t})$.
Suppose that $d(\sigma, f(\overline{\tau}[\sigma/i])) \leq \epsilon$. Let $m = \mu i.f(\overline{\tau} \setminus i)$. Then, $m = f(\overline{\tau}[m/i])$. We have
$d(\sigma, m) \leq d(\sigma, f(\overline{\tau}[\sigma/i])) + d(f(\overline{\tau}[\sigma/i]), m)$
$= d(\sigma, f(\overline{\tau}[\sigma/i])) + d(f(\overline{\tau}[\sigma/i]), f(\overline{\tau}[m/i]))$
$\leq \epsilon + ad(\sigma, m)$. Hence, $d(\sigma, m) \leq \frac{\epsilon}{1-a}$ implying that
$s =_\epsilon f(\overline{t}[s/i]) \models_{\mathcal{A},\iota} s =_{\frac{\epsilon}{1-a}} \mu i.f(\overline{t} \setminus i)$. This concludes the proof. ∎

Hence, the class of models of $\mathcal{U}$ and of $\mathcal{U}^\mu$ coincide in the category of 1-bounded complete metric spaces.

Remark 5.8 *Note that all the terms in $\hat{\Omega}^\mu X$ are nonexpansive in all their variables and with well-defined Banach patterns. Consequently, we can think of theses terms as elements of a larger Banach signature $\hat{\Omega}^\mu$ that contains all the terms as function symbols, and this is a "legal" quantitative algebra signature. Similarly, one can think of the fixed-point extension $\mathcal{U}^\mu$ of a quantitative equational theory $\mathcal{U}$ over a Banach signature as a quantitative equational theory over the signature $\hat{\Omega}^\mu X$ as originally defined in [MPP16].*

This remark together with the result of Theorem 5.7 allows us to conclude this section with a completeness result.

Theorem 5.9 (Completeness for fixed-point theories) *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega} X$ and $\mathcal{U}^\mu$ its fixed-point extension. Then for any fixed-point quantitative judgement $\Gamma \vdash \phi \in \mathcal{J}(\hat{\Omega}^\mu X)$,*

$$\Gamma \models_{\mathbf{C}(\mathcal{U})} \phi \text{ iff } \Gamma \vdash \phi \in \mathcal{U}^\mu.$$

Proof. Following the Remark 5.8, $\hat{\Omega}^\mu$ is a Banach signature and $\mathcal{U}^\mu$ is a quantitative equational theory over $\hat{\Omega}^\mu X$. From the completeness result for quantitative algebras, stated in 3.4 and proven in [MPP16], we get that

$$\Gamma \models_{\mathbf{C}(\mathcal{U}^\mu)} \phi \text{ iff } \Gamma \vdash \phi \in \mathcal{U}^\mu.$$

Applying Theorem 5.7, which asserts that $\mathbf{C}(\mathcal{U}) = \mathbf{C}(\mathcal{U}^\mu)$, we conclude the proof. ∎

6 Quantitative Fixed-Point Theories

In this section we investigate the relation between the fixed-point extension of Banach theories and the traditional concepts of Conway theories and iteration theories [BE93, SP00].

Notation. In what follows, for a term $f : n : \theta \triangleright i$ and a sequence $\overline{x} = \langle x_1..x_n \rangle$ of variables, we will also use the usual variable-binding fixed-point syntax and write $\mu x_i.f(\overline{x})$ to denote $\mu i.f(\overline{x} \setminus i)$. This notation allows us to present a series of results in a more familiar format.

6.1 Quantitative Conway theories

The Conway theories [BE93, SP00], are defined by two properties Dinaturality and Diagonal property. We prove here that quantitative versions of these can be proven in any fixed-point Banach theory.

Lemma 6.1 *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega} X$ and let $f : p : \theta \triangleright i$, $g : q : \zeta \triangleright j \in \hat{\Omega}^\mu X$. We focus on the i-th variable of f, $f(\!(x_i)\!)$ and on the j-th variable of g, $g(\!(y_j)\!)$.*
Then, for any $s \in \hat{\Omega}^\mu X$ we have that $\forall \epsilon > 0 \; \exists n \; \forall m$,

$$\vdash [f(\!(g(\!(y_j)\!))\!)]_{i+j-1}^{n+m}(s) =_\epsilon f(\!([g(\!(f(\!(x_i)\!))\!)]_{i+j-1}^{n+m}(s))\!) \in \mathcal{U}^\mu.$$

14

Proof. Let $a = \max\{\alpha_i \mid \overline{\alpha} \in \theta\}$, $b = \max\{\beta_j \mid \overline{\beta} \in \zeta\}$. From (1-bound) we have $\vdash s =_1 f(\!|s|\!) \in \mathcal{U}^\mu$ and applying (Banach) to this, we conclude $\vdash g(\!|s|\!) =_b g(\!|f(\!|s|\!)|\!) \in \mathcal{U}^\mu$. We again apply (Banach) to this last equation and get $\vdash f(\!|g(\!|s|\!)|\!) =_{ab} f(\!|g(\!|f(\!|s|\!)|\!)|\!) \in \mathcal{U}^\mu$. Hence,
$\vdash [f(\!|g(\!|y_j|\!)|\!)]^1_{i+j-1}(s) =_{ab} f(\!|[g(\!|f(\!|x_i|\!)|\!)]^1_{i+j-1}(s)|\!) \in \mathcal{U}^\mu$. Repeating these steps we obtain
$\vdash [f(\!|g(\!|y_j|\!)|\!)]^n_{i+j-1}(s) =_{(ab)^n} f(\!|[g(\!|f(\!|x_i|\!)|\!)]^n_{i+j-1}(s)|\!) \in \mathcal{U}^\mu$.
Since $ab < 1$, we can make $(ab)^n$ as small as we want. ∎

This lemma allows us to prove a quantitative version of the Dinaturality property [SP00].

Theorem 6.2 (Quantitative Dinaturality) *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega}X$ and $f : n : \theta$, $g : m : \zeta \in \hat{\Omega}^\mu X$ such that $\theta[\zeta/i]$ and $\zeta[\theta/j]$ are $i + j - 1$-contractive. We focus on the i-th variable of f, $f(\!|x_i|\!)$, and on j-th variable of g, $g(\!|y_j|\!)$. Then,*

$$\vdash \mu y_j.f(\!|g(\!|y_j|\!)|\!) =_0 f(\!|\mu x_i.g(\!|f(\!|x_i|\!)|\!)|\!) \in \mathcal{U}^\mu.$$

Proof. Let $a = \max\{\alpha_i \mid \overline{\alpha} \in \theta\}$.
By using Lemma 6.1 and Theorem 5.3 (4) together, we obtain $\forall \epsilon > 0 \; \exists n \; \forall m$ such that the following three statements are satisfied.
$\vdash [f(\!|g(\!|y_j|\!)|\!)]^{n+m}_{i+j-1}(s) =_{\frac{\epsilon}{3}} f(\!|[g(\!|f(\!|x_i|\!)|\!)]^{n+m}_{i+j-1}(s)|\!) \in \mathcal{U}^\mu$,
$\vdash [f(\!|g(\!|y_j|\!)|\!)]^{m+n}_{i+j-1}(s) =_{\frac{\epsilon}{3}} \mu y_j.f(\!|g(\!|y_j|\!)|\!) \in \mathcal{U}^\mu$ and
$\vdash [g(\!|f(\!|x_i|\!)|\!)]^{m+n}_{i+j-1}(s) =_{\frac{\epsilon}{3a}} \mu x_i.g(\!|f(\!|x_i|\!)|\!) \in \mathcal{U}^\mu$.
The last one implies
$\vdash f(\!|[g(\!|f(\!|x_i|\!)|\!)]^{m+n}_{i+j-1}(s)|\!) =_{\frac{\epsilon}{3}} f(\!|\mu x_i.g(\!|f(\!|x_i|\!)|\!)|\!) \in \mathcal{U}^\mu$.
Using this one and the first two with (Triang) we conclude that for any $\epsilon > 0$,

$$\vdash \mu y_j.f(\!|g(\!|y_j|\!)|\!) =_\epsilon f(\!|\mu x_i.g(\!|f(\!|x_i|\!)|\!)|\!) \in \mathcal{U}^\mu.$$

Now we apply (Cont) and complete the proof. ∎

This type of "$\epsilon/3$-argument" is common in analysis.

With these results in hand we can proceed and prove a quantitative version of the diagonal property for fixed-point theories.

Theorem 6.3 (Quantitative Diagonal property) *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega}X$ and $f : n : \theta \in \hat{\Omega}^\mu X$. Suppose there exists $i < j \leq n$ s.t. for any $\overline{\alpha} \in \theta$, $\alpha_i + \alpha_j < 1$. We focus on the i-th and j-th variables of f, $f(\!|x_i, x_j|\!)$. Then,*
$$\vdash \mu x_i.f(\!|x_i, x_i|\!) =_0 \mu x_j.\mu x_i.f(\!|x_i, x_j|\!) \in \mathcal{U}^\mu.$$

Proof. Let $s = \mu x_i.f(\!|x_i, x_i|\!)$ and $t(\!|x_i|\!) = \mu x_j.f(\!|x_i, x_j|\!)$. Theorem 5.5 (1) guarantees that
$\vdash s =_0 t(\!|s, s|\!) \in \mathcal{U}^\mu$ and $\vdash t(\!|x_i|\!) =_0 f(\!|x_i, t(\!|x_i|\!)|\!) \in \mathcal{U}^\mu$.
Let $a = \max\{\alpha_i, \alpha_j \mid \overline{\alpha} \in \theta\}$. Then applying (Banach),
$x_i =_\epsilon s \vdash f(\!|x_i, t(\!|x_i|\!)|\!) =_{a\epsilon} f(\!|s, t(\!|x_i|\!)|\!) \in \mathcal{U}^\mu$ and
$t(\!|x_i|\!) =_\delta s \vdash f(\!|s, t(\!|x_i|\!)|\!) =_{a\delta} f(\!|s, s|\!) \in \mathcal{U}^\mu$. We apply (Triang) and obtain

15

$\{x_i =_\epsilon s, \ t(\!|x_i|\!) =_\delta s\} \vdash f(\!|x_i, t(\!|x_i|\!)|\!) =_{a(\epsilon+\delta)} f(\!|s, s|\!) \in \mathcal{U}^\mu$. Now if we instantiate this with $x_i = s$, we get

$t(\!|s|\!) =_\delta s \vdash f(\!|s, t(\!|s|\!)|\!) =_{a\delta} f(\!|s, s|\!) \in \mathcal{U}^\mu$. We already know that

$\vdash s =_0 t(\!|s, s|\!) \in \mathcal{U}^\mu$ and $\vdash t(\!|x_i|\!) =_0 f(\!|x_i, t(\!|x_i|\!)|\!) \in \mathcal{U}^\mu$. Combining these three, we obtain

$t(\!|s|\!) =_\delta s \vdash t(\!|s|\!) =_{a\delta} s \in \mathcal{U}^\mu$. By applying this repeatedly and eventually using (Cont), we obtain

$$\vdash t(\!|s|\!) =_0 s \in \mathcal{U}^\mu,$$

which is the desired result. ∎

Theorems 6.2 and 6.3 establish that the fixed-point extension of any Banach theory is a Conway theory, in the sense of [SP00].

6.2 Quantitative Iteration Theories

In this subsection we show that the fixed-point Banach theories are not only Conway theories, but they are iteration theories in the sense of [SP00]; meaning that, in addition to quantitative dinaturality and the quantitative diagonal property, they also satisfy a quantitative version of the amalgamation property.

Lemma 6.4 *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega}X$ and $f : n : \theta, g : m : \zeta \in \hat{\Omega}^\mu X$ such that there exist $i < j \leq n$ and $u < v \leq m$ with $a_i(1 - b_v) + b_u < 1$ and $a_i + 2a_j b_v < 1$, where $a_i = \max\{\alpha_i \mid \overline{\alpha} \in \theta\}$, $a_j = \max\{\alpha_j \mid \overline{\alpha} \in \theta\}$, $b_u = \max\{\beta_u \mid \overline{\beta} \in \zeta\}$ and $b_v = \max\{\beta_v \mid \overline{\beta} \in \zeta\}$. We focus on the variables i and j in f, $f(\!|x_i, x_j|\!)$ and on variables u and v in g, $g(\!|y_u, y_v|\!)$. Then,*

$$\vdash \mu x.f(\!|x, \mu y.g(\!|x, y|\!)|\!) =_0 \mu x.\mu y.f(\!|x, g(\!|x, y|\!)|\!) \in \mathcal{U}^\mu.$$

Proof. The inequalities $a_i(1 - b_v) + b_u < 1$ and $a_i + 2a_j b_v < 1$ guarantee that the fixed-points are properly defined.

Observe now that by repeatedly applying 3.7 (2) we can prove that for any two sequences $(s_k)_{k \geq 1} \subseteq \hat{\Omega}^\mu X$ and $(t_r)_{r \geq 1} \subseteq \hat{\Omega}^\mu X$ convergent in $\mathcal{U}^\mu$, $(\lim_r f(\!|s_k, g(\!|s_k, t_r|\!)|\!))_{k \geq 1}$ and $(f(\!|s_k, \lim_r g(\!|s_k, t_r|\!)|\!))_{k \geq 1}$ are convergent in $\mathcal{U}^\mu$ and moreover,

$$\vdash \lim_k f(\!|s_k, \lim_r g(\!|s_k, t_r|\!)|\!) =_0 \lim_k (\lim_r f(\!|s_k, g(\!|s_k, t_r|\!)|\!)) \in \mathcal{U}^\mu.$$

Applying this in the context of Corollary 5.4, we get the desired result. ∎

With the result of the previous lemma, we are ready to state the quantitative amalgamation theorem.

Theorem 6.5 (Quantitative amalgamation) *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega}X$ and let $f_i : n : \theta^i \in \hat{\Omega}^\mu X$ for $i \leq n$ be a family of functions such that for any $i \neq j$ and any $\overline{\alpha}^i \in \theta^i$ and $\overline{\alpha}^j \in \theta^j$,*

$$\sum_{k \leq n} \alpha_k^i = \sum_{k \leq n} \alpha_k^j = \alpha < 1.$$

Suppose there exists $g : 1 : \{\langle \alpha \rangle\} \in \hat{\Omega}^\mu X$ s.t. for all $i \leq n$,

$$\vdash f_i(x..x) =_0 g(x) \in \mathcal{U}^\mu.$$

If there exists $s_1, .., s_n \in \hat{\Omega}^\mu X$ s.t. for all $i \leq n$,

$$\vdash s_i =_0 f_i(s_1..s_n) \in \mathcal{U}^\mu,$$

then for all $i \leq n$,

$$\vdash s_i =_0 \mu x.g(x) \in \mathcal{U}^\mu.$$

Proof. We only sketch the proof for the case $n = 2$ that is simpler to present. The general case is proven in the same way, but one needs to keep track of more indices.

From the hypothesis,
$\vdash s_1 =_0 f(s_1, s_2) \in \mathcal{U}^\mu$ and $\vdash s_2 = f_2(s_1, s_2) \in \mathcal{U}^\mu$. Let t denote $\mu x.g(x)$. From Theorem 5.5 we know that $\vdash t =_0 g(t) \in \mathcal{U}^\mu$.
We will prove that for any i, $\vdash s_i =_0 t \in \mathcal{U}^\mu$.

Let $\phi_1(x_2) = \mu x_1.f_1(x_1, x_2)$ and $\phi_2(x_1) = \mu x_2.f_2(x_1, x_2)$. From the hypothesis we have $\vdash \phi_1(s_2) =_0 s_1 \in \mathcal{U}^\mu$ and $\vdash \phi_2(s_1) =_0 s_2 \in \mathcal{U}^\mu$. Consequently, $\vdash s_1 = \phi_1(\phi_2(s_1)) \in \mathcal{U}^\mu$. Applying Theorem 5.5, we get then
$\vdash s_1 =_0 \mu z.\phi_1(\phi_2(z)) \in \mathcal{U}^\mu$. By extending ϕ_2 we get further $\vdash s_1 =_0 \mu z.\phi_1(\mu x_2.f_2(z, x_2)) \in \mathcal{U}^\mu$, and after extending ϕ_1 we get
$\vdash s_1 =_0 \mu z.\mu x_1.f_1(x_1, \mu x_2.f_2(z, x_2)) \in \mathcal{U}^\mu$. We use the quantitative diagonal property and get
$\vdash s_1 =_0 \mu v.f_1(v, \mu x_2.f_2(v, x_2)) \in \mathcal{U}^\mu$. Now we apply Lemma 6.4 to obtain
$\vdash s_1 =_0 \mu v.\mu x_2.f_1(v, f_2(v, x_2)) \in \mathcal{U}^\mu$.
Lemma 6.4 also gives us[2]
$\vdash s_1 =_0 \mu w.f_1(w, f_2(w, w)) \in \mathcal{U}^\mu$.
From the hypothesis we know that $\vdash t =_0 f_1(t, t) \in \mathcal{U}^\mu$ and $\vdash t =_0 f_2(t, t) \in \mathcal{U}^\mu$. Hence, $\vdash t =_0 f_1(t, f_2(t, t)) \in \mathcal{U}^\mu$. Next Theorem 5.5 guarantees that
$\vdash t =_0 \mu w.f_1(w, f_2(w, w)) \in \mathcal{U}^\mu$. Combining this with the previous fixed-point description that we derived for s_1, we get $\vdash s_1 =_0 t \in \mathcal{U}^\mu$. Similarly one can prove $\vdash s_2 =_0 t \in \mathcal{U}^\mu$. ∎

Note that all three theorems have a top-level statement that is stated in terms of exact equality but the proofs use approximate equality. In the amalgamation proof the approximate reasoning is isolated into Lemma 6.4. In addition to Theorems 6.2 and 6.3, Theorem 6.5 guarantees that any fixed-point extension of a Banach theory is an iteration theory as defined in [SP00].

7   The metric coinduction principle

In this section we will investigate the relation of these theories with a very interesting and useful coinduction principle proposed by Kozen in [Koz06, Koz07].

---

[2]The same sequence of operations can be used n times if the arity of f is n and get a fixed-point as the one we get here.

We will demonstrate that the metric coinduction principle can be proven within any Banach fixed-point theorem and that this principle is equivalent to our rule (Approx). That being said, however, if we restrict ourselves to *finitary* proofs, we suspect that the metric coinduction principle is more powerful.

The context in which the metric coinduction principle is stated in [Koz06, Koz07] is a bit more liberal than the syntax of fixed-point Banach theories, as it involves the concept of closed predicate, defined as a predicate whose extension is a closed set in any bounded complete metric space. For this reason, we will work in this section at a metalevel, where semantics concepts, *i.e.* metric and topological concepts, are used together with the syntax of Banach theories.

Consider a Banach theory $\mathcal{U}$ over $\hat{\Omega}X$ and its fixed-point extension $\mathcal{U}^\mu$. A *closed predicate* in this context is any predicate P whose extension, when interpreted in any model in $\mathbf{C}(\mathcal{U})$, is a closed set in the open-ball topology induced by the metric. In this setting, the metric coinduction principle for the closed predicate P is stated as follows, for any map $f : n : \theta \rhd i \in \hat{\Omega}^\mu X$, any $\overline{y} \in X^n$ and an arbitrary $t \in \hat{\Omega}^\mu X$.

$$\textbf{(MCoind)} \quad \frac{\vdash P(t) \quad P(x) \vdash P(f(\overline{y}[x/i]))}{\vdash P(\mu i.f(\overline{y}))}.$$

Given a Banach theory $\mathcal{U}$ over $\hat{\Omega}X$, let $\mathcal{U}^M$ be the smallest extension of $\mathcal{U}$ over $\hat{\Omega}^\mu X$ that is closed under the metric coinduction principle (MCoind) - we call it the coinductive extension of $\mathcal{U}$.

The next two theorems will relate $\mathcal{U}^M$ and $\mathcal{U}^\mu$.

Theorem 7.1 *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega}X$, and let $\mathcal{U}^M$ and $\mathcal{U}^\mu$ be its coinductive extension and fixed-point extension respectively. Then*

$$\mathcal{U}^\mu \subseteq \mathcal{U}^M.$$

Proof. To prove this result it is sufficient to show that $\mathcal{U}^M$ is closed under the rule (Approx).

For simplicity, we focus on the i-th variable of f, $f(\!|x_i|\!)$ and let $a = \max\{\alpha_i \mid \overline{\alpha} \in \theta\}$.

Consider the predicate

$$R(y) = \forall x.(x =_\epsilon f(\!|x|\!) \vdash x =_{\frac{\epsilon}{1-a}} y \in \mathcal{U}^M).$$

and let $\overline{B_\epsilon(x)} = \{z \in \hat{\Omega}^\mu X \mid \ \vdash x =_\epsilon z\}$, which is interpreted in any model as the ϵ-closed ball centred at x.
Then we can characterize R as follows

$$R(y) = \forall x(f(\!|x|\!) \in \overline{B_\epsilon(x)} \Rightarrow y \in \overline{B_{\frac{\epsilon}{1-a}}(x)} \ )$$

$$= y \in \bigcap_{z \in \{x \mid \ f(\!|x|\!) \in \overline{B_\epsilon(x)}\}} \overline{B_{\frac{\epsilon}{1-a}(x)}}.$$

Hence, R is a closed predicate and we can use it to instantiate (MCoind) and conclude that $\mathcal{U}^M$ is closed under the rule

$$\frac{\vdash R(t) \quad R(x) \vdash R(f(\!|x|\!))}{\vdash R(\mu x.f(\!|x|\!))}.$$

We prove now that for any x, $R(x) \vdash R(f(\!|x|\!)) \in \mathcal{U}^M$ which is equivalent to proving that

$$\forall z [\forall x (x =_\epsilon f(\!|x|\!) \vdash x =_{\frac{\epsilon}{1-a}} z) \in \mathcal{U}^M$$

$$\Rightarrow \forall x (x =_\epsilon f(\!|x|\!) \vdash x =_{\frac{\epsilon}{1-a}} f(\!|x|\!)) \in \mathcal{U}^M].$$

Suppose that for any x, $x =_\epsilon f(\!|x|\!) \vdash x =_{\frac{\epsilon}{1-a}} z \in \mathcal{U}^M$. Since f is contractive, (Banach) guarantees that
$x =_\epsilon y \vdash f(\!|x|\!) =_{a\epsilon} f(\!|y|\!) \in \mathcal{U}^M$. Hence,
$x =_\epsilon f(\!|x|\!) \vdash f(\!|x|\!) =_{\epsilon \frac{a}{1-a}} f(\!|z|\!) \in \mathcal{U}^M$. Next (Triang) proofs
$x =_\epsilon f(\!|x|\!) \vdash x =_{\frac{\epsilon}{1-a}} f(\!|z|\!) \in \mathcal{U}^M$, hence for any x, $R(x) \vdash R(f(\!|x|\!)) \in \mathcal{U}^M$.

Now it is not difficult to notice that $x =_\epsilon f(\!|x|\!) \vdash x =_{\epsilon \frac{a^n}{1-a}} [f]_i^n(\!|s|\!)$ - Theorem 5.3.

So, since (1-bound) guarantees that for any $s \in \hat{\Omega}^\mu X$, $\vdash s =_1 f(\!|s|\!) \in \mathcal{U}^M$, we get that the sequence $([f]_i^k(\!|s|\!))_{k \geq 1}$ is convergent in $\mathcal{U}^M$ and its limit t is such that $\vdash R(t) \in \mathcal{U}^M$.

Hence both hypothesis of (MCoind) for R are satisfied, meaning that its conclusion has to be true, which is

$$\vdash R(\mu i.f) \in \mathcal{U}^M,$$

but this is exactly (Approx). ■

The next theorem says that whenever we have a closed predicate, any consequences proved using (Mcoind) with this predicate can be proved in $\mathcal{U}^\mu$.

Theorem 7.2 *Let $\mathcal{U}$ be a Banach theory over $\hat{\Omega}X$, let $\mathcal{U}^\mu$ be its fixed-point extension and let P be a closed predicate. Then any consequences of P obtained using (Mcoind) can be established in $\mathcal{U}^\mu$.*

Proof. Let P be a closed predicate. Then it must be the complement of an open predicate B, i.e.,
$$P = B^c.$$

Let
$$B_\epsilon(x) = \{y \in \hat{\Omega}^\mu X \mid \ \vdash x =_\delta y \in \mathcal{U}^\mu \text{ for some } \delta < \epsilon\},$$

be the x-centred open ball of radius $\epsilon > 0$. These sets for a base in the open ball topology, hence there must exist a set I of indices and a set of I-indexed terms $s_i \in \hat{\Omega}^\mu X$ such that
$$B = \bigcup_{i \in I} B_{\epsilon_i}(s_i).$$

Consequently,
$$P = \bigcap_{i \in I} B_{\epsilon_i}^c(s_i).$$

19

Now we have
$$B_\epsilon^c(x) = \{y \mid\Vdash x =_\delta y \in \mathcal{U}^\mu \Rightarrow \delta \geq \epsilon\}.$$

Hence,
$$P = \{y \mid \forall i \in I, \vdash s_i =_\delta y \in \mathcal{U}^\mu \Rightarrow \delta \geq \epsilon_i\}.$$

Hence we can define any closed predicate P as

$$P(x) = \forall i \in I (\vdash s_i =_\delta x \in \mathcal{U}^\mu \Rightarrow \delta \geq \epsilon_i).$$

Now we prove that $\mathcal{U}^\mu$ is closed under (MCoind) for P.
Suppose that for some $s \in \hat{\Omega}^\mu X$, $\vdash P(s) \in \mathcal{U}^\mu$, and that $P(x) \vdash P(f(\!|x|\!)) \in \mathcal{U}^\mu$.
The second one means

$$\forall x [\forall i (\vdash x =_\delta s_i \in \mathcal{U}^\mu \Rightarrow \delta \geq \epsilon_i)$$

$$\Rightarrow \forall i (\vdash f(\!|x|\!) =_\delta s_i \in \mathcal{U}^\mu \Rightarrow \delta \geq \epsilon_i)].$$

Iterating this over
$$\forall i \in I (\vdash s_i =_\delta s \in \mathcal{U}^\mu \Rightarrow \delta \geq \epsilon_i),$$

which is an equivalent statement for $\vdash P(s) \in \mathcal{U}^\mu$, we get

$$\forall k \; \forall i \; [\vdash [f]_i^k (\!|s|\!) =_\delta s_i \Rightarrow \delta \geq \epsilon_i] (*).$$

We need to prove that

$$\forall i [\vdash s_i =_\delta \mu x.f(\!|x|\!) \Rightarrow \delta \geq \epsilon_i].$$

Suppose this is not the case and there exists some $j \in I$ so that for some $r > 0$,

$$\vdash \mu x.f(\!|x|\!) =_r s_j \in \mathcal{U}^\mu \ \wedge \ r < \epsilon_j.$$

We know from Corollary 5.4 that for any $0 < p < \epsilon_j - r$ there exists some k s.t.

$$\vdash \mu x.f(\!|x|\!) =_{\epsilon_j - r - p} [f]_i^k (\!|s_j|\!) \in \mathcal{U}^\mu.$$

Finally (Triang) gives us

$$\vdash [f]_i^k (\!|s_j|\!) =_{\epsilon_j - p} s_j \in \mathcal{U}^\mu,$$

but this contradicts the statement $(*)$ above since $\epsilon_j - p < \epsilon_j$. ∎

The results stated in Theorems 7.1 and 7.2 show that the metric coinduction principle, despite its more semantic flavour and its quantification over all closed predicates, has the same power as our fixed point Banach theories. However, it is often easier to use and is a very attractive proof principle.

8 Markov Decision Processes and the Bellman equation

Markov decision processes [Put94] are a well known formalism used in operations research and extensively in reinforcement learning [SB98]. The Bellman equation is perhaps the most common application of the Banach fixed-point theorem. This section is an extended example showing how one can reason about the Bellman equation in our setting. Indeed this research project began from a desire to treat the Bellman equation as an example within the quantitative equational logic framework before we developed the general theory reported here.

8.1 Markov decision processes

Definition 8.1 *A Markov decision process is a tuple*

$$\mathcal{M} = (S, A, (P^a)_{a \in A}, (R^a)_{a \in A})$$

where

- *S is a finite set of* states*; let ΔS represent the set of probability distributions on S.*

- *A is a finite set of* actions*; let ΔA represent the set of probability distributions on A.*

- *For each $a \in A$, $P^a : S \longrightarrow \Delta(S)$ are the labelled probabilistic* transitions.

- *For each $a \in A$, $R^a : S \longrightarrow [0,1]$ is the* reward *function.*

One can think of these as transition systems where an external agent controls the system choosing actions according to some policy. The system responds by changing state according to the transition function and returning a reward. The reward is accumulated, with a multiplicative discount factor, and the goal of reinforcement learning is to find the best policy for optimizing the reward.

The effectiveness of a particular policy is captured by what are called *value functions* which summarize the aggregated discounted rewards associated with a policy. Mathematically, value functions are elements of the space $\mathcal{V} = [0,1]^S$; this is a metric space endowed with the metric

$$d(f, g) = \max_{s \in S} |f(s) - g(s)|.$$

A *policy* is a map $\pi : S \longrightarrow \Delta A$ that associates to each state a probability distribution over the actions. Let Π denote the set of policies for $\mathcal{M}$. For arbitrary $a \in A$ we write $\hat{a}$ for the constant policy that associates to any state the Dirac distribution concentrated at a.

For an arbitrary policy $\pi \in \Pi$, the expected immediate reward of π is the value function $R^\pi \in \mathcal{V}$ defined for arbitrary $s \in S$, by

$$R^\pi(s) = \sum_{a \in A} \pi(s)(a) R^a(s).$$

Given a policy $\pi \in \Pi$ and a discount factor $\gamma \in (0,1)$, the Bellman operator of π is the operator $T^\pi : \mathcal{V} \to \mathcal{V}$ defined for arbitrary $f \in \mathcal{V}$ and $s \in S$ as follows

$$T^\pi(f)(s) = (1-\gamma)R^\pi(s) + \gamma \sum_{a \in A} \sum_{s' \in S} \pi(s)(a)P^a(s)(s')f(s')$$

The Bellman equation for the policy $\pi \in \Pi$ and discount factor $\gamma \in (0,1)$ is the following fixed point equation over $\mathcal{V}$

$$X = T^\pi(X).$$

The discount factor makes this operator contractive and thus has a unique fixed point: this is the value function of the policy π.

8.2 Reward Barycentric Algebra

Assumptions For the rest of this section, we assume a fixed Markov decision process $\mathcal{M} = (S, A, (P^a)_{a \in A}, (R^a)_{a \in A})$ and a fixed discount factor $\gamma \in (0,1)$.

We develop a particular Banach theory, designed for solving the Bellman equation for $\mathcal{M}$ and γ. Its signature extends the barycentric signature and the theory extends the quantitative barycentric theory developed in [MPP16]. The models of our theory will be called reward barycentric algebras (RBA), and will be a specialised class of barycentric algebras, as defined in [MPP16], devised with additional algebraic structure.

Signature. Consider the Banach signature Σ containing the following basic operators.

- For each $\epsilon \in [0,1]$, $+_\epsilon : 2 : \{\langle \epsilon, 1-\epsilon \rangle\} \in \Sigma$;

- For each $\pi \in \Pi$, $\langle \pi \rangle : 1 : \{\langle 1 \rangle\} \in \Sigma$;

- For each $\pi \in \Pi$, $|\pi| : 1 : \{\langle \gamma \rangle\} \in \Sigma$.

Consider now the Banach theory $\mathcal{B}$ over $\hat{\Sigma}X$ axiomatized by the following two sets of axioms

Barycentric axioms:
for arbitrary $\epsilon, \epsilon' \in [0,1]$, $p, q \in \mathbb{R}_+$, $x, x', y, y' \in X$

(B1) $\vdash x +_1 x' =_0 x$

(B2) $\vdash x +_\epsilon x =_0 x$

(SC) $\vdash x +_\epsilon x' =_0 x' +_{1-\epsilon} x$

(SA) $\vdash (x +_\epsilon x') +_{\epsilon'} y =_0 x +_{\epsilon\epsilon'} (x' +_{\frac{\epsilon'-\epsilon\epsilon'}{1-\epsilon\epsilon'}} y)$ for $\epsilon\epsilon' < 1$

(BA) $\{x =_p x', y =_q y'\} \vdash x +_\epsilon x' =_{\epsilon p + (1-\epsilon)q} y +_\epsilon y'$

22

Reward axioms:

for arbitrary $\pi, \pi' \in \Pi$, $\epsilon \in [0,1]$ and $x, y \in X$

(R1) $\vdash \langle \epsilon\pi + (1-\epsilon)\pi'\rangle x =_0 \langle\pi\rangle x +_\epsilon \langle\pi'\rangle x$

(R2) $\vdash |\epsilon\pi + (1-\epsilon)\pi'|x =_0 |\pi|x +_\epsilon |\pi'|x$

(R3) $x =_\epsilon y \vdash |\pi|x =_{\gamma\epsilon} |\pi|y$

Algebra of value functions. The space $(\mathcal{V}, d)$ of value functions of $\mathcal{M}$ is a 1-bounded complete metric space and has a natural σ-algebra of Borel sets. We interpret the basic functions in Σ, for arbitrary $f, g \in \mathcal{V}$, $\pi \in \Pi$ and $s \in S$ as follows

- $(f +_\epsilon g)^{\mathcal{V}} = \epsilon f^{\mathcal{V}} + (1-\epsilon)g^{\mathcal{V}}$

- $(\langle\pi\rangle f)^{\mathcal{V}}(s) = \sum_{a \in A} \pi(s)(a) \sum_{s' \in S} P^a(s)(s')f^{\mathcal{V}}(s')$

- $(|\pi|f)^{\mathcal{V}} = (1-\gamma)R^\pi + \gamma f^{\mathcal{V}}$

It is not difficult to verify that the functions have indeed the expected Banach patterns, hence $\mathcal{V}$ with the previous interpretation is indeed an algebra of the right form. Consider now $\hat{\Sigma}^\mu X$ the fixed-point extension of $\hat{\Sigma}X$.

For simplicity, in what follows we denote the interpretation of any $t \in \hat{\Sigma}^\mu X$ in $\mathcal{V}$ by $[\![t]\!]$. We can now prove that $\mathcal{V}$ satisfies indeed the axioms of $\mathcal{B}$.

Theorem 8.2 *The space $\mathcal{V}$ of value functions of $\mathcal{M}$ is a model for $\mathcal{B}$, $\mathcal{V} \models \mathcal{B}$.*

Proof. The fact that the Barycentric axioms are satisfied by $\mathcal{V}$ is already proven in [MPP16]. We prove here the soundness of the reward axioms.

(R1): for any $t \in \hat{\Sigma}^\mu X$,
$[\![\langle\epsilon\pi + (1-\epsilon)\pi'\rangle t]\!](s)$
$= \sum_{a \in A}(\epsilon\pi(s)(a) + (1-\epsilon)\pi'(s)(a)) \sum_{s' \in S} P^a(s)(s')[\![t]\!](s')$
$= \epsilon \sum_{a \in A} \pi(s)(a) \sum_{s' \in S} P^a(s)(s')[\![t]\!](s') +$
$+ (1-\epsilon) \sum_{a \in A} \pi'(s)(a) \sum_{s' \in S} P^a(s)(s')[\![t]\!](s')$
$= \epsilon[\![\langle\pi\rangle t]\!](s) + (1-\epsilon)[\![\langle\pi'\rangle t]\!](s)$
$= [\![\langle\pi\rangle t +_\epsilon \langle\pi'\rangle t]\!](s).$

(R2): for any $t \in \hat{\Sigma}^\mu X$,
$[\![|\epsilon\pi + (1-\epsilon)\pi'|t]\!](s) =$
$= (1-\gamma)R^{\epsilon\pi + (1-\epsilon)\pi'}(s) + \gamma[\![t]\!](s)$
$= (1-\gamma) \sum_{a \in A} R^a(s)(\epsilon\pi + (1-\epsilon)\pi')(s)(a) + \gamma[\![t]\!](s)$
$= \epsilon((1-\gamma) \sum_{a \in A} R^a(s)\pi(s)(a) + \gamma[\![t]\!](s)) +$

$$+(1-\epsilon)((1-\gamma)\sum_{a\in A}R^a(s)\pi'(s)(a)+\gamma[\![t]\!](s))$$
$$=[\![|\pi|t+_\epsilon|\pi'|t]\!](s).$$

(R3): for any $t,t'\in\hat{\Sigma}^\mu X$,
$$|[\![|\pi|t]\!](s)-[\![|\pi|t']\!](s)|$$
$$=|(1-\gamma)R^\pi(s)+\gamma[\![t]\!](s)-(1-\gamma)R^\pi(s)-\gamma[\![t']\!](s)|$$
$$=\gamma|[\![t]\!](s)-[\![t']\!](s)|. \qquad\blacksquare$$

8.3 Solving the Bellman equation iteratively

We define now, for any $\pi\in\Pi$ a derived operator O^π inductively on the structure of the policy π as follows.

- For $a\in A$, $O^{\hat{a}}t=|\hat{a}|\langle\hat{a}\rangle t$.

- For $\pi,\pi'\in\Pi$ and $\epsilon\in[0,1]$,

$$O^{\epsilon\pi+(1-\epsilon)\pi'}t=O^\pi+_\epsilon O^{\pi'}.$$

Since all the distributions with finite support can be represented as convex combinations of Dirac distributions, any policy can be represented by a term with appropriately nested $+_\epsilon$ operators on top of constant policies. Hence the definition of O^π is complete.

The following theorem states that O^π is the syntactic counterpart of the Bellman operator T^π.

Theorem 8.3 *For any $\pi\in\Pi$ and any $t\in\hat{\Sigma}^\mu X$,*

$$[\![O^\pi t]\!]=T^\pi[\![t]\!].$$

Proof. We prove this inductively on the structure of $\pi\in\Pi$. Let $s\in S$.

For $\pi=\hat{a}$, $a\in A$,
$$[\![O^{\hat{a}}t]\!](s)=[\![\hat{a}\langle\hat{a}\rangle t]\!](s)$$
$$=(1-\gamma)R^a(s)+\gamma[\![\langle\hat{a}\rangle t]\!](s)$$
$$=(1-\gamma)R^a(s)+\gamma\sum_{s'\in S}P^a(s)(s')[\![t]\!](s')=T^{\hat{a}}[\![t]\!](s).$$
For $\epsilon\pi+(1-\epsilon)\pi'$ under the inductive hypothesis for $O^\pi t$ and $O^{\pi'}t$. We have
$$[\![O^{\epsilon\pi+(1-\epsilon)\pi'}t]\!](s)=[\![O^\pi t+_\epsilon O^{\pi'}t]\!](s)$$
$$=\epsilon[\![O^\pi t]\!](s)+(1-\epsilon)[\![O^{\pi'}t]\!](s)$$
$$=\epsilon T^\pi[\![t]\!](s)+(1-\epsilon)T^{\pi'}[\![t]\!](s)$$
$$=\epsilon\sum_{a\in A}\pi(s)(a)[(1-\gamma)R^a(s)+$$
$$\gamma\sum_{s'\in S}P^a(s)(s')[\![t]\!](s')]+$$
$$+(1-\epsilon)\sum_{a\in A}\pi'(s)(a)[(1-\gamma)R^a(s)+$$
$$+\gamma\sum_{s'\in S}P^a(s)(s')[\![t]\!](s')]$$
$$=\sum_{a\in A}(\epsilon\pi(s)+(1-\epsilon)\pi'(s))(a)[(1-\gamma)R^a(s)+$$
$$+\gamma\sum_{s'\in S}P^a(s)(s')[\![t]\!](s')]=T^{\epsilon\pi+(1-\epsilon)\pi'}[\![t]\!](s). \qquad\blacksquare$$

Next we verify that O^π has Banach pattern $\{\langle\gamma\rangle\}$.

Lemma 8.4 *For any $\pi \in \Pi$,*

$$O^\pi : 1 : \{\langle\gamma\rangle\} \in \hat{\Sigma}^\mu.$$

Proof. We prove, inductively on the structure of π, that
$x =_\epsilon y \vdash O^\pi x =_{\gamma\epsilon} O^\pi y \in \mathcal{B}$.
For $\pi = \hat{a}$, (NExp) for $\langle\hat{a}\rangle$ gives us
$x =_\epsilon y \vdash \langle\hat{a}\rangle x =_\epsilon \langle\hat{a}\rangle y \in \mathcal{B}$ and instantiating (R2),
$\langle\hat{a}\rangle x =_\epsilon \langle\hat{a}\rangle y \vdash |\hat{a}|\langle\hat{a}\rangle x =_{\epsilon\gamma} |\hat{a}|\langle\hat{a}\rangle y \in \mathcal{B}$,
hence, $x =_\epsilon y \vdash |\hat{a}|\langle\hat{a}\rangle x =_{\epsilon\gamma} |\hat{a}|\langle\hat{a}\rangle y \in \mathcal{B}$.
For $\epsilon\pi + (1 - \epsilon)\pi'$, consider the inductive hypotheses
$x =_\epsilon y \vdash O^\pi x =_{\epsilon\gamma} O^\pi y \in \mathcal{B}$ and
$x =_\epsilon y \vdash O^{\pi'} x =_{\epsilon\gamma} O^{\pi'} y \in \mathcal{B}$. (NExp) of $+_\epsilon$ gives

$$\{O^\pi x =_{\epsilon\gamma} O^\pi y, O^{\pi'} x =_{\epsilon\gamma} O^{\pi'} y\} \vdash$$

$$\vdash O^\pi x +_\epsilon O^{\pi'} x =_{\gamma\epsilon} O^\pi y +_\epsilon O^{\pi'} y \in \mathcal{B}.$$

Hence, $x =_\epsilon y \vdash O^\pi x +_\epsilon O^{\pi'} x =_{\gamma\epsilon} O^\pi y +_\epsilon O^{\pi'} y \in \mathcal{B}$
i.e., $x =_\epsilon y \vdash O^{\epsilon\pi+(1-\epsilon)\pi'} x =_{\gamma\epsilon} O^{\epsilon\pi+(1-\epsilon)\pi'} y \in \mathcal{B}$. ∎

Since our working hypothesis is that $\gamma < 1$, the previous lemma ensures that in the fixed-point extension of $\mathcal{B}$, which is $\mathcal{B}^\mu$, we have judgements involving $\mu x.O^\pi x$. We use this to show how the Bellman equation can be solved.

Recall that $[O^\pi]_1^k(s)$ represents the k-th iteration of O^π on s. Since O^π has only one variable, we drop the lower index 1 and write $[O^\pi]^k(s)$ for the k-th iteration on s.

The next theorem is a direct consequence of the Corollary 5.4 and Theorem 8.3.

Theorem 8.5 (Bellman equation) *For any $\pi \in \Pi$ and any $s \in \hat{\Sigma}^\mu X$, the sequence $([O^\pi]^k(s))_{k\geq 1}$ is convergent in $\mathcal{B}^\mu$ and its limit is $\mu x.O^\pi x$, i.e., $\forall \epsilon > 0 \; \exists n \; \forall m$,*

$$\vdash [O^\pi]^{m+n}(s) =_\epsilon \mu x.O^\pi x.$$

Moreover, $[\![\mu x.O^\pi x]\!]$ is the unique solution of Bellman equation

$$X = T^\pi X.$$

Note that the fixed-point Banach theory $\mathcal{B}^\mu$ gives us not only the solution to Bellman equation, but the apparatus for controlling "the speed" of convergence of the iteration sequence to the solution of Bellman equation. In this way, we can build an approximation theory directly inside $\mathcal{B}^\mu$.

9 Conclusions and related work

We have developed a quantitative fixed point theory extending the quantitative equational logic of [MPP16] by introducing fixed point operators and appropriate axioms. The key ingredients needed were the Banach patterns that capture the contractiveness of functions in their different arguments. We were able to mimic, in this setting, the standard iteration theories as described in [BE93] and [SP00]. We also developed an extended example showing that the notion of Bellman equations, which are the centrepiece of reinforcement learning, can be described in our framework.

A very general and interesting categorical treatment of iteration comes from the theory of traced monoidal categories [AJV96]. Recent work by Goncharov and Schröder [GS18] develops the notion of guarded traced categories which, like our Banach patterns, controls when traces can be taken. The monumental treatise of Bloom and Esik [BE93] also gives a very general treatment of iteration and mentions fixed points in metric spaces as an example. However, these theories are all in the traditional setting of equational logic and do not have the quantitative notions that we have here with approximate equality. Thus, for example, we can discuss the geometric rate of convergence in value iteration.

A very interesting formulation of the coinduction principle due to Dexter Kozen [Koz06, Koz07] is closely related to our rule for reasoning about fixed points. It is equivalent in power to our fixed-point approximation axiom, as we have argued. However his rule is very flexible and probably more convenient to use in various situations. It would certainly make an interesting variation to our formulation. We did consider both alternatives when we were developing our framework and at the moment we do not see a compelling reason to choose one over the other. This is definitely a topic which should be explored further.

While the fixed-point theory in this paper is infinitary, it would be interesting, as well as potentially useful, to develop a finitary version of it, and in this context, the Kozen principle of coinduction may be more powerful.

We have developed an example showing that some nontrivial situations can be modelled and reasoned about in our framework. Of course, whatever we have shown about Bellman equations has been long known, but it does show the potential power of the framework. In recent work Amortila et al. [APPB20] have proven convergence, using coupling techniques, of a variety of more recent reinforcement learning algorithms. It would be fascinating to see if the present framework could help to organize and reason about situations where the convergence has not yet been established.

References

[AJV96]   Ross Street Andre Joyal and Dominic Verity. Traced monoidal categories. *Math. Proc. Camb. Phil. Soc.*, 119:447–468, 1996.

[APPB20] Philip Amortila, Doina Precup, Prakash Panangaden, and Marc Bellemare. A distributional analysis of sampling-based reinforcement learning algorithms. In *The 23rd International Conference on Artificial Intelligence and Statistics*, 2020.

[Bak71]   Jaco W. De Bakker. *Recursive procedures*. Number 24 in Mathematical Centre Tracts. Mathematisch Centrum, Amsterdam, 1971.

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

[BE93]   S. Bloom and Z. Ésik. Iteration theories. *EATCS Monographs on Theoretical Computer Science*, 1993.

[EB95]   Z. Esik and L. Bernatsky. Scott Induction and Equational Proofs. *ENTCS*, 1:154–181, 1995.

[GS18]   Sergey Goncharov and Lutz Schröder. Guarded traced categories. In Christel Baier and Ugo Dal Lago, editors, *Foundations of Software Science and Computation Structures - 21st International Conference, FOSSACS 2018, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2018, Thessaloniki, Greece, April 14-20, 2018, Proceedings*, volume 10803 of *Lecture Notes in Computer Science*, pages 313–330. Springer, 2018.

[Has99]   M. Hasegawa. *Models of Sharing Graphs: A Categorical Semantics of let and letrec.* Distinguished Dissertation Series, Springer-Verlag, 1999.

[Kle52]   S. C. Kleene. *Introduction to metamathematics*. North-Holland, Amsterdam, 1952.

[Koz06]   Dexter Kozen. Coinductive proof principles for stochastic processes. In Rajeev Alur, editor, *Proceedings of the 21st Annual IEEE Symposium On Logic In Computer Science LICS'06*, pages 359–366, August 2006.

[Koz07]   Dexter Kozen. Coinductive proof principles for stochastic processes. *Logical Methods In Computer Science*, 3(4:8):1–14, 2007.

[MPP16]   Radu Mardare, Prakash Panangaden, and Gordon Plotkin. Quantitative algebraic reasoning. In *Proceedings of the 31st Annual ACM-IEEE Symposium on Logic in Computer Science*, pages 700–709, 2016.

[MPP17]   Radu Mardare, Prakash Panangaden, and Gordon Plotkin. On the axiomatizability of quantitative algebras. In *Proceedings of the 32nd Annual ACM-IEEE Symposium on Logic in Computer Science*, 2017.

[Put94]   Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 1994.

[SB]   Dana Scott and Jaco W. De Bakker. A theory of programs. Unpublished notes, IBM Seminar, Vienna.

[SB98]   Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[SP00]   Alex Simpson and Gordon Plotkin. Complete axioms for categorical fixed-point operators. In *Proceedings of the 15th Annual IEEE Symposium on Logic in Computer Science (LICS 2000)*, pages 30–41. IEEE, June 2000.