Computable exchangeable sequences have computable de Finetti measures

Cameron E. Freer[1] and Daniel M. Roy[2]

[1] Department of Mathematics
Massachusetts Institute of Technology
freer@math.mit.edu
[2] Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
droy@csail.mit.edu

Abstract. We prove a uniformly computable version of de Finetti's theorem on exchangeable sequences of real random variables. In the process, we develop machinery for computably recovering a distribution from its sequence of moments, which suffices to prove the theorem in the case of (almost surely) continuous directing random measures. In the general case, we give a proof inspired by a randomized algorithm which succeeds with probability one. Finally, we show how, as a consequence of the main theorem, exchangeable stochastic processes in probabilistic functional programming languages can be rewritten as procedures that do not use mutation.

Keywords: de Finetti, exchangeability, computable probability theory, probabilistic programming languages, mutation.

1 Introduction

This paper examines the computable probability theory of exchangeable sequences of real-valued random variables; the main contribution is a uniformly computable version of de Finetti's theorem.

The classical result states that an exchangeable sequence of real random variables is a mixture of independent and identically distributed (i.i.d.) sequences of random variables. Moreover, there is an (almost surely unique) measure-valued random variable, called the *directing random measure*, conditioned on which the random sequence is i.i.d. The distribution of the directing random measure is called the *de Finetti measure*.

We show that computable exchangeable sequences of real random variables have computable de Finetti measures. In the process, we show that a distribution on $[0,1]^\omega$ is computable if and only if its moments are uniformly computable.

This work is formulated in essentially the type-2 theory of effectivity (TTE) framework for computable analysis, though it is also related to domain theoretic representations of measures. Furthermore, our formalism coincides with those distributions from which we can generate exact samples on a computer.

In particular, we highlight the relationship between exchangeability and mutation in probabilistic programming languages. The computable de Finetti theorem can be used to uniformly transform procedures which induce exchangeable stochastic processes into equivalent procedures which do not use mutation (see Section 5).

1.1 de Finetti's Theorem

We assume familiarity with the standard measure-theoretic formulation of probability theory (see, e.g., [Bil95] or [Kal02]). Fix a basic probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{B}_{\mathbf{R}}$ denote the Borel sets of $\mathbf{R}$. By a *random measure* we mean a random element in the space of Borel measures on $\mathbf{R}$, i.e., a kernel from $(\Omega, \mathcal{F})$ to $(\mathbf{R}, \mathcal{B}_{\mathbf{R}})$. An event $A \in \mathcal{F}$ is said to occur *almost surely* (a.s.) if $\mathbb{P}A = 1$. We denote the indicator function of a set B by $\mathbf{1}_B$.

Let $X = \{X_i\}_{i\geq 1}$ be an infinite sequence of real random variables. An infinite sequence X is *exchangeable* if, for any finite set $\{k_1, \dots, k_j\}$ of distinct indices, $(X_{k_1}, \dots, X_{k_j}) \stackrel{d}{=} (X_1, \dots, X_j)$, where $\stackrel{d}{=}$ denotes equality in distribution.

Theorem 1 (de Finetti [Kal05, Chap. 1.1]). *Let $X = \{X_i\}_{i\geq 1}$ be an exchangeable sequence of real-valued random variables. There is a random probability measure ν on $\mathbf{R}$ such that $\{X_i\}_{i\geq 1}$ is conditionally i.i.d. with respect to ν. That is,*

$$\mathbb{P}[X \in \cdot \mid \nu] = \nu^\infty \quad a.s. \tag{1}$$

Moreover, ν is a.s. unique and given by

$$\nu(B) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_B(X_i) \quad a.s., \tag{2}$$

where B ranges over $\mathcal{B}_{\mathbf{R}}$. □

We call ν the *directing random measure*, and its distribution, $\mu := \mathcal{L}(\nu)$, the *de Finetti measure*. As in [Kal05, Chap. 1, Eq. 3], we may take expectations on both sides of (1) to arrive at a characterization

$$\mathbb{P}\{X \in \cdot\} = \mathbb{E}\nu^\infty = \int m^\infty \mu(dm) \tag{3}$$

of an exchangeable sequence as a mixture of i.i.d. sequences.

A Bayesian perspective suggests the following interpretation: exchangeable sequences arise from independent observations from some latent random measure. Posterior analysis follows from placing a prior distribution on ν.

In 1931, de Finetti [dF31] proved the classical result for binary exchangeable sequences, in which case the de Finetti measure is simply a mixture of Bernoulli distributions; the exchangeable sequence is equivalent to repeatedly flipping a coin whose weight is drawn from some distribution on $[0, 1]$. Later, Hewitt and

Savage [HS55] and Ryll-Nardzewski [RN57] extended the result to arbitrary real-valued exchangeable sequences. We will refer to this more general version as the *de Finetti theorem*. Hewitt and Savage [HS55] provide a history of the early developments, and a history of more recent extensions can be found in Kingman [Kin78], Diaconis and Freedman [DF84], Lauritzen [Lau84], and Aldous [Ald85].

1.2 The Computable de Finetti Theorem

Consider an exchangeable sequence of $[0, 1]$-valued random variables. In this case, the de Finetti measure is a distribution on the (Borel) measures on $[0, 1]$. For example, if the de Finetti measure is a Dirac delta on the uniform distribution on $[0, 1]$, then the induced exchangeable sequence consists of independent, uniformly distributed random variables on $[0, 1]$.

As another example, let p be a random variable, uniformly distributed on $[0, 1]$, and let $\nu = \delta_p$. Then the de Finetti measure is the uniform distribution on Dirac delta measures on $[0, 1]$, and the corresponding exchangeable sequence is $p, p, \dots$, i.e., a constant sequence, marginally uniformly distributed.

Clearly, sampling a measure ν, and then repeated sampling from ν, induces an exchangeable sequence; de Finetti's theorem states that all exchangeable sequences have an (a.s.) unique representation of this form.

In both examples, the de Finetti measures are *computable measures.* (In Section 2, we make this and related notions precise. For a more detailed example, see our discussion of the Pólya urn construction of the Beta-Bernoulli process in Section 5.) A natural question to ask is whether computable exchangeable sequences always arise from independent samples from computable random distributions. In fact, sampling from a computable de Finetti measure gives a computable directing random measure ν; repeated sampling from ν induces a computable exchangeable sequence (Proposition 1). Our main result is the converse: every computable real-valued exchangeable sequence arises from a computable de Finetti measure.

Theorem 2 (Computable de Finetti). *Let X be a real-valued exchangeable sequence and let μ be the distribution of its directing random measure ν. Then X is computable iff μ is computable. Moreover, μ is uniformly computable in X, and conversely.*

1.3 Outline of the Proof

Let $\mathcal{B}_{\mathbf{R}}$ denote the Borel sets of $\mathbf{R}$, let $\mathcal{I}_{\mathbf{R}}$ denote the set of finite unions of open intervals, and let $\mathcal{I}_{\mathbf{Q}}$ denote the set of finite unions of open intervals with rational endpoints. For $k \geq 1$ and $\boldsymbol{\beta} \in \mathcal{B}_{\mathbf{R}}^k = \mathcal{B}_{\mathbf{R}} \times \cdots \times \mathcal{B}_{\mathbf{R}}$, we write $\boldsymbol{\beta}(i)$ to denote the ith coordinate of $\boldsymbol{\beta}$.

Let $X = \{X_i\}_{i \geq 1}$ be an exchangeable sequence of real random variables, with directing random measure ν. For every $\boldsymbol{\gamma} \in \mathcal{B}_{\mathbf{R}}$, we define a $[0, 1]$-valued random variable $V_{\boldsymbol{\gamma}} := \nu\boldsymbol{\gamma}$. A classical result in probability theory [Kal02, Lem. 1.17] implies that a Borel measure on $\mathbf{R}$ is uniquely characterized by the mass it places on the open intervals with rational endpoints. Therefore, the distribution of the stochastic process $\{V_\tau\}_{\tau\in\mathcal{I}_\mathbf{Q}}$ determines the de Finetti measure μ (the distribution of ν).

The *mixed moments* of the variables $\{V_\gamma\}_{\gamma\in C}$, for a subset $C \subseteq \mathcal{B}_\mathbf{R}$, are defined to be the expectations of the monomials $\prod_{i=1}^{k} V_{\boldsymbol{\beta}(i)}$, for $k \geq 1$ and $\boldsymbol{\beta} \in C^k$. The following relationship between the finite dimensional distributions of the sequence X and the mixed moments of $\{V_\beta\}_{\beta\in\mathcal{B}_\mathbf{R}}$ is an immediate consequence of the characterization of an exchangeable sequence as a mixture of i.i.d. sequences given by (3).

Lemma 1. $\mathbb{P}\big(\bigcap_{i=1}^{k}\{X_i \in \boldsymbol{\beta}(i)\}\big) = \mathbb{E}\big(\prod_{i=1}^{k} V_{\boldsymbol{\beta}(i)}\big)$ *for* $k \geq 1$ *and* $\boldsymbol{\beta} \in \mathcal{B}_\mathbf{R}^k$. □

Assume that X is computable. If ν is a.s. continuous, we can compute the mixed moments of $\{V_\tau\}_{\tau\in\mathcal{I}_\mathbf{Q}}$. In Section 3, we show how to computably recover a distribution from its moments. This suffices to recover the de Finetti measure.

In the general case, point masses in ν prevent us from computing the mixed moments. Here we use a proof inspired by a randomized algorithm which (a.s.) avoids the point masses and recovers the de Finetti measure. For the complete proof, see Section 4.4.

2 Computable Measures

We assume familiarity with the standard notions of computability theory and computably enumerable (c.e.) sets (see, e.g., [Rog67] or [Soa87]). Recall that $r \in \mathbf{R}$ is called a c.e. real when the set of all rationals less than r is a c.e. set. Similarly, r is a co-c.e. real when the set of all rationals greater than r is c.e. A real r is a computable real when it is both a c.e. and co-c.e. real.

The following representations for probability measures on computable T_0 spaces [GSW07, §3] are devised from more general TTE representations in [Sch07] and [Bos08], and agree with [Wei99] in the case of the unit interval. This work is also related to domain theoretic representations of measure theory (given in [Eda95], [Eda96], and [AES00]) and to constructive analysis (see [Bau05]), though we do not explore this here.

Schröder [Sch07] has connected TTE representations with *probabilistic processes* [SS06] which sample representations. In Section 5 we explore consequences of the computable de Finetti theorem for probabilistic programming languages.

2.1 Representations of Probability Measures

Let S be a T_0 second-countable space with a countable basis $\mathcal{S}$ that is closed under finite unions. Assume that there is (and fix) an enumeration of $\mathcal{S}$ for which the sets $\{C \in \mathcal{S} : C \subseteq A \cap B\}$ are c.e., uniformly in the indices of $A, B \in \mathcal{S}$.

Let $\mathcal{M}_1(S)$ denote the set of Borel probability measures on S (i.e., the probability measures on the σ-algebra generated by $\mathcal{S}$). Such measures are determined by the measure they assign to elements of $\mathcal{S}$; we define the representation of

$\eta \in \mathcal{M}_1(S)$ with respect to $\mathcal{S}$ to be the set $\{(B, q) \in \mathcal{S} \times \mathbf{Q} : \eta B > q\}$. We say that η is a *computable measure* when its representation is a c.e. set (in terms of the fixed enumeration of $\mathcal{S}$). In other words, η is computable if the measure it assigns to a basis set $B \in \mathcal{S}$ is a c.e. real, uniformly in the index of B in the enumeration of $\mathcal{S}$. This representation for $\mathcal{M}_1(S)$ is admissible with respect to the weak topology, hence computably equivalent (see [Wei00, Chap. 3]) to the canonical representation for Borel measures given in [Sch07]. (Consequences for measurable functions with respect to admissible representations can be found in [BG07, §2].) We will be interested in representing measures $\eta \in \mathcal{M}_1(S)$ where S is either $\mathbf{R}^\omega$, $[0,1]^k$, or $\mathcal{M}_1(\mathbf{R})$, topologized as described below.

Consider $\mathbf{R}^\omega$ under the product topology, where $\mathbf{R}$ is endowed with its standard topology. Fix the canonical enumeration of $\mathcal{I}_\mathbf{Q}$ and the enumeration of $\bigcup_{k\geq 1} \mathcal{I}_\mathbf{Q}^k$ induced by interleaving (and the pairing function). A suitable basis for this product topology is given by the cylinders $\{\sigma \times \mathbf{R}^\omega : \sigma \in \bigcup_{k\geq 1} \mathcal{I}_\mathbf{Q}^k\}$. Let $\vec{x} = \{x_i\}_{i\geq 1}$ be a sequence of real-valued random variables (e.g., an exchangeable sequence X, or $\{V_\tau\}_{\tau \in \mathcal{I}_\mathbf{Q}}$ under the canonical enumeration of $\mathcal{I}_\mathbf{Q}$). The representation of the joint distribution of $\vec{x}$ is $\bigcup_{k\geq 1}\big\{(\sigma, q) \in \mathcal{I}_\mathbf{Q}^k \times \mathbf{Q} : \mathbb{P}\big(\bigcap_{i=1}^k \{x_i \in \sigma(i)\}\big) > q\big\}$. We say that $\vec{x}$ is computably distributed when this set is c.e. Informally, we will call a sequence $\vec{x}$ computable when it is computably distributed.

Fix $k \geq 1$, and consider the *right order topology* on $[0,1]$ generated by the basis $\mathcal{T} = \big\{(c, 1] : c \in \mathbf{Q},\ 0 \leq c < 1\big\} \cup \{[0,1]\}$. Let $\mathcal{S} = \mathcal{T}^k$ be a basis for the product of the right order topology on $[0,1]^k$. Let $\vec{w} = (w_1, \ldots, w_k)$ be a random vector in $[0,1]^k$. The representation of the joint distribution of $\vec{w}$ with respect to $\mathcal{S}$ is $\{(\vec{c}, q) \in \mathbf{Q}^k \times \mathbf{Q} : \mathbb{P}\big(\bigcap_{i=1}^k \{w_i > c_i\}\big) > q\}$. We will use this representation in Proposition 1 for $\{V_{\sigma(i)}\}_{i\leq k}$ where $\sigma \in \mathcal{I}_\mathbf{Q}^k$.

The de Finetti measure μ is the distribution of the directing random measure ν, an $\mathcal{M}_1(\mathbf{R})$-valued random variable. For $k \geq 1$, $\boldsymbol{\beta} \in \mathcal{B}_\mathbf{R}^k$, and $\vec{c} \in \mathbf{Q}^k$, we define the event $Y_{\boldsymbol{\beta},\vec{c}} := \bigcap_{i=1}^k \{\nu \boldsymbol{\beta}(i) > c_i\}$. The sets $\{\nu : \nu\tau > c\}$ for $\tau \in \mathcal{I}_\mathbf{Q}$ and $c \in \mathbf{Q}$ form a subbasis for the weak topology on $\mathcal{M}_1(\mathbf{R})$. We therefore represent the distribution μ by $\bigcup_{k\geq 1}\{(\sigma, \vec{c}, q) \in \mathcal{I}_\mathbf{Q}^k \times \mathbf{Q}^k \times \mathbf{Q} : \mathbb{P}Y_{\sigma,\vec{c}} > q\}$, and say that μ is computable when $\mathbb{P}Y_{\sigma,\vec{c}}$ is a c.e. real, uniformly in σ and $\vec{c}$.

2.2 Computable de Finetti Measures

In the remainder of the paper, let X be a real-valued exchangeable sequence with directing random measure ν and de Finetti measure μ.

Proposition 1. *X is uniformly computable in μ.*

Proof. We give a proof that relativizes to μ; without loss of generality, assume that μ is computable. Then $\mathbb{P}\big(\bigcap_{i=1}^k \{\nu : \nu\sigma(i) > c_i\}\big)$ is a c.e. real, uniformly in $\sigma \in \mathcal{I}_\mathbf{Q}^k$ and $\vec{c} \in \mathbf{Q}^k$. Therefore, the joint distribution of $\{V_{\sigma(i)}\}_{i\leq k}$ under the right order topology is computable, uniformly in σ. By Lemma 1, $\mathbb{P}\big(\bigcap_{i=1}^k \{X_i \in \sigma(i)\}\big) = \mathbb{E}\big(\prod_{i=1}^k V_{\sigma(i)}\big)$. Because the monomial maps $[0,1]^k$ into $[0,1]$ and is continuous, this expectation is a c.e. real, uniformly in σ [Sch07, Prop. 3.6]. ☐

3 The Computable Moment Problem

One often has access to the moments of a distribution, and wishes to recover the underlying distribution. Let η be a distribution on $[0,1]^\omega$, and let $\vec{x} = \{x_i\}_{i\geq 1}$ be a sequence of $[0,1]$-valued random variables with distribution η. Classically, the distribution of η is uniquely determined by the mixed moments of $\vec{x}$. We show that a distribution on $[0,1]^\omega$ is computable iff its sequence of mixed moments is uniformly computable.

To show that η is computable, it suffices to show that $\eta(\sigma \times [0,1]^\omega)$ is a c.e. real for $\sigma \in \mathcal{I}^k_{\mathbf{Q}}$, uniformly in σ. Recall that $\eta(\sigma \times [0,1]^\omega) = \mathbb{E}\mathbf{1}_{\sigma\times[0,1]^\omega}$. Using the effective Weierstrass theorem ([Wei00, p. 161] or [PR89, p. 45]), we can approximate the indicator function by a sequence of rational polynomials that converge pointwise from below:

Lemma 2 (Polynomial approximations). *There is a computable sequence* $\{p_{n,\sigma} : n \in \omega,\ \sigma \in \bigcup_{k\geq 1} \mathcal{I}^k_{\mathbf{Q}}\}$ *of rational polynomials where* $p_{n,\sigma}$ *is in* k *variables (for* $\sigma \in \mathcal{I}^k_{\mathbf{Q}}$*) and, for* $\vec{x} \in [0,1]^k$ *we have* $-1 \leq p_{n,\sigma}(\vec{x}) \leq \mathbf{1}_\sigma(\vec{x})$ *and* $\lim_n p_{n,\sigma}(\vec{x}) = \mathbf{1}_\sigma(\vec{x})$*.* □

By the dominated convergence theorem, the expectations of the sequence converge to $\eta(\sigma \times [0,1]^\omega)$ from below:

Lemma 3. *Fix* $k \geq 1$ *and* $\sigma \in \mathcal{I}^k_{\mathbf{Q}}$*, and let* $\vec{x} = (x_1, \ldots, x_k)$ *be a random vector in* $[0,1]^k$*. Then* $\mathbb{E}\big(\mathbf{1}_\sigma(\vec{x})\big) = \sup_n \mathbb{E}\big(p_{n,\sigma}(\vec{x})\big)$*.* □

Theorem 3 (Computable moments). *Let* $\vec{x} = (x_i)_{i\geq 1}$ *be a random vector in* $[0,1]^\omega$ *with distribution* η*. Then* η *is computable iff the mixed moments of* $\{x_i\}_{i\geq 1}$ *are uniformly computable.*

Proof. All monomials in $\{x_i\}_{i\geq 1}$ are bounded and continuous on $[0,1]^\omega$. If η is computable, then the mixed moments are uniformly computable by a computable integration result [Sch07, Prop. 3.6].

Now suppose the mixed moments of $\{x_i\}_{i\geq 1}$ are uniformly computable. To establish the computability of η, it suffices to show that $\eta(\sigma\times[0,1]^\omega)$ is a c.e. real, uniformly in σ. Note that $\eta(\sigma \times [0,1]^\omega) = \mathbb{E}\big(\mathbf{1}_{\sigma\times[0,1]^\omega}(\vec{x})\big) = \mathbb{E}\big(\mathbf{1}_\sigma(x_1,\ldots,x_k)\big)$. By Lemma 3, $\mathbb{E}\big(\mathbf{1}_\sigma(x_1,\ldots,x_k)\big) = \sup_n \mathbb{E}\big(p_{n,\sigma}(x_1,\ldots,x_k)\big)$. The sequence of polynomials $\big(p_{n,\sigma}(x_1,\ldots,x_k)\big)_{n\in\omega}$ is uniformly computable by Lemma 2. The expectation of each polynomial is a $\mathbf{Q}$-linear combination of moments, hence computable. □

4 The Computable de Finetti Theorem

4.1 Continuous Directing Random Measures

For $k \geq 1$ and $\boldsymbol{\psi} \in \mathcal{I}^k_{\mathbf{R}}$, we say that $\boldsymbol{\psi}$ has *no mass on its boundary* when $\mathbb{P}\big(\bigcup_{i=1}^k \{X_i \in \partial\boldsymbol{\psi}(i)\}\big) = 0$.

Lemma 4. *Suppose X is computable. Then the mixed moments of $\{V_\tau\}_{\tau\in\mathcal{I}_{\mathbf{Q}}}$ are uniformly c.e. reals and the mixed moments of $\{V_{\overline{\tau}}\}_{\tau\in\mathcal{I}_{\mathbf{Q}}}$ are uniformly co-c.e. reals. In particular, if $\sigma \in \mathcal{I}_{\mathbf{Q}}^k$ has no mass on its boundary, then the moment $\mathbb{E}\big(\prod_{i=1}^k V_{\sigma(i)}\big)$ is a computable real.*

Proof. Let $k \geq 1$ and $\sigma \in \mathcal{I}_{\mathbf{Q}}^k$. By Lemma 1, $\mathbb{E}\big(\prod_{i=1}^k V_{\sigma(i)}\big) = \mathbb{P}\big(\bigcap_{i=1}^k \{X_i \in \sigma(i)\}\big)$, which is a c.e. real because X is computable. Similarly, $\mathbb{E}\big(\prod_{i=1}^k V_{\overline{\sigma(i)}}\big) = \mathbb{P}\big(\bigcap_{i=1}^k \{X_i \in \overline{\sigma(i)}\}\big)$, which is the probability of a finite union of closed rational cylinders, and thus a co-c.e. real (as we can computably enumerate all $\tau \in \mathcal{I}_{\mathbf{Q}}^k$ contained in the complement of a given closed rational cylinder). When σ has no mass on its boundary, $\mathbb{E}\big(\prod_{i=1}^k V_{\sigma(i)}\big) = \mathbb{E}\big(\prod_{i=1}^k V_{\overline{\sigma(i)}}\big)$, which is both c.e. and co-c.e. □

Proposition 2 (Computable de Finetti – a.s. continuous ν). *Assume that ν is continuous with probability one. Then μ is uniformly computable in X, and conversely.*

Proof. By Proposition 1, X is uniformly computable in μ. We now give a proof of the other direction that relativizes to X; without loss of generality, assume that X is computable. Let $k \geq 1$ and consider $\sigma \in \mathcal{I}_{\mathbf{Q}}^k$. The (a.s.) continuity of ν implies that $\mathbb{P}\big(\bigcup_{i=1}^k \{X_i \in \partial\sigma(i)\}\big) = 0$, i.e., σ has no mass on its boundary. Therefore the moment $\mathbb{E}\big(\prod_{i=1}^k V_{\sigma(i)}\big)$ is computable by Lemma 4. By the computable moment theorem (Theorem 3), the joint distribution of the variables $\{V_\tau\}_{\tau\in\mathcal{I}_{\mathbf{Q}}}$ is computable. □

4.2 "Randomized" Proof Sketch

The task of lower bounding the measure of the events $Y_{\sigma,\vec{c}}$ for $\sigma \in \mathcal{I}_{\mathbf{Q}}^k$ and $\vec{c} \in \mathbf{Q}^k$ is complicated by the potential existence of point masses on any of the boundaries $\partial\sigma(i)$ of $\sigma(i)$. If X is a computable exchangeable sequence for which ν is discontinuous at some rational point with non-zero probability, the mixed moments of $\{V_\tau\}_{\tau\in\mathcal{I}_{\mathbf{Q}}}$ are c.e., but not co-c.e., reals (by Lemma 4). Therefore the computable moment theorem (Theorem 3) is inapplicable. For arbitrary directing random measures, we give a proof of the computable de Finetti theorem which works regardless of the location of point masses.

Consider the following sketch of a "randomized algorithm": We draw a countably infinite set of real numbers $A \subseteq \mathbf{R}$ from a continuous distribution with support on the entire real line (e.g., Gaussian or Cauchy). Note that, with probability one (over the draw), A will be dense in $\mathbf{R}$, and no point mass of ν will be in A. Let $\mathcal{I}_A$ denote the set of finite unions of intervals with endpoints in A. Because there is (a.s.) no mass on any boundary of any $\boldsymbol{\psi} \in \mathcal{I}_A$, we can proceed as in the case of (a.s.) continuous ν, in terms of this alternative basis.

If X is computable, we can compute all moments (relative to the infinite representations of the points of A) by giving c.e. (in A) lower and (as in Lemma 4) upper bounds. By the computable moment theorem (Theorem 3) we can compute (in A) the joint distribution of $\{V_{\boldsymbol{\psi}}\}_{\boldsymbol{\psi} \in \mathcal{I}_A}$. This joint distribution classically determines the de Finetti measure. Moreover, we can compute (in A) all rational lower bounds on any $\mathbb{P}Y_{\sigma,\vec{c}}$ for $\sigma \in \mathcal{I}^k_{\mathbf{Q}}$, our original basis.

This algorithm essentially constitutes a (degenerate) probabilistic process (see [SS06]) which returns a representation of the de Finetti measure with probability one. A proof along these lines could be formalized using the TTE and probabilistic process frameworks. Here, however, we provide a construction that makes explicit the representation in terms of our rational basis, and which can be seen as a "derandomization" of this algorithm.

Weihrauch [Wei99, Thm. 3.6] proves a computable integration result via an argument that could likewise be seen as a derandomization of an algorithm which densely subdivides the unit interval at random locations to avoid mass on the division boundaries. Similar arguments are used to avoid mass on the boundaries of open hypercubes [Mül99, Thm. 3.7] and open balls [Bos08, Lem. 2.15]. These arguments also resemble the classic proof of the Portmanteau theorem [Kal02, Thm. 4.25], in which an uncountable family of sets with disjoint boundaries is defined, almost all of which have no mass on their boundaries.

4.3 Rational Refinements and Polynomial Approximations

We call $\boldsymbol{\psi} \in \mathcal{I}^k_{\mathbf{R}}$ a *refinement* of $\boldsymbol{\varphi} \in \mathcal{I}^k_{\mathbf{R}}$, and write $\boldsymbol{\psi} \lhd \boldsymbol{\varphi}$, when $\overline{\boldsymbol{\psi}(i)} \subseteq \boldsymbol{\varphi}(i)$ for all $i \leq k$. For $n \in \omega$ and $\vec{c} = (c_1, \dots, c_k) \in \mathbf{Q}^k$, we will denote by $p_{n,\vec{c}}$ the polynomial $p_{n,\sigma}$ (defined in Lemma 2), where $\sigma = (c_1, 2) \times \cdots \times (c_k, 2)$. Let $\vec{x} = (x_1, \dots, x_k)$, and similarly with $\vec{y}$. We can write $p_{n,\vec{c}}(\vec{x}) = p^+_{n,\vec{c}}(\vec{x}) - p^-_{n,\vec{c}}(\vec{x})$, where $p^+_{n,\vec{c}}$ and $p^-_{n,\vec{c}}$ are polynomials with positive coefficients. Define the $2k$-variable polynomial $q_{n,\vec{c}}(\vec{x},\vec{y}) := p^+_{n,\vec{c}}(\vec{x}) - p^-_{n,\vec{c}}(\vec{y})$. We will denote $q_{n,\vec{c}}(V_{\boldsymbol{\psi}(1)}, \dots, V_{\boldsymbol{\psi}(k)}, V_{\boldsymbol{\zeta}(1)}, \dots, V_{\boldsymbol{\zeta}(k)})$ by $q_{n,\vec{c}}(V_{\boldsymbol{\psi}}, V_{\boldsymbol{\zeta}})$, and similarly with $p_{n,\vec{c}}$.

Proposition 3. *Let $\sigma \in \mathcal{I}^k_{\mathbf{Q}}$ and let $\vec{c} \in \mathbf{Q}^k$ and $n \in \omega$. If X is computable, then $\mathbb{E}q_{n,\vec{c}}(V_\sigma, V_{\overline{\sigma}})$ is a c.e. real (uniformly in $\sigma, \vec{c}$, and n).*

Proof. By Lemma 4, each monomial of $p^+_{n,\vec{c}}(V_\sigma)$ has a c.e. real expectation, and each monomial of $p^-_{n,\vec{c}}(V_{\overline{\sigma}})$ has a co-c.e. real expectation, and so by linearity of expectation, $\mathbb{E}q_{n,\vec{c}}(V_\sigma, V_{\overline{\sigma}})$ is a c.e. real. ◻

4.4 Proof of the Computable de Finetti Theorem

Proof of Theorem 2 (Computable de Finetti). The sequence X is uniformly computable in μ by Proposition 1. We now give a proof of the other direction that relativizes to X; without loss of generality, suppose that X is computable. Let $\pi \in \mathcal{I}^k_{\mathbf{Q}}$ and $\vec{c} \in \mathbf{Q}^k$. It suffices to show that the probability $\mathbb{P}Y_{\pi,\vec{c}}$ is a c.e. real (uniformly in π and $\vec{c}$). We do this by a series of reductions to quantities of the form $\mathbb{E}q_{n,\vec{c}}(V_\sigma, V_{\overline{\sigma}})$ for $\sigma \in \mathcal{I}^k_{\mathbf{Q}}$, which by Proposition 3 are c.e. reals.

Note that $\mathbb{P}Y_{\pi,\vec{c}} = \mathbb{E}\mathbf{1}_{Y_{\pi,\vec{c}}}$. If $\boldsymbol{\psi} \in \mathcal{I}^k_{\mathbf{R}}$ satisfies $\boldsymbol{\psi} \lhd \pi$, then $Y_{\boldsymbol{\psi},\vec{c}} \subseteq Y_{\pi,\vec{c}}$. Furthermore, $\mathcal{I}^k_{\mathbf{R}}$ is dense in $\mathcal{I}^k_{\mathbf{Q}}$, and so by the dominated convergence theorem we have that $\mathbb{E}\mathbf{1}_{Y_{\pi,\vec{c}}} = \sup_{\boldsymbol{\psi}\lhd\pi} \mathbb{E}\mathbf{1}_{Y_{\boldsymbol{\psi},\vec{c}}}$. By the definitions of Y and V we have $\mathbb{E}\mathbf{1}_{Y_{\boldsymbol{\psi},\vec{c}}} = \mathbb{E}(\mathbf{1}_{(c_1,2)\times\cdots\times(c_k,2)}(V_{\boldsymbol{\psi}(1)},\dots,V_{\boldsymbol{\psi}(k)}))$, which, by a Weierstrass approximation (Lemma 3) equals $\sup_n \mathbb{E}p_{n,\vec{c}}(V_{\boldsymbol{\psi}})$.

If $\boldsymbol{\zeta},\boldsymbol{\varphi} \in \mathcal{I}_{\mathbf{R}}$ satisfy $\boldsymbol{\zeta} \lhd \boldsymbol{\varphi}$ then $V_{\boldsymbol{\zeta}} \le V_{\boldsymbol{\varphi}}$ a.s. Multiplication is continuous, and so the dominated convergence theorem gives us

$$\mathbb{E}\left(\prod_{i=1}^k V_{\boldsymbol{\psi}(i)}\right) = \sup_{\sigma\lhd\boldsymbol{\psi}} \mathbb{E}\left(\prod_{i=1}^k V_{\sigma(i)}\right) \quad \text{and} \tag{4}$$

$$\mathbb{E}\left(\prod_{i=1}^k V_{\overline{\boldsymbol{\psi}(i)}}\right) = \inf_{\tau\rhd\boldsymbol{\psi}} \mathbb{E}\left(\prod_{i=1}^k V_{\overline{\tau(i)}}\right), \tag{5}$$

where σ and τ range over $\mathcal{I}_{\mathbf{Q}}^k$. By the linearity of expectation, $\mathbb{E}p^+_{n,\vec{c}}(V_{\boldsymbol{\psi}}) = \sup_{\sigma\lhd\boldsymbol{\psi}} \mathbb{E}p^+_{n,\vec{c}}(V_\sigma)$ and $\mathbb{E}p^-_{n,\vec{c}}(V_{\boldsymbol{\psi}}) = \inf_{\tau\rhd\boldsymbol{\psi}} \mathbb{E}p^-_{n,\vec{c}}(V_{\overline{\tau}})$.

If $\boldsymbol{\psi}$ has no mass on its boundary then $V_{\boldsymbol{\psi}(i)} = V_{\overline{\boldsymbol{\psi}(i)}}$ a.s. for $i \le k$, and so

$$\begin{aligned}
\mathbb{E}p_{n,\vec{c}}(V_{\boldsymbol{\psi}}) &= \mathbb{E}q_{n,\vec{c}}(V_{\boldsymbol{\psi}}, V_{\boldsymbol{\psi}}) && (6)\\
&= \mathbb{E}q_{n,\vec{c}}(V_{\boldsymbol{\psi}}, V_{\overline{\boldsymbol{\psi}}}) && (7)\\
&= \mathbb{E}p^+_{n,\vec{c}}(V_{\boldsymbol{\psi}}) - \mathbb{E}p^-_{n,\vec{c}}(V_{\overline{\boldsymbol{\psi}}}) && (8)\\
&= \sup_{\sigma\lhd\boldsymbol{\psi}} \mathbb{E}p^+_{n,\vec{c}}(V_\sigma) - \inf_{\tau\rhd\boldsymbol{\psi}} \mathbb{E}p^-_{n,\vec{c}}(V_{\overline{\tau}}) && (9)\\
&= \sup_{\sigma\lhd\boldsymbol{\psi}\lhd\tau} \mathbb{E}q_{n,\vec{c}}(V_\sigma, V_{\overline{\tau}}). && (10)
\end{aligned}$$

Because there are only countably many point masses, the $\boldsymbol{\psi} \in \mathcal{I}_{\mathbf{R}}^k$ with no mass on their boundaries are dense in $\mathcal{I}_{\mathbf{Q}}^k$. Therefore,

$$\sup_{\boldsymbol{\psi}\lhd\pi} \mathbb{E}p_{n,\vec{c}}(V_{\boldsymbol{\psi}}) = \sup_{\boldsymbol{\psi}\lhd\pi}\ \sup_{\sigma\lhd\boldsymbol{\psi}\lhd\tau} \mathbb{E}q_{n,\vec{c}}(V_\sigma, V_{\overline{\tau}}). \tag{11}$$

Note that $\{(\sigma,\tau) : (\exists\boldsymbol{\psi} \lhd \pi)\ \sigma \lhd \boldsymbol{\psi} \lhd \tau\} = \{(\sigma,\tau) : \sigma \lhd \pi \text{ and } \sigma \lhd \tau\}$. Hence

$$\sup_{\boldsymbol{\psi}\lhd\pi}\ \sup_{\sigma\lhd\boldsymbol{\psi}}\ \sup_{\tau\rhd\boldsymbol{\psi}} \mathbb{E}q_{n,\vec{c}}(V_\sigma, V_{\overline{\tau}}) = \sup_{\sigma\lhd\pi}\ \sup_{\tau\rhd\sigma} \mathbb{E}q_{n,\vec{c}}(V_\sigma, V_{\overline{\tau}}). \tag{12}$$

Again by dominated convergence we have $\sup_{\tau\rhd\sigma} \mathbb{E}q_{n,\vec{c}}(V_\sigma, V_{\overline{\tau}}) = \mathbb{E}q_{n,\vec{c}}(V_\sigma, V_{\overline{\sigma}})$. In summary, we have $\mathbb{P}Y_{\pi,\vec{c}} = \sup_n \sup_{\sigma\lhd\pi} \mathbb{E}q_{n,\vec{c}}(V_\sigma, V_{\overline{\sigma}})$. Finally, by Proposition 3, $\{\mathbb{E}q_{n,\vec{c}}(V_\sigma, V_{\overline{\sigma}}) : \sigma \lhd \pi \text{ and } n \in \omega\}$ is a set of uniformly c.e. reals. ☐

5 Exchangeability in Probabilistic Programs

Functional programming languages with probabilistic choice operators have recently been proposed as universal languages for statistical modeling (e.g., IBAL [Pfe01], $\lambda_\circ$[PPT08], and Church [GMR+08]). Although the semantics of probabilistic programs have been studied extensively in theoretical computer science in the context of randomized algorithms (e.g., [Koz81] and [JP89]), this application of probabilistic programs to universal statistical modeling has a different character which has raised a number of interesting theoretical questions (e.g., [RP02], [PPT08], [GMR+08], [RMG+08], and [Man09]).

The computable de Finetti theorem has implications for the semantics of probabilistic programs, especially concerning the choice of whether or not to allow mutation (i.e., the ability of a program to modify its own state as it runs). For concreteness, we will explore this connection using Church, a probabilistic functional programming language. Church extends Scheme (a dialect of LISP) with a boolean-valued `flip` procedure, which denotes the Bernoulli distribution. In Church, an expression denotes the distribution induced by evaluation. For example, `(+ (flip) (flip) (flip))` denotes a Binomial$(n = 3, p = \frac{1}{2})$ distribution and `(λ (x) (if (flip) x 0))` denotes the probability kernel $x \mapsto \frac{1}{2}(\delta_x + \delta_0)$. Church is call-by-value and so `(= (flip) (flip))` denotes a Bernoulli distribution, while the application of the procedure `(λ (x) (= x x))` to the argument `(flip)`, written `((λ (x) (= x x)) (flip))`, denotes δ_1. (For more details, see [GMR+08].) If an expression modifies its environment using mutation it might not denote a fixed distribution. For example, a procedure may keep a counter variable and return an increasing sequence of integers on repeated calls.

Consider the Beta-Bernoulli process and the Pólya urn scheme written in Church. While these two processes look different, they induce the same distribution on sequences. We will define the procedure `sample-coin` such that calling `sample-coin` returns a new procedure which itself returns random binary values. The probabilistic program

```
(define my-coin (sample-coin))
(my-coin) (my-coin) (my-coin) (my-coin) (my-coin) ...
```

defines a random binary sequence. Fix $a, b > 0$. Consider the following two implementations of `sample-coin` (and recall that the `(λ () ... )` special form creates a procedure of no arguments):

(i)

```
(define (sample-coin)
  (let ((weight (beta a b)))
    (λ () (flip weight)) ) )
```

(ii)

```
(define (sample-coin)
  (let ((red a)
        (total (+ a b)) )
    (λ () (let ((x (flip red/total)))
            (set! red (+ red x))
            (set! total (+ total 1))
            x ) ) )
```

In case *(i)*, evaluating `(my-coin)` returns a 1 with probability `weight` and a 0 otherwise, where the shared `weight` parameter is itself drawn from a Beta(a, b) distribution on $[0, 1]$. Note that the sequence of values obtained by evaluating `(my-coin)` is exchangeable but not i.i.d. (e.g., an initial sequence of ten 1's leads one to predict that the next draw is more likely to be 1 than 0). However, conditioned on the `weight` (a variable within the opaque procedure `my-coin`) the sequence is i.i.d. A second random coin constructed by `(define your-coin` `(sample-coin))` will have a different weight (a.s.) and will generate a sequence that is independent of `my-coin`'s.

The sequence induced by *(i)* is exchangeable because applications of ***`beta`*** and ***`flip`*** return independent samples. The code in *(ii)* implements the Pólya urn scheme with a red balls and b black balls (see [dF75, Chap. 11.4]), and the sequence of return values is exchangeable because its joint distribution depends only on the number of red and black balls and not on their order.

Because the sequence induced by *(ii)* is exchangeable, de Finetti's theorem implies that its distribution is equivalent to that induced by i.i.d. draws from some some random measure (the directing random measure). In the case of the Pólya urn scheme, the directing random measure is a random Bernoulli whose weight parameter has a Beta(a, b) distribution. Therefore *(i)* denotes the de Finetti measure, and *(i)* and *(ii)* induce the same distribution over sequences.

However, there is an important difference between these two implementations. While *(i)* denotes the de Finetti measure, *(ii)* does not, because samples from it do not denote fixed distributions: The state of a procedure `my-coin` sampled using *(ii)* changes after each iteration, as the sufficient statistics are updated (using the mutation operator `set!`). Therefore, each element of the sequence is generated from a different distribution. Even though the sequence of calls to such a `my-coin` has the same *marginal* distribution as those given by repeated calls to a `my-coin` sampled using *(i)*, a procedure `my-coin` sampled using *(ii)* could be thought of as a probability kernel which depends on the state.

In contrast, a `my-coin` sampled using *(i)* does not modify itself via mutation; the value of `weight` does not change after it is randomly initialized and therefore `my-coin` denotes a fixed distribution — a particular Bernoulli. Its marginal distribution is a random Bernoulli, precisely the directing random measure of the Beta-Bernoulli process.

The computable de Finetti theorem could be used to uniformly transform *(ii)* into a procedure ***`sample-weight`*** which does not use mutation and whose application is equivalent in distribution to the evaluation of `(beta a b)`. In general, an implementation of Theorem 2 transforms code which generates an exchangeable sequence (like *(ii)*) into code representing the de Finetti measure (i.e., a procedure of the form *(i)* which does not use mutation). In addition to their simpler semantics, mutation-free procedures are often desirable for practical reasons. For example, having sampled the directing random measure, an exchangeable sequence of random variables can be efficiently sampled in parallel without the overhead necessary to communicate sufficient statistics.

5.1 Partial Exchangeability of Arrays and Other Data Structures

The example above involved binary sequences, but the computable de Finetti theorem can be used to transform implementations of real exchangeable sequences. For example, given a probabilistic program outputting repeated draws from a distribution with a Dirichlet process prior via the Chinese restaurant process [Ald85] representation, we can automatically recover a version of the random directing measure characterized by Sethuraman's stick-breaking construction [Set94].

More general de Finetti-type results deal with variables taking values other than reals, or weaken exchangeability to various notions of partial exchangeability. The Indian buffet process, defined in [GG05], can be interpreted as a set-valued exchangeable sequence and can be written in a way analogous to the Pólya urn in *(ii)*. A corresponding stick-breaking construction given by [TGG07] (following [TJ07]) is analogous to the code in *(i)*, but gives only a Δ_1-index for the (a.s.) finite set of indices (with respect to a random countable set of reals sampled from the base measure), rather than its canonical index (see [Soa87, II.2]). This observation was first noted in [RMG$^+$08]. In the case of the Indian buffet process on a discrete base measure (where we can fix an enumeration of all finite sets of elements in the support), the computable de Finetti theorem implies the existence of a computable de Finetti measure that gives canonical indices for the sets. In the general case, the computable de Finetti theorem is not directly applicable, and the question of its computability is open.

Probabilistic models of relational data ([KTG$^+$06] and [RT09]) can be viewed as random binary (2-dimensional) arrays whose distributions are *separately* (or *jointly*) exchangeable, i.e., invariant under (simultaneous) permutations of the rows and columns. Aldous [Ald81] and Hoover [Hoo79] gave de Finetti-type results for infinite arrays that satisfy these two notions of partial exchangeability. An extension of the computable de Finetti theorem to this setting would provide an analogous uniform transformation on arrays.

Data structures such as trees or graphs might also be approached in this manner. Diaconis and Janson [DJ08] and Austin [Aus08] explore many connections between partially exchangeable random graphs and the theory of graph limits.

Acknowledgements
The authors would like to thank Vikash Mansinghka, Hartley Rogers, and the anonymous referees for helpful comments.

References

[AES00] Alvarez-Manilla, M., Edalat, A., Saheb-Djahromi, N.: An extension result for continuous valuations. J. London Math. Soc. (2) 61(2), 629–640 (2000)

[Ald81] Aldous, D.J.: Representations for partially exchangeable arrays of random variables. J. Multivariate Analysis 11(4), 581–598 (1981)

[Ald85] Aldous, D.J.: Exchangeability and related topics. In: École d'été de probabilités de Saint-Flour, XIII—1983. Lecture Notes in Math., vol. 1117. Springer, Berlin, 1–198 (1985)

[Aus08] Austin, T.: On exchangeable random variables and the statistics of large graphs and hypergraphs. Probab. Surv. 5, 80–145 (2008)

[Bau05] Bauer, A.: Realizability as the connection between constructive and computable mathematics. In: CCA 2005: Second Int. Conf. on Comput. and Complex. in Analysis (2005)

[BG07] Brattka, V., Gherardi, G.: Borel complexity of topological operations on computable metric spaces. In: CiE 2007, LNCS, vol. 4497, Springer, Heidelberg, 83–97 (2007)

[Bil95] Billingsley, P.: Probability and measure. Third edn. John Wiley & Sons Inc., New York (1995)

[Bos08] Bosserhoff, V.: Notions of probabilistic computability on represented spaces. J. of Universal Comput. Sci. 14(6), 956–995 (2008)

[dF31] de Finetti, B.: Funzione caratteristica di un fenomeno aleatorio. Atti della R. Accademia Nazionale dei Lincei, Ser. 6. Memorie, Classe di Scienze Fisiche, Matematiche e Naturali 4, 251–299 (1931)

[dF75] de Finetti, B.: Theory of probability. Vol. 2. John Wiley & Sons Ltd., London (1975)

[DF84] Diaconis, P., Freedman, D.: Partial exchangeability and sufficiency. In: Statistics: applications and new directions (Calcutta, 1981). Indian Statist. Inst., Calcutta, 205–236 (1984)

[DJ08] Diaconis, P., Janson, S.: Graph limits and exchangeable random graphs. Rendiconti di Matematica, Ser. VII 28(1), 33–61 (2008)

[Eda95] Edalat, A.: Domain theory and integration. Theoret. Comput. Sci. 151(1), 163–193 (1995)

[Eda96] Edalat, A.: The Scott topology induces the weak topology. In: 11th Ann. IEEE Symp. on Logic in Comput. Sci. IEEE Comput. Soc. Press, Los Alamitos, CA, 372–381 (1996)

[GG05] Griffiths, T.L., Ghahramani, Z.: Infinite latent feature models and the Indian buffet process. In: Adv. in Neural Inform. Processing Syst. 17, MIT Press, Cambridge, MA, 475–482 (2005)

[GMR+08] Goodman, N.D., Mansinghka, V.K., Roy, D.M., Bonawitz, K., Tenenbaum, J.B.: Church: a language for generative models. In: Uncertainty in Artificial Intelligence (2008)

[GSW07] Grubba, T., Schröder, M., Weihrauch, K.: Computable metrization. Math. Logic Q. 53(4-5), 381–395 (2007)

[Hoo79] Hoover, D.N.: Relations on probability spaces and arrays of random variables. Preprint, Institute for Advanced Study, Princeton, NJ (1979)

[HS55] Hewitt, E., Savage, L.J.: Symmetric measures on Cartesian products. Trans. Amer. Math. Soc. 80, 470–501 (1955)

[JP89] Jones, C., Plotkin, G.: A probabilistic powerdomain of evaluations. In: Proc. of the Fourth Ann. Symp. on Logic in Comp. Sci., IEEE Press, 186–195 (1989)

[Kal02] Kallenberg, O.: Foundations of modern probability. Second edn. Springer, New York (2002)

[Kal05] Kallenberg, O.: Probabilistic symmetries and invariance principles. Springer, New York (2005)

[Kin78] Kingman, J.F.C.: Uses of exchangeability. Ann. Probability 6(2), 183–197 (1978)

[Koz81] Kozen, D.: Semantics of probabilistic programs. J. Comp. System Sci. 22(3), 328–350 (1981)

[KTG+06] Kemp, C., Tenenbaum, J., Griffiths, T., Yamada, T., Ueda, N.: Learning systems of concepts with an infinite relational model. In: Proc. of the 21st Nat. Conf. on Artificial Intelligence (2006)

[Lau84] Lauritzen, S.L.: Extreme point models in statistics. Scand. J. Statist. 11(2), 65–91 (1984)

[Man09] Mansinghka, V.K.: Natively Probabilistic Computing. PhD thesis, Massachusetts Institute of Technology (2009)

[Mül99] Müller, N.Th.: Computability on random variables. Theor. Comput. Sci. 219(1-2), 287–299 (1999)

[Pfe01] Pfeffer, A.: IBAL: A probabilistic rational programming language. In: Proc. of the 17th Int. Joint Conf. on Artificial Intelligence, Morgan Kaufmann Publ., 733–740 (2001)

[PPT08] Park, S., Pfenning, F., Thrun, S.: A probabilistic language based on sampling functions. ACM Trans. Program. Lang. Syst. 31(1), 1–46 (2008)

[PR89] Pour-El, M.B., Richards, J.I.: Computability in analysis and physics. Springer, Berlin (1989)

[RMG+08] Roy, D.M., Mansinghka, V.K., Goodman, N.D., Tenenbaum, J.B.: A stochastic programming perspective on nonparametric Bayes. In: Nonparametric Bayesian Workshop, Int. Conf. on Machine Learning (2008)

[RN57] Ryll-Nardzewski, C.: On stationary sequences of random variables and the de Finetti's equivalence. Colloq. Math. 4, 149–156 (1957)

[Rog67] Rogers, Jr., H.: Theory of recursive functions and effective computability. McGraw-Hill, New York (1967)

[RP02] Ramsey, N., Pfeffer, A.: Stochastic lambda calculus and monads of probability distributions. Proc. of the 29th ACM SIGPLAN-SIGACT Symp. on Principles of Program. Lang., 154–165 (2002)

[RT09] Roy, D.M., Teh, Y.W.: The Mondrian process. In: Adv. in Neural Inform. Processing Syst. Vol. 21. (2009)

[Sch07] Schröder, M.: Admissible representations for probability measures. Math. Logic Q. 53(4-5), 431–445 (2007)

[Set94] Sethuraman, J.: A constructive definition of Dirichlet priors. Statistica Sinica 4, 639–650 (1994)

[Soa87] Soare, R.I.: Recursively enumerable sets and degrees. Springer, Berlin (1987)

[SS06] Schröder, M., Simpson, A.: Representing probability measures using probabilistic processes. J. Complex. 22(6), 768–782 (2006)

[TGG07] Teh, Y.W., Görür, D., Ghahramani, Z.: Stick-breaking construction for the Indian buffet process. In: Proc. of the 11th Conf. on A.I. and Stat. (2007)

[TJ07] Thibaux, R., Jordan, M.I.: Hierarchical beta processes and the Indian buffet process. In: Proc. of the 11th Conf. on A.I. and Stat. (2007)

[Wei99] Weihrauch, K.: Computability on the probability measures on the Borel sets of the unit interval. Theoret. Comput. Sci. 219(1–2), 421–437 (1999)

[Wei00] Weihrauch, K.: Computable analysis: an introduction. Springer, Berlin (2000)