Presenting convex sets of probability distributions by convex semilattices and unique bases

Filippo Bonchi, Ana Sokolova, Valeria Vignudelli

Presenting convex sets of probability distributions by convex semilattices and unique bases

Filippo Bonchi
University of Pisa, Italy

Ana Sokolova
University of Salzburg, Austria

Valeria Vignudelli
Univ Lyon, CNRS, ENS Lyon, UCB Lyon 1, LIP, France

Abstract

We prove that every finitely generated convex set of finitely supported probability distributions has a unique base. We apply this result to provide an alternative proof of a recent result: the algebraic theory of convex semilattices presents the monad of convex sets of probability distributions.

2012 ACM Subject Classification Theory of computation → Logic and verification; Theory of computation → Axiomatic semantics; Theory of computation → Categorical semantics

Keywords and phrases Convex sets of distributions monad, Convex semilattices, Unique base

Digital Object Identifier 10.4230/LIPIcs.CALCO.2021.13

Category (Co)algebraic pearls

Funding *Filippo Bonchi*: Supported by the Ministero dell'Università e della Ricerca of Italy under Grant No. 201784YSZ5, PRIN2017 – ASPRA (Analysis of Program Analyses).
Valeria Vignudelli: Supported by the French projects ANR-20-CE48-0005 QuaReMe and ANR-16-CE25-0011 REPAS, the European Research Council (ERC) under the European Union's Horizon 2020 programme (CoVeCe, grant agreement No 678157), the LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

1 Introduction

Models of computations exhibiting both nondeterministic and probabilistic behaviour are abundantly used in computed assisted verification [1, 12, 19, 5, 35, 11, 27], Artificial Intelligence [4, 17, 26], and studied from semantics perspective [14, 29, 13]. Indeed, probability is needed to quantitatively model uncertainty and belief, whereas nondeterminism enables modelling of incomplete information, unknown environment, implementation freedom, or concurrency.

Since several decades, computer scientists have found it convenient to exploit algebraic methods to analyse computing systems. From an algebraic perspective, the interplay of nondeterminism and probability has been posing some remarkable challenges [34, 18, 20, 16, 33, 24, 9, 31, 23]. Nevertheless, several fundamental algebraic structures have been identified and studied in depth.

In this paper we focus on one such structure, namely *convex sets of probability distributions*. These sets give rise to a monad that is well known in the literature and has found applications in several works [24, 9, 31, 33, 34, 16, 10, 22]. In recent work [3], we proved that this monad is presented by the algebraic theory of *convex semilattices*. In this paper, we provide an alternative proof based on a simple property: We show that every (finitely generated) convex set of distributions has a *unique base*.

This alternative proof technique is based on a categorical machinery together with a more syntax-based approach, which has already proven useful in extensions of the presentation results to the setting of metric spaces and quantitative equational theories [22, 21].

Synopsis: In Section 2, we show the unique base theorem. Our alternative proof of the presentation of the monad is based on exhibiting a monad map which is an isomorphism. We recall the relevant categorical notions in Section 3, and introduce a general recipe for building a monad map. In Section 4 we illustrate the monad of interest as well as the theory of convex semilattices, and in Section 5 we apply the recipe from Section 3 to build a monad map relating the monad and the theory. In Section 6 we prove that this monad map is an isomorphism, by relying on the unique base theorem to derive a normal-form argument.

2 A unique base theorem for convex sets of probability distributions

Given a set X, a probability distribution is a function $d\colon X \to [0,1]$ such that $\sum_{x\in X} d(x) = 1$. A probability distribution d is finitely supported if $d(x) \neq 0$ for finitely many x. We call $\mathcal{D}(X)$ the set of finitely suported probability distributions over X. A probability distribution $d \in \mathcal{D}(X)$ is a convex combination of the distributions $d_1, \dots d_n \in \mathcal{D}(X)$ if there exist $\alpha_1, \dots, \alpha_n \in [0,1]$ such that $\sum_i \alpha_i = 1$ and for all x, $d(x) = \sum_i \alpha_i d_i(x)$. Hereafter we will just write the latter condition as $d = \sum_i \alpha_i d_i$. The *convex closure* of a subset $S \subseteq \mathcal{D}(X)$, written $\mathrm{conv}(S)$, is the set of all the convex combinations of the distributions in S. A subset $S \subseteq \mathcal{D}(X)$ is *convex* if $S = \mathrm{conv}(S)$. A convex set is *finitely generated* if there exist $d_1, \dots, d_n \in \mathcal{D}(X)$ such that $S = \mathrm{conv}(\{d_1, \dots, d_n\})$. We let $C(X)$ denote the set of non-empty, finitely-generated convex sets of distributions over X. A *base* for $S \in C(X)$ is a set $\{d_1, \dots, d_n\}$ such that $S = \mathrm{conv}(\{d_1, \dots, d_n\})$ and for all $i \in 1 \dots n$, $d_i \notin \mathrm{conv}(\{d_j \mid j \neq i, 1 \leq j \leq n\})$.

▶ **Theorem 1.** *For every $S \in C(X)$, there exists a unique base.*

We show here a direct proof (Proof I) and an alternative proof using functional analysis tools and the strong theorem of Krein-Milman [25] (Proof II).

Proof I. Existence of the base comes from the property that S is finitely generated. In the rest of this section we prove uniqueness; namely if $\{d_1, \dots, d_n\}$ and $\{d'_1, \dots, d'_m\}$ are two bases for some $S \in \mathcal{D}(X)$, then $\{d_1, \dots, d_n\} = \{d'_1, \dots, d'_m\}$.

Let $\{d_1, \dots, d_n\}$ and $\{d'_1, \dots, d'_m\}$ be two bases for $S \in \mathcal{D}(X)$. Then for all $i \in 1 \dots n$ it holds $d_i \in \mathrm{conv}(\{d'_1, \dots, d'_m\})$ and for all $j \in 1 \dots m$ it holds $d'_j \in \mathrm{conv}(\{d_1, \dots, d_n\})$. By unfolding the definition of conv, this means that for all i there exist $\alpha_{i,j}$ such that $\sum_j \alpha_{i,j} = 1$ and for all j there exist $\alpha'_{j,i}$ such that $\sum_i \alpha'_{j,i} = 1$ and such that

$$d_i = \sum_{j\in\{1\dots m\}} \alpha_{i,j} d'_j \quad \text{and} \quad d'_j = \sum_{i\in\{1\dots n\}} \alpha'_{j,i} d_i. \tag{1}$$

Hence, for all i it holds

$$d_i = \sum_{j\in\{1\dots m\}} \alpha_{i,j} \left(\sum_{k\in\{1\dots n\}} \alpha'_{j,k} d_k \right) = \sum_{k\in\{1\dots n\}} \left(\sum_{j\in\{1\dots m\}} \alpha_{i,j}\alpha'_{j,k} \right) d_k$$

where the fist equality follows by replacing the d'_j in the left equation in (1) with the one in the right equation in (1). So we have

$$d_i = \left(\sum_{j\in\{1\dots m\}} \alpha_{i,j}\alpha'_{j,i} \right) d_i + \sum_{k\in\{1\dots n\}\setminus\{i\}} \left(\sum_{j\in\{1\dots m\}} \alpha_{i,j}\alpha'_{j,k} \right) d_k \tag{2}$$

We now prove by contradiction that

$$\sum_{j\in\{1\dots m\}} \alpha_{i,j}\alpha'_{j,i} = 1 \text{ for all } i \in \{1\dots n\} \tag{3}$$

Let $i \in \{1\dots n\}$ and let $\beta_i = \sum_{j\in\{1\dots m\}} \alpha_{i,j}\alpha'_{j,i}$. If $\beta_i \neq 1$, then by (2) we have

$$d_i = \beta_i d_i \,+\, (1-\beta_i) \sum_{k\in\{1\dots n\}\setminus\{i\}} \left(\frac{\sum_{j\in\{1\dots m\}} \alpha_{i,j}\alpha'_{j,k}}{1-\beta_i} \right) d_k$$

and from this we derive

$$d_i = \sum_{k\in\{1\dots n\}\setminus\{i\}} \left(\frac{\sum_{j\in\{1\dots m\}} \alpha_{i,j}\alpha'_{j,k}}{1-\beta_i} \right) d_k$$

This means that d_i is expressible as a convex combination of $\{d_1\dots,d_n\}\setminus\{d_i\}$, which contradicts the hypothesis that $\{d_1\dots,d_n\}$ is a base. Hence, $\beta_i = 1$, which proves (3).

From (3) and (2) we derive that for all $k \in \{1\dots n\}\setminus\{i\}$, $\sum_{j\in\{1\dots m\}} \alpha_{i,j}\alpha'_{j,k} = 0$. Since all the summands are non-negative, this entails that

$$\alpha_{i,j}\alpha'_{j,k} = 0 \text{ for all } i \in \{1\dots n\},\, k \in \{1\dots n\}\setminus\{i\} \text{ and } j \in \{1\dots m\}. \tag{4}$$

By reasoning in the same way, we obtain the following

$$\alpha'_{j,i}\alpha_{i,l} = 0 \text{ for all } j \in \{1\dots m\},\, l \in \{1\dots m\}\setminus\{j\} \text{ and } i \in \{1\dots n\}. \tag{5}$$

We now prove that for all i there exists one j such that $\alpha_{i,j} = 1$. As $\sum_j \alpha_{i,j} = 1$, there is at least one j such that $\alpha_{i,j} > 0$. By this and (4) one has that for all $k \in \{1\dots n\}\setminus\{i\}$, $\alpha'_{j,k} = 0$. Since $\sum_{k\in\{1\dots n\}} \alpha'_{j,k} = 1$, we have that $\alpha'_{j,i} = 1$. Hence we derive by (5) that $\alpha_{i,l} = 0$ for all $l \in \{1\dots m\}\setminus\{j\}$. Since $\sum_{l\in\{1\dots m\}} \alpha_{i,l} = 1$, we have $\alpha_{i,j} = 1$.

Using this fact, we conclude by the left equation in (1) that for every i there exists one j such that $d_i = d'_j$. Hence, we have $\{d_1,\dots,d_n\} \subseteq \{d'_1,\dots,d'_m\}$. The opposite inclusion follows symmetrically. ◀

Proof II. Let $S \in C(X)$. Note that then S is a subset of $\mathcal{D}(X) \subseteq \mathbb{R}^X$ and hence a subset of a locally convex topological vector space ($\mathbb{R}^X$ with the product topology). Consider the family $\mathcal{B} = \{B \subseteq S \mid S = \mathrm{conv}(B)\}$. It is obvious that B is minimal in $\mathcal{B}$ if and only if no element $d \in B$ satisfies $d \in \mathrm{conv}(B\setminus\{d\})$. We now show that $\mathcal{B}$ contains a smallest element.

First, note that for all $B \in \mathcal{B}$, $\mathrm{Ext}(S) \subseteq B$, with $\mathrm{Ext}(S)$ being the set of extreme points of S. Indeed, let $d \in \mathrm{Ext}(S)$. Then $d \in S$ and can be written as $d = \sum_{d_i\in B} p_i d_i = p_i \cdot d_i + (1-p_i)\cdot e$ for some $p_i \neq 0$ and $e \in S$, and hence by extremality of d we have $d = d_i = e$ yielding $d \in B$.

Next, we show that $S = \mathrm{conv}(\mathrm{Ext}(S))$, which means that $\mathrm{Ext}(S) \in \mathcal{B}$ and hence together with $\mathrm{Ext}(S) \subseteq B$ shows that $\mathrm{Ext}(S)$ is the smallest element of $\mathcal{B}$. This smallest element $\mathrm{Ext}(S)$ is the unique base of S. Pick a finite $B_0 = \{d_1,\dots,d_n\} \in \mathcal{B}$. Then $S = \Phi(\Delta_n)$ for

$$\Delta_n = \{(x_1,\dots,x_n) \in \mathbb{R}^n \mid x_i \in [0,1], \sum_i x_i = 1\}$$

and $\Phi\colon \mathbb{R}^n \to \mathbb{R}^X$ given by $\Phi(x_1,\dots,x_n) = \sum_i x_i d_i$. Note that Δ_n is compact, by Heine-Borel, as it is a closed and bounded subset of $\mathbb{R}^n$, and Φ is continuous, since we are in a topological vector space and hence algebraic operations are continuous. As a consequence, S

is compact as a continuous image of a compact set. Now, Krein-Milman applies, yielding that $S = \overline{\text{conv}}(\text{Ext}(S))$ with $\overline{\text{conv}}$ denoting the closed convex hull and hence

$$S = \overline{\text{conv}}(\text{Ext}(S)) = \text{conv}(\text{Ext}(S))$$

since by the same argument as above $\text{conv}(\text{Ext}(S))$ is compact and hence closed. ◀

Instead of the Krein-Milman theorem, one could use in this proof its predecessor from classical convex analysis in $\mathbb{R}^n$, e.g. [32, Theorem 18.5]. The reason is that since we deal with finitely generated convex subsets of finitely supported distributions, such subsets are actually elements of $C(X)$ for a finite set X.

3 Monads and presentations

Theorem 1 states the existence of a unique base for every finitely generated convex set of probability distributions. In the remainder of this paper, we exploit this result to illustrate an alternative proof of Theorem 4 in [3] that provides a presentation of the monad C [24, 9, 31, 33, 34, 16]. In Section 4, we recall the monad as well as its presentation given in [3]. In this section, we recall some basic facts about monads and presentations.

A *monad* on **Sets** is a functor $\mathcal{M}\colon \mathbf{Sets} \to \mathbf{Sets}$ together with two natural transformations: a unit $\eta\colon \text{Id} \Rightarrow \mathcal{M}$ and multiplication $\mu\colon \mathcal{M}^2 \Rightarrow \mathcal{M}$ that satisfy the laws $\mu \circ \eta\mathcal{M} = \mu \circ \mathcal{M}\eta = id$ and $\mu \circ \mathcal{M}\mu = \mu \circ \mu\mathcal{M}$.

A *monad map* from a monad $\mathcal{M}$ to a monad $\hat{\mathcal{M}}$ is a natural transformation $\sigma\colon \mathcal{M} \Rightarrow \hat{\mathcal{M}}$ that makes the following diagrams commute, with η, μ and $\hat{\eta}, \hat{\mu}$ denoting the unit and multiplication of $\mathcal{M}$ and $\hat{\mathcal{M}}$, respectively, and $\sigma\sigma = \sigma \circ \mathcal{M}\sigma = \hat{\mathcal{M}}\sigma \circ \sigma_{\mathcal{M}}$.

$$\begin{array}{ccc} X & \xrightarrow{\eta} & \mathcal{M}X \\ & \searrow^{\hat{\eta}} & \downarrow{\scriptstyle \sigma} \\ & & \hat{\mathcal{M}}X \end{array} \qquad \begin{array}{ccc} \mathcal{M}\mathcal{M}X & \xrightarrow{\sigma\sigma} & \hat{\mathcal{M}}\hat{\mathcal{M}}X \\ {\scriptstyle \mu}\downarrow & & \downarrow{\scriptstyle \hat{\mu}} \\ \mathcal{M}X & \xrightarrow[\sigma]{} & \hat{\mathcal{M}}X \end{array}$$

If $\sigma\colon \mathcal{M}X \to \hat{\mathcal{M}}X$ is an iso, the two monads are isomorphic.

An important example of monad is provided by the *free monad of terms*. Given a signature Σ, namely a set of operation symbols equipped with an arity, the free monad $T_\Sigma\colon \mathbf{Sets} \to \mathbf{Sets}$ of terms over Σ maps a set X to the set of all Σ-terms with variables in X, and $f\colon X \to Y$ to the function that maps a term over X to a term over Y obtained by substitution according to f. The unit maps a variable in X to itself, and the multiplication is term composition.

Given a set of axioms E over Σ-terms, one can define the smallest congruence generated by the axioms, denoted by $=_E$. Hereafter we write $[t]_E$ for the $=_E$-equivalence class of the Σ-term t and $T_{\Sigma,E}(X)$ for the set of E-equivalence classes of Σ-terms with variables in X. The assignment $X \mapsto T_{\Sigma,E}(X)$ gives rise to a functor $T_{\Sigma,E}\colon \mathbf{Sets} \to \mathbf{Sets}$ where the behaviour on functions is defined as for T_Σ. Such functor carries the structure of a monad: the unit $\eta^E\colon \text{Id} \Rightarrow T_{\Sigma,E}$ and the multiplication $\mu^E\colon T_{\Sigma,E}T_{\Sigma,E} \Rightarrow T_{\Sigma,E}$ are defined as $\eta^E(x) = [x]_E$ and $\mu^E[t\{[t_i]_E/x_i\}]_E = [t\{t_i/x_i\}]_E$.

An *algebraic theory* is a pair (Σ, E) of signature Σ and a set of equations E. We say that (Σ, E) provides a *presentation* for a monad $\mathcal{M}$ if $T_{\Sigma,E}$ is isomorphic to $\mathcal{M}$.

We next introduce several monads on **Sets** together with their presentations.

Nondeterminism. The non-empty finite powerset monad $\mathcal{P}_{ne}$ maps a set X to the set of non-empty finite subsets $\mathcal{P}_{ne}X = \{U \mid U \subseteq X,\ U \text{ is finite and non-empty}\}$ and a function

$f\colon X \to Y$ to $\mathcal{P}_{ne}f\colon \mathcal{P}_{ne}X \to \mathcal{P}_{ne}Y$, $\mathcal{P}_{ne}f(U) = \{f(u) \mid u \in U\}$. The unit η of $\mathcal{P}_{ne}$ is given by singleton, i.e., $\eta(x) = \{x\}$, and the multiplication μ is given by union, i.e., $\mu(S) = \bigcup_{U \in S} U$ for $S \in \mathcal{P}_{ne}\mathcal{P}_{ne}X$.

Let Σ_N be the signature consisting of a binary operation $\oplus$. Let E_N be the following set of axioms, the axioms of semilattice:

$$(x \oplus y) \oplus z \overset{(A)}{=} x \oplus (y \oplus z) \qquad x \oplus y \overset{(C)}{=} y \oplus x \qquad x \oplus x \overset{(I)}{=} x$$

It is easy to show that the algebraic theory (Σ_N, E_N) provides a presentation for the monad $\mathcal{P}_{ne}$, in the sense that there exists an isomorphism of monads $\iota^N\colon T_{\Sigma_N,E_N} \Rightarrow \mathcal{P}_{ne}$.

Probability. The finitely supported probability distribution monad $\mathcal{D}$ is defined, for a set X and a function $f\colon X \to Y$, as $\mathcal{D}X = \{\varphi\colon X \to [0,1] \mid \sum_{x\in X} \varphi(x) = 1, \text{supp}(\varphi) \text{ is finite}\}$ and $\mathcal{D}f(\varphi)(y) = \sum_{x \in f^{-1}(y)} \varphi(x)$. The unit of $\mathcal{D}$ is given by a Dirac distribution $\eta(x) = \delta_x = (x \mapsto 1)$ for $x \in X$ and the multiplication by $\mu(\Phi)(x) = \sum_{\varphi \in \text{supp}(\Phi)} \Phi(\varphi) \cdot \varphi(x)$ for $\Phi \in \mathcal{D}\mathcal{D}X$. We sometimes write $\sum_{i\in I} p_i x_i$ for a distribution φ with $\text{supp}(\varphi) = \{x_i \mid i \in I\}$ and $\varphi(x_i) = p_i$.

Let Σ_P be the signature consisting of a binary operation $+_p$ for all $p \in (0,1)$. Let E_P be the following set of axioms, the axioms of a barycentric algebra, also called convex algebra:[1]

$$(x +_q y) +_p z \overset{(A_p)}{=} x +_{pq} (y +_{\frac{p(1-q)}{1-pq}} z) \qquad x +_p y \overset{(C_p)}{=} y +_{1-p} x \qquad x +_p x \overset{(I_p)}{=} x$$

The algebraic theory (Σ_P, E_P) provides a presentation for the monad $\mathcal{D}$ [30, 28, 7, 8, 15], in the sense that there exists an isomorphism of monads $\iota^P\colon T_{\Sigma_P,E_P} \Rightarrow \mathcal{D}$.

3.1 A well known recipe for constructing monad morphisms

To prove that an algebraic theory (Σ, E) presents a monad $\mathcal{M}$, one has to provide $\iota\colon T_{\Sigma,E} \Rightarrow \mathcal{M}$ that (a) is a monad map and (b) is an isomorphism. While the proof of (b) often requires some specific normal form arguments, the proof of (a) can be significantly simplified by using some standard categorical machinery.

In this section, we illustrate a well known recipe which allows for constructing a monad map $\iota\colon T_{\Sigma,E} \Rightarrow \mathcal{M}$ in a principled way. We begin by recalling Eilenberg-Moore algebras.

To each monad $\mathcal{M}$, one associates the Eilenberg-Moore category $\text{EM}(\mathcal{M})$ of $\mathcal{M}$-algebras. Objects of $\text{EM}(\mathcal{M})$ are pairs $\mathbb{A} = (A, a)$ of a set $A \in \mathbf{Sets}$ and a map $a\colon \mathcal{M}A \to A$, making the first two diagrams below commute.

$$\begin{array}{ccc} A \xrightarrow{\eta} \mathcal{M}A & \mathcal{M}^2A \xrightarrow{\mathcal{M}a} \mathcal{M}A & \mathcal{M}A \xrightarrow{\mathcal{M}h} \mathcal{M}B \\ \searrow \downarrow a & \mu\downarrow \quad \downarrow a & a\downarrow \quad \downarrow b \\ A & \mathcal{M}A \xrightarrow[a]{} A & A \xrightarrow[h]{} B \end{array}$$

A homomorphism from an algebra $\mathbb{A} = (A, a)$ to an algebra $\mathbb{B} = (B, b)$ is a map $h\colon A \to B$ between the underlying sets making the third diagram above commute.

It is well known that, when $\mathcal{M}$ is the monad $T_{\Sigma,E}$ for some algebraic theory (Σ, E), $\text{EM}(\mathcal{M})$ is isomorphic to the category $\text{Alg}(\Sigma, E)$ of (Σ, E)-algebras and their morphisms. A Σ-algebra (X, Σ_X) consist of a set X together with a set Σ_X of operations $\hat{o}_X\colon X^n \to X$, one for each operation symbol $o \in \Sigma$ of arity n. A (Σ, E)-algebra is a Σ-algebra where all the

[1] There is another equivalent presentation for convex algebras with a signature involving arbitrary convex combinations and two axioms, projection and barycenter. In this paper we will mainly use the binary convex operations.

equations in E hold. A homomorphism h from a (Σ, E)-algebra (X, Σ_X) to a (Σ, E)-algebra (Y, Σ_Y) is a function $h\colon X \to Y$ that commutes with the operations, i.e., $h \circ \hat{o}_X = \hat{o}_Y \circ h^n$ for all n-ary $o \in \Sigma$.

For instance, (Σ_N, E_N)-algebras are semilattices, namely a set X equipped with a binary operation $\hat{\oplus}_X$ that is associative, commutative and idempotent. A semilattice homomorphism is a function $h\colon X \to Y$ such that $h(x_1 \hat{\oplus}_X x_2) = h(x_1) \hat{\oplus}_Y h(x_2)$ for all $x_1, x_2 \in X$.

Now we can display an abstract recipe for constructing a monad map $\iota\colon T_{\Sigma,E} \Rightarrow \mathcal{M}$, which consists of three steps:

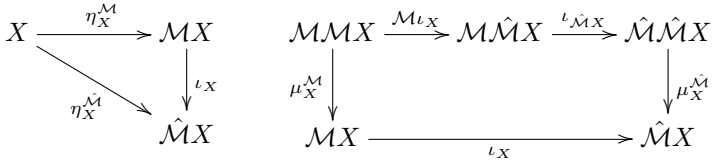
(A) For each set X, provide $\mathcal{M}X$ with the structure of a (Σ, E)-algebra, namely functions $\hat{o}_X\colon (\mathcal{M}X)^n \to \mathcal{M}X$ for each $o \in \Sigma$, that satisfy the equations in E;
(B) Prove that for each function $f\colon X \to Y$, $\mathcal{M}f$ is a (Σ, E)-algebra homomorphism;
(C) Prove that for each set X, $\mu^{\mathcal{M}}_X\colon \mathcal{M}\mathcal{M}X \to \mathcal{M}X$ is a (Σ, E)-algebra homomorphism.

By the correspondence of (Σ, E)-algebras and Eilenberg-Moore algebras for $T_{\Sigma,E}$ and (A), we obtain a $T_{\Sigma,E}$-algebra $\alpha^\sharp_X\colon T_{\Sigma,E}\mathcal{M}X \to \mathcal{M}X$ for each set X. These $\alpha^\sharp_X$ give rise to a natural transformation $\alpha^\sharp\colon T_{\Sigma,E}\mathcal{M} \Rightarrow \mathcal{M}$ by (B) and the correspondence of (Σ, E)-homomorphisms and $T_{\Sigma,E}$-homomorphisms. The monad morphism $\iota\colon T_{\Sigma,E} \Rightarrow \mathcal{M}$ is then obtained by (C) and the following theorem[2].

▶ **Theorem 2.** *Let $(\mathcal{M}, \eta^{\mathcal{M}}, \mu^{\mathcal{M}})$ and $(\hat{\mathcal{M}}, \eta^{\hat{\mathcal{M}}}, \mu^{\hat{\mathcal{M}}})$ be two monads. Let $\alpha^\sharp\colon \mathcal{M}\hat{\mathcal{M}} \Rightarrow \hat{\mathcal{M}}$ be a natural transformation such that $\alpha^\sharp_X\colon \mathcal{M}\hat{\mathcal{M}}X \to \hat{\mathcal{M}}X$ is an Eilenberg-Moore algebra for $\mathcal{M}$ and that $\mu^{\hat{\mathcal{M}}}_X\colon \hat{\mathcal{M}}\hat{\mathcal{M}}X \to \hat{\mathcal{M}}X$ is an $\mathcal{M}$-algebra morphism from $(\hat{\mathcal{M}}\hat{\mathcal{M}}X, \alpha^\sharp_{\hat{\mathcal{M}}X})$ to $(\hat{\mathcal{M}}X, \alpha^\sharp_X)$. Then the following is a monad map:*

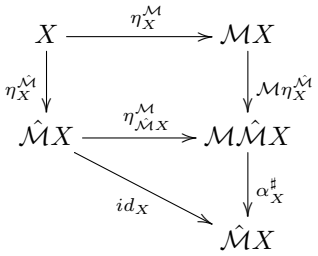
$$\iota := \mathcal{M} \xRightarrow{\mathcal{M}\eta^{\hat{\mathcal{M}}}} \mathcal{M}\hat{\mathcal{M}} \xRightarrow{\alpha^\sharp} \hat{\mathcal{M}} .$$

Proof. In order to prove that ι is a monad map, we need to prove that the following two diagrams commute.



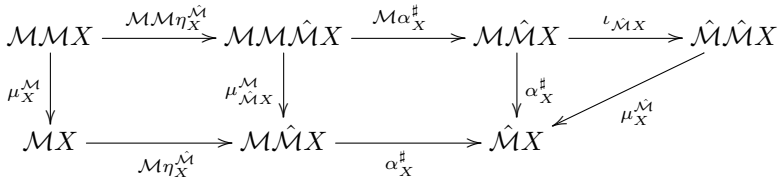
(6)

For the diagram on the left, it is enough to recall that $\iota = \alpha^\sharp \circ \mathcal{M}\eta^{\hat{\mathcal{M}}}$ and observe that the following diagram commutes: the top square commutes by naturality of $\eta^{\mathcal{M}}$ and the bottom triangle commutes since $\alpha^\sharp_X$ is an Eilenberg Moore algebra for $\mathcal{M}$.



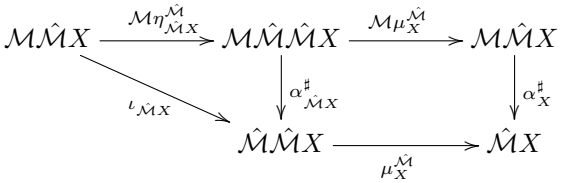
[2] This theorem is known, but it is not easy to find an original reference for it. We thank Jurriaan Rot for recalling the theorem and the proof with us.

In order to prove the commutation of the diagram on the right in (6), by $\iota = \alpha^\sharp \circ \mathcal{M}\eta^{\hat{\mathcal{M}}}$ it is enough to prove that the following commutes:



The left square commutes by naturality of $\mu^{\mathcal{M}}$. The central square commutes since $\alpha^\sharp_X$ is an Eilenberg-Moore algebra for $\mathcal{M}$. It remains to prove that the right triangle commutes.

First, observe that the diagram below commutes: the left triangle commutes by definition of ι, and the right square commutes by the assumption that $\mu^{\hat{\mathcal{M}}}_X$ is an $\mathcal{M}$-algebra morphism.



This completes the proof as $\mathcal{M}\mu^{\hat{\mathcal{M}}}_X \circ \mathcal{M}\eta^{\hat{\mathcal{M}}}_{\hat{\mathcal{M}}X} = \mathcal{M}(\mu^{\hat{\mathcal{M}}}_X \circ \eta^{\hat{\mathcal{M}}}_{\hat{\mathcal{M}}X}) = \mathcal{M}(id_{\hat{\mathcal{M}}X}) = id_{\mathcal{M}\hat{\mathcal{M}}X}$. ◀

The function $\iota_X : T_{\Sigma,E}X \to \mathcal{M}X$ obtained by the above recipe can be inductively defined for all $x \in X$, $t_1, \ldots, t_n \in T_\Sigma X$ and n-ary operations o in Σ as follows.

$$\iota_X([x]_E) = \eta^{\mathcal{M}}_X(x) \qquad \iota_X([o(t_1, \ldots, t_n)]_E) = \hat{o}_X(\iota_X[t_1]_E, \ldots, \iota_X[t_n]_E). \tag{7}$$

The fact that the functions $\hat{o}_X$ form a (Σ, E)-algebra ensures that ι is a well defined function, namely if $t =_E t'$, then $\iota([t]_E) = \iota([t']_E)$.

We conclude this section by shortly illustrating how to apply the above recipe to the monad for nondeterminism and the one for probability discussed above. To construct a monad map $\iota^N : T_{\Sigma_N,E_N} \Rightarrow \mathcal{P}_{ne}$, we define for all sets X the binary function $\hat{\oplus} : \mathcal{P}_{ne}(X) \times \mathcal{P}_{ne}(X) \to \mathcal{P}_{ne}(X)$ as the union $\cup$. This is associative, commutative and idempotent, so the axioms in E_N are satisfied, or in other words, this forms a semilattice. This corresponds to point (A) of the recipe. It is not difficult to check (B) and (C). The resulting monad map is defined for all sets X as

$$\iota^N_X([x]_{E_N}) = \{x\} \qquad \iota^N_X([t_1 \oplus t_2]_{E_N}) = \iota^N_X([t_1]_{E_N}) \cup \iota^N_X([t_2]_{E_N}).$$

To construct the monad map $\iota^P : T_{\Sigma_P,E_P} \Rightarrow \mathcal{D}$, we define for all $p \in (0,1)$ and all sets X the binary function $\hat{+}_p : \mathcal{D}(X) \times \mathcal{D}(X) \to \mathcal{D}(X)$ as $d_1 \hat{+}_p d_2 = p\, d_1 + (1-p) d_2$. One can check that the three axioms in E_P are satisfied (distributions form a convex algebra), and that points (B) and (C) of the recipe hold. The resulting monad map is defined for all sets X as

$$\iota^P_X([x]_{E_P}) = \delta_x \qquad \iota^P_X([t_1 +_p t_2]_{E_P}) = p\, \iota^P_X([t_1]_{E_P}) + (1-p)\iota^P_X([t_2]_{E_P}). \tag{8}$$

4 The monad for nondeterminism and probability

In this section, we recall the monad for nondeterminism and probability, its presentation, and we illustrate some interesting properties.

The monad $C\colon \mathbf{Sets} \to \mathbf{Sets}$ maps a set X into CX, namely the set of non-empty, finitely-generated convex subsets of distributions on X (as defined in Section 2). For a function $f\colon X \to Y$, $Cf\colon CX \to CY$ is given by $Cf(S) = \{\mathcal{D}f(d) \mid d \in S\}$. The unit of C is $\eta\colon X \to CX$ given by $\eta(x) = \{\delta_x\}$. The multiplication $\mu\colon CCX \to CX$ of C can be expressed in concrete terms as follows [16]. Given $S \in CCX$,

$$\mu(S) = \bigcup_{\Phi \in S} \Big\{ \sum_{U \in \operatorname{supp}\Phi} \Phi(U) \cdot d \mid d \in U \Big\}.$$

Let Σ be the signature $\Sigma_N \cup \Sigma_P$. Let E be the sets of axioms consisting of E_N, E_p and the following distributivity axiom:

$$(x \oplus y) +_p z \overset{(D)}{=} (x +_p z) \oplus (y +_p z)$$

This theory (Σ, E) is the algebraic theory of *convex semilattices*, introduced in [3].

▶ **Theorem 3.** *(Σ, E) is a presentation of the monad C.*

The above theorem has been proved in [3]. In the remainder of this paper, we will provide an alternative proof of this fact by exploiting the unique base theorem (Theorem 1).

We begin by observing that the assignment $S \mapsto \operatorname{conv}(S)$ gives rise to a natural transformation $\operatorname{conv}\colon \mathcal{P}_{ne}\mathcal{D} \Rightarrow C$ [20, 2]. Theorem 1 provides a way of going backward, from C to $\mathcal{P}_{ne}\mathcal{D}$: we call $\mathrm{UB}_X\colon CX \to \mathcal{P}_{ne}\mathcal{D}X$ the function assigning to each convex subset S its unique base. However such UB_X does not give rise to a natural transformation, in the sense that the diagram on the left in (9) only commutes laxly for arbitrary functions $f\colon X \to Y$.

$$\begin{array}{ccc} CX & \xrightarrow{Cf} & CY \\ {\scriptstyle \mathrm{UB}_X}\downarrow & \swarrow & \downarrow{\scriptstyle \mathrm{UB}_Y} \\ \mathcal{P}_{ne}\mathcal{D}X & \xrightarrow[\mathcal{P}_{ne}\mathcal{D}f]{} & \mathcal{P}_{ne}\mathcal{D}Y \end{array} \qquad \begin{array}{ccc} CX & \xrightarrow{Cf} & CY \\ {\scriptstyle \mathrm{UB}_X}\downarrow & & \uparrow{\scriptstyle \mathrm{conv}_Y} \\ \mathcal{P}_{ne}\mathcal{D}X & \xrightarrow[\mathcal{P}_{ne}\mathcal{D}f]{} & \mathcal{P}_{ne}\mathcal{D}Y \end{array} \tag{9}$$

It holds that $\mathrm{UB}_Y \circ Cf \subseteq \mathcal{P}_{ne}\mathcal{D}f \circ \mathrm{UB}_X$ but not the other way around, as shown by the next example.

▶ **Example 4.** Let $X = \{x, y, z\}$, $Y = \{a, b\}$ and $f\colon X \to Y$ be the function mapping both x and y to a and z to b. Consider the set $S = \{\frac{1}{2}x + \frac{1}{2}y,\ \frac{1}{2}x + \frac{1}{2}z,\ \delta_z\}$: this set is a base since none of its element can be expressed as convex combination of the others. However, the set $\mathcal{P}_{ne}\mathcal{D}f(S) = \{\delta_a,\ \frac{1}{2}a + \frac{1}{2}b,\ \delta_b\}$ is not a base since $\frac{1}{2}a + \frac{1}{2}b$ can be expressed as a linear combination of δ_a and δ_b. Now, by taking the convex set $\operatorname{conv}(S) \in CX$ one can easily see that $\mathrm{UB}_Y \circ Cf \not\supseteq \mathcal{P}_{ne}\mathcal{D}f \circ \mathrm{UB}_X$. Indeed $\mathcal{P}_{ne}\mathcal{D}f \circ \mathrm{UB}_X(\operatorname{conv}(S)) = \mathcal{P}_{ne}\mathcal{D}f(S) = \{\delta_a,\ \frac{1}{2}a + \frac{1}{2}b,\ \delta_b\}$, while $\mathrm{UB}_Y \circ Cf(\operatorname{conv}(S)) = \{\delta_a,\ \delta_b\}$ since $Cf(\operatorname{conv}(S)) = \operatorname{conv}(\mathcal{D}f(S))$ by Lemma 5 below.

Interestingly enough, while the diagram on the left in (9) does not commute, the diagram on the right in (9) does. This is closely related to Lemma 37 from [3], which provides a slightly different formulation. Below, we illustrate a proof: to simplify the notation of the natural transformations, we avoid to specify the set X whenever it is clear from the context.

▶ **Lemma 5.** *Let $S \in C(X)$ and $f : X \to Y$. Then $Cf(S) = \operatorname{conv}(\{\mathcal{D}f(d) \mid d \in \mathrm{UB}(S)\})$.*

Proof. We prove $Cf(S) \subseteq \operatorname{conv}(\bigcup_{d \in \mathrm{UB}(S)} \{\mathcal{D}f(d)\})$. Let $e \in Cf(S)$. Then $e = \mathcal{D}f(d)$ for some $d \in S$, which implies that d is a convex combination of elements of $\mathrm{UB}(S)$, that is, $d = \sum_i p_i \cdot d_i$ with $d_i \in \mathrm{UB}(S)$ for all i. Hence, $e = \sum_i p_i \cdot \mathcal{D}f(d_i) \in \operatorname{conv}(\bigcup_{d \in \mathrm{UB}(S)} \{\mathcal{D}f(d)\})$.

For the opposite inclusion, let $e \in \mathrm{conv}(\bigcup_{d\in \mathrm{UB}(S)} \{\mathcal{D}f(d)\})$. Hence, $e = \sum_i p_i \cdot \mathcal{D}f(d_i)$ with $d_i \in \mathrm{UB}(S)$ for all i. We have $\sum_i p_i \cdot \mathcal{D}f(d_i) = \mathcal{D}f(\sum_i p_i \cdot d_i)$ and, from $\sum_i p_i \cdot d_i \in S$, we conclude $e \in Cf(S)$.

◀

5 The monad map $\iota\colon T_{\Sigma,E} \Rightarrow C$

In this section we apply the standard recipe from Section 3.1 to construct a monad map $\iota\colon T_{\Sigma,E} \Rightarrow C$.

For this aim, we first recall two well-known operations on convex sets: the convex union $\oplus\colon C(X) \times C(X) \to C(X)$ defined for all $S_1, S_2 \in C(X)$ as

$$S_1 \oplus S_2 = \mathrm{conv}(S_1 \cup S_2)$$

and, for all $p \in (0,1)$, the Minkowski sum $+_p\colon C(X) \times C(X) \to C(X)$ defined as

$$S_1 +_p S_2 = \{d \mid d = pd_1 + (1-p)d_2 \text{ for some } d_1 \in S_1 \text{ and } d_2 \in S_2\}.$$

Points (A) and (B) of the recipe hold by the following result from [3, Lemma 38].

▶ **Lemma 6.** *With the above defined operations $(CX, \oplus, +_p)$ is a convex semilattice. Moreover, for a map $f\colon X \to Y$, the map $Cf\colon CX \to CY$ is a convex semilattice homomorphism from $(CX, \oplus, +_p)$ to $(CY, \oplus, +_p)$.* ◀

The following lemma proves point (C) explicitly, namely that μ is a (Σ, E)-homomorphism.[3]

▶ **Lemma 7.** *For all $S_1, S_2 \in CC(X)$, it holds that:*

1. $\mu(S_1 \oplus S_2) = \mu(S_1) \oplus \mu(S_2)$
2. $\mu(S_1 +_p S_2) = \mu(S_1) +_p \mu(S_2)$

Proof. Through this proof, we will often use the following key observation: $d \in \mu(S)$ iff

$$\exists \Phi \in S \text{ such that } d = \sum_{U\in \mathrm{supp}(\Phi)} \Phi(U)\cdot f(U) \text{ , for } f\colon \mathrm{supp}(\Phi) \to \mathcal{D}(X) \text{ such that } f(U) \in U.$$

1. We first prove the inclusion $\mu(S_1) \oplus \mu(S_2) \subseteq \mu(S_1 \oplus S_2)$. As $S_1 \subseteq S_1 \oplus S_2$ we derive that

$$\begin{array}{rcl} \mu(S_1) & \stackrel{\text{def}}{=} & \bigcup_{\Phi\in S_1} \{\sum_{U\in \mathrm{supp}(\Phi)} \Phi(U)\cdot d \mid d \in U\} \\ & \subseteq & \bigcup_{\Phi\in S_1\oplus S_2} \{\sum_{U\in \mathrm{supp}(\Phi)} \Phi(U)\cdot d \mid d \in U\} \\ & \stackrel{\text{def}}{=} & \mu(S_1 \oplus S_2) \end{array} \tag{10}$$

Symmetrically, by $S_2 \subseteq S_1 \oplus_p S_2$ we have

$$\begin{array}{rcl} \mu(S_2) & \stackrel{\text{def}}{=} & \bigcup_{\Phi\in S_2} \{\sum_{U\in \mathrm{supp}(\Phi)} \Phi(U)\cdot d \mid d \in U\} \\ & \subseteq & \bigcup_{\Phi\in S_1\oplus S_2} \{\sum_{U\in \mathrm{supp}(\Phi)} \Phi(U)\cdot d \mid d \in U\} \\ & \stackrel{\text{def}}{=} & \mu(S_1 \oplus S_2) \end{array} \tag{11}$$

[3] In [3], we show that $(CX, \oplus, +_p)$ is the free convex semilattice generated by X and then prove that $\mu = id^{\#}_{CX}$, see [3, Lemma 41]. An implicit consequence of this is that μ is the unique homomorphism from the free convex semilattice generated by CX to the free convex semilattice generated by X that extends the identity map on CX.

Hence,

$$\begin{aligned}
&\mu(S_1) \oplus \mu(S_2) \\
&\quad = \operatorname{conv}\Big(\bigcup_{\Phi \in S_1} \Big\{ \sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot d \mid d \in U \Big\} \cup \bigcup_{\Phi \in S_2} \Big\{ \sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot d \mid d \in U \Big\} \Big) \\
&\quad \subseteq \operatorname{conv}(\mu(S_1 \oplus S_2)) && \text{(by (10), (11))} \\
&\quad = \mu(S_1 \oplus S_2) && \text{(by } \mu(S_1 \oplus S_2) \text{ a convex set)}
\end{aligned}$$

We then prove the inclusion $\mu(S_1 \oplus S_2) \subseteq \mu(S_1) \oplus \mu(S_2)$. Take $d \in \mu(S_1 \oplus S_2)$. Then there is a $\Phi \in S_1 \oplus S_2$ such that $d = \sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot f(U)$, with $f : \operatorname{supp}(\Phi) \to \mathcal{D}(X)$ such that $f(U) \in U$. As Φ is a convex combination of distributions in $S_1 \cup S_2$, we have $\Phi = \sum_i p_i \cdot \Phi_i$ with $\Phi_i \in (S_1 \cup S_2)$ for all i. Then for all $x \in X$ we have

$$\begin{aligned}
\sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot f(U)(x) &= \sum_{U \in \cup_i \operatorname{supp}(\Phi_i)} \Big((\sum_i p_i \cdot \Phi_i)(U) \cdot f(U)(x) \Big) \\
&= \sum_{U \in \cup_i \operatorname{supp}(\Phi_i)} \Big(\sum_i p_i \cdot \Phi_i(U) \cdot f(U)(x) \Big) \\
&= \sum_i p_i \cdot \Big(\sum_{U \in \cup_i \operatorname{supp}(\Phi_i)} \Phi_i(U) \cdot f(U)(x) \Big) \\
&= \sum_i p_i \cdot \Big(\sum_{U \in \operatorname{supp}(\Phi_i)} \Phi_i(U) \cdot f(U)(x) \Big)
\end{aligned}$$

Hence, the result follows as

$$\begin{aligned}
d &= \sum_i p_i \cdot \Big(\sum_{U \in \operatorname{supp}(\Phi_i)} \Phi_i(U) \cdot f(U) \Big) \\
&\in \operatorname{conv}\Big(\bigcup_{\Phi \in (S_1 \cup S_2)} \Big\{ \sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot d \mid d \in U \Big\} \Big) \\
&= \mu(S_1 \oplus S_2).
\end{aligned}$$

2. We first prove $\mu(S_1) +_p \mu(S_2) \subseteq \mu(S_1 +_p S_2)$. Let $d \in \mu(S_1) +_p \mu(S_2)$. Then

$$d = \Big(\sum_{U \in \operatorname{supp}(\Phi_1)} \Phi_1(U) \cdot f(U) \Big) +_p \Big(\sum_{U \in \operatorname{supp}(\Phi_2)} \Phi_2(U) \cdot g(U) \Big)$$

with $\Phi_1 \in S_1, \Phi_2 \in S_2$, with $f : \operatorname{supp}(\Phi_1) \to \mathcal{D}(X)$ such that $f(U) \in U$, and with

$g : \mathrm{supp}(\Phi_2) \to \mathcal{D}(X)$ such that $g(U) \in U$. For all $x \in X$, we have

$$\begin{aligned}
d(x) &= \Big(\big(\sum_{U \in \mathrm{supp}(\Phi_1)} \Phi_1(U) \cdot f(U) \big) +_p \big(\sum_{U \in \mathrm{supp}(\Phi_2)} \Phi_2(U) \cdot g(U) \big) \Big)(x) \\
&= \big(\sum_{U \in \mathrm{supp}(\Phi_1)} (p \cdot \Phi_1(U) \cdot f(U)(x)) \big) + \big(\sum_{U \in \mathrm{supp}(\Phi_2)} ((1-p) \cdot \Phi_2(U) \cdot g(U)(x)) \big) \\
&= \big(\sum_{U \in \mathrm{supp}(\Phi_1) \setminus \mathrm{supp}(\Phi_2)} (p \cdot \Phi_1(U) \cdot f(U)(x)) \big) \\
&\quad + \big(\sum_{U \in \mathrm{supp}(\Phi_2) \setminus \mathrm{supp}(\Phi_1)} ((1-p) \cdot \Phi_2(U) \cdot g(U)(x)) \big) \\
&\quad + \Big(\sum_{U \in \mathrm{supp}(\Phi_1) \cap \mathrm{supp}(\Phi_2)} \big((p \cdot \Phi_1(U) \cdot f(U)(x)) + ((1-p) \cdot \Phi_2(U) \cdot g(U)(x)) \big) \Big) \\
&\overset{(*)}{=} \big(\sum_{U \in \mathrm{supp}(\Phi_1) \setminus \mathrm{supp}(\Phi_2)} ((\Phi_1 +_p \Phi_2)(U) \cdot f(U)(x)) \big) \\
&\quad + \big(\sum_{U \in \mathrm{supp}(\Phi_2) \setminus \mathrm{supp}(\Phi_1)} ((\Phi_1 +_p \Phi_2)(U) \cdot g(U)(x)) \big) \\
&\quad + \Big(\sum_{U \in \mathrm{supp}(\Phi_1) \cap \mathrm{supp}(\Phi_2)} \big((\Phi_1 +_p \Phi_2)(U) \cdot (f(U)(x) +_{\frac{p \cdot \Phi_1(U)}{(\Phi_1 +_p \Phi_2)(U)}} g(U)(x)) \big) \Big) \\
&= \sum_{U \in \mathrm{supp}(\Phi_1 +_p \Phi_2)} \big((\Phi_1 +_p \Phi_2)(U) \cdot h(U)(x) \big)
\end{aligned}$$

where $h : \mathrm{supp}(\Phi_1 +_p \Phi_2) \to \mathcal{D}(X)$ is defined as:

$$h(U) = \begin{cases} f(U) & \text{if } U \in (\mathrm{supp}(\Phi_1) \setminus \mathrm{supp}(\Phi_2)) \\ g(U) & \text{if } U \in (\mathrm{supp}(\Phi_2) \setminus \mathrm{supp}(\Phi_1)) \\ (f(U) +_{\frac{p \cdot \Phi_1(U)}{(\Phi_1 +_p \Phi_2)(U)}} g(U)) & \text{if } U \in (\mathrm{supp}(\Phi_1) \cap \mathrm{supp}(\Phi_2)) \end{cases}$$

and the equality $(*)$ holds by $(p_1 \cdot q_1) + (p_2 \cdot q_2) = (p_1 + p_2) \cdot (q_1 +_{\frac{p_1}{p_1+p_2}} q_2)$, $\forall p_1, p_2, q_1, q_2$. Then, observe that for every $U \in \mathrm{supp}(\Phi_1 +_p \Phi_2)$ we have $h(U) \in U$, since every U is a convex set, and thus if U contains $f(U)$ and $g(U)$ then it also contains $f(U) +_q g(U)$, for all q. Thereby, we conclude $d \in \mu(S_1 +_p S_2)$.

We now prove the remaining inclusion, i.e., $\mu(S_1 +_p S_2) \subseteq \mu(S_1) +_p \mu(S_2)$.

Let $\Phi \in S_1 +_p S_2$ and let $d = \sum_{U \in \mathrm{supp}(\Phi)} \Phi(U) \cdot f(U)$, with $f : \mathrm{supp}(\Phi) \to \mathcal{D}(X)$ such that $f(U) \in U$, be an element of $\mu(S_1 +_p S_2)$. Then, $\Phi = \Phi_1 +_p \Phi_2$, with $\Phi_1 \in S_1, \Phi_2 \in S_2$. For every $x \in X$ we have

$$\begin{aligned}
d(x) &= \sum_{U \in \mathrm{supp}(\Phi_1) \cup \mathrm{supp}(\Phi_2)} \Big((\Phi_1 +_p \Phi_2)(U) \cdot f(U)(x) \Big) \\
&= \sum_{U \in \mathrm{supp}(\Phi_1) \cup \mathrm{supp}(\Phi_2)} \Big((p \cdot \Phi_1(U) \cdot f(U)(x)) + ((1-p) \cdot \Phi_2(U) \cdot f(U)(x)) \Big) \\
&= \Big(\sum_{U \in \mathrm{supp}(\Phi_1)} p \cdot \Phi_1(U) \cdot f(U)(x) \Big) + \Big(\sum_{U \in \mathrm{supp}(\Phi_2)} (1-p) \cdot \Phi_2(U) \cdot f(U)(x) \Big) \\
&= \Big(\sum_{U \in \mathrm{supp}(\Phi_1)} \Phi_1(U) \cdot f(U)(x) \Big) +_p \Big(\sum_{U \in \mathrm{supp}(\Phi_2)} \Phi_2(U) \cdot f(U)(x) \Big)
\end{aligned}$$

which implies $d \in \mu(S_1) +_p \mu(S_2)$. ◀

By applying the recipe from Section 3.1, we obtain from Lemmas 6 and 7 a monad map.

▶ **Proposition 8.** *The natural transformation* $\iota\colon T_{\Sigma,E} \Rightarrow C$ *is a monad map, defined as:*

$$\iota([x]_E) = \{\delta_x\} \quad \iota([t_1 \oplus t_2]_E) = \iota([t_1]_E) \oplus \iota([t_2]_E) \quad \iota([t_1 +_p t_2]_E) = \iota([t_1]_E) +_p \iota([t_2]_E)$$

Lemma 7, together with the existence of unique bases, also allows us to derive a useful characterization of the multiplication μ of the monad C.

▶ **Lemma 9.** *For* $S \in CCX$,

$$\mu(S) = \operatorname{conv}\Big(\bigcup_{\Phi \in \mathrm{UB}(S)} \Big\{ \sum_{U \in \operatorname{supp}\Phi} \Phi(U) \cdot d \mid d \in \mathrm{UB}(U) \Big\} \Big).$$

Proof. We have $S = \operatorname{conv}(\bigcup_{\Phi \in \mathrm{UB}(S)} \{\Phi\})$ which means that S is a convex union of the sets $\{\Phi\}$, for $\Phi \in \mathrm{UB}(S)$. Then by Lemma 7 we derive $\mu(S) = \operatorname{conv}(\bigcup_{\Phi \in \mathrm{UB}(S)} \mu\{\Phi\})$. By definition, $\mu\{\Phi\} = \{\sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot d \mid d \in U\}$, hence

$$\mu(S) = \operatorname{conv}\Big(\bigcup_{\Phi \in \mathrm{UB}(S)} \Big\{ \sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot d \mid d \in U \Big\} \Big). \tag{12}$$

Observe that the Minkowski sum operation, which is equivalently defined on arbitrary (i.e., not necessarily convex) sets of distributions, enjoys the following property:

$$\text{for any sets of distributions } X, Y,\ \operatorname{conv}(X) +_p \operatorname{conv}(Y) = \operatorname{conv}(X +_p Y). \tag{13}$$

Indeed, $X +_p Y \subseteq \operatorname{conv}(X) +_p \operatorname{conv}(Y)$, and as the Minkowski sum of convex sets is convex we have $\operatorname{conv}(X +_p Y) \subseteq \operatorname{conv}(\operatorname{conv}(X) +_p \operatorname{conv}(Y)) = \operatorname{conv}(X) +_p \operatorname{conv}(Y)$. For the other direction, take $p(\sum_i p_i x_i) + (1-p)(\sum_j q_j y_j) \in \operatorname{conv}(X) +_p \operatorname{conv}(Y)$. We have:

$$p(\sum_i p_i x_i) + (1-p)(\sum_j q_j y_j) = p(\sum_{i,j} (p_i q_j) x_i) + (1-p)(\sum_{i,j} (p_i q_j) y_j) = \sum_{i,j} (p_i q_j)(p x_i + (1-p) y_j)$$

which is then an element of $\operatorname{conv}(X +_p Y)$. This shows (13).

For every Φ, the set $\{\sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot d \mid d \in U\}$ is a Minkowski sum over the elements U of $\operatorname{supp}(\Phi)$, which are themselves convex sets satisfying $U = \operatorname{conv}(\mathrm{UB}(U))$. Then by (13) we derive:

$$\Big\{ \sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot d \mid d \in U \Big\} = \operatorname{conv}\Big(\Big\{ \sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot d \mid d \in \mathrm{UB}(U) \Big\} \Big). \tag{14}$$

By (12) and (14) it holds:

$$\mu(S) = \operatorname{conv}\left(\bigcup_{\Phi \in \mathrm{UB}(S)} \operatorname{conv}\Big(\Big\{ \sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot d \mid d \in \mathrm{UB}(U) \Big\} \Big) \right).$$

Then, by using the property that $\operatorname{conv}(\operatorname{conv}(X) \cup Y) = \operatorname{conv}(X \cup Y)$ for any sets of distributions X, Y (as shown in the proof of [3, Lemma 38]), we conclude that the latter is equal to

$$\operatorname{conv}\Big(\bigcup_{\Phi \in \mathrm{UB}(S)} \Big\{ \sum_{U \in \operatorname{supp}(\Phi)} \Phi(U) \cdot d \mid d \in \mathrm{UB}(U) \Big\} \Big).$$

◀

6 Proving the isomorphism

In the previous section we have constructed a monad map $\iota\colon T_{\Sigma,E} \Rightarrow C$ (Proposition 8). In this section, we prove that it is an isomorphism by exploting Theorem 1.

We start with a simple observation: for each set X, there is a trivial injection $i_X\colon T_{\Sigma_P}(X) \to T_\Sigma(X)$. A term in T_Σ is said to be a *purely probabilistic term* (p-term, for short) if and only if it lays in the image of i. We overload the notation and also denote with i its extension to equivalence classes $i_X\colon T_{\Sigma_P,E_P}(X) \to T_{\Sigma,E}(X)$, which is well defined as $E_P \subseteq E$.

▶ **Lemma 10.** *Let $\{-\}_X\colon \mathcal{D}(X) \to C(X)$ be the function mapping every distribution d into the convex set $\{d\}$ and let $\iota^P\colon T_{\Sigma_P,E_P} \Rightarrow \mathcal{D}$ be the monad map from* (8)*. The following diagram commutes.*

$$\begin{array}{ccc} T_{\Sigma_P,E_P}X & \xrightarrow{i_X} & T_{\Sigma,E}X \\ {\scriptstyle \iota^P_X}\downarrow & & \downarrow{\scriptstyle \iota_X} \\ \mathcal{D}X & \xrightarrow[\{-\}_X]{} & CX \end{array}$$

Proof. We prove by induction that $\{\iota^P_X([t]_{E_P})\}_X = \iota_X(i_X([t]_{E_P}))$ for all $t \in T_{\Sigma_P}$. If $t = x \in X$, then $\{\iota^P_X([x]_{E_P})\}_X = \{\delta_x\} = \iota_X([x]_E) = \iota_X(i_X([t]_{E_P}))$. If $t = t_1 +_p t_2$, then

$$\begin{aligned} \{\iota^P_X([t_1 +_p t_2]_{E_P})\}_X &= \{p \cdot \iota^P_X([t_1]_{E_P}) + (1-p) \cdot \iota^P_X([t_1]_{E_P})\} \\ &= \{\iota^P_X([t_1]_{E_P})\} +_p \{\iota^P_X([t_2]_{E_P})\} \\ &= \iota_X(i_X([t_1]_{E_P})) +_p \iota_X(i_X([t_2]_{E_P})) \\ &= \iota_X([t_1]_E) +_p \iota_X([t_2]_E) \\ &= \iota_X([t_1 +_p t_2]_E) \\ &= \iota_X(i_X([t_1 +_p t_2]_{E_P})). \end{aligned}$$

◀

Recall that the monad map $\iota^P\colon T_{\Sigma_P,E_P} \Rightarrow \mathcal{D}$ defined in (8) is an isomorphism. We call $\kappa^P\colon \mathcal{D} \Rightarrow T_{\Sigma_P,E_P}$ its inverse. By exploiting κ^P and Theorem 1, it is easy to define a function $\kappa_X\colon C(X) \to T_{\Sigma,E}(X)$ as follows: for $S \in C(X)$ with base $\{d_1,\dots,d_n\}$

$$\kappa_X(S) = [\, i(\kappa^P(d_1)) \oplus \dots \oplus i(\kappa^P(d_n))\,]_E. \tag{15}$$

▶ **Proposition 11.** $\iota \circ \kappa = id_C$

Proof. Let $S \in C(X)$ be a convex set with base $\{d_1,\dots,d_n\}$. By definition of κ and ι,

$$\iota(\kappa(S)) = \iota([i(\kappa^P(d_1))]_E) \oplus \dots \oplus \iota([i(\kappa^P(d_n))]_E).$$

By Lemma 10, $\iota(\kappa(S)) = \{d_1\} \oplus \dots \oplus \{d_n\}$ which is exactly S. ◀

▶ Remark 12. Proposition 11 and Lemma 10 entail that $i_X\colon T_{\Sigma_P,E_P}(X) \to T_{\Sigma,E}(X)$ is injective. Hence, two p-terms are equal in E if and only if they are also equal in E_P.

We are now left to prove that $\kappa \circ \iota = id_{T_{\Sigma,E}}$. This means that any term t is in the equivalence class of $\kappa \circ \iota([t]_E)$, which by definition of κ is $[i(\kappa^P(d_1)) \oplus \dots \oplus i(\kappa^P(d_n))]_E$ where $\{d_1,\dots,d_n\}$ is the base of $\iota([t]_E)$.

The first step consists in showing that every term is equivalent, modulo E, with a term of a certain shape: a term $t \in T_\Sigma(X)$ is said to be in *nondeterministic-probablistic form*, *n-p*

form for short, if there exists $t_1, \dots, t_n \in T_{\Sigma_P}(X)$ such that $t = i(t_1) \oplus \dots \oplus i(t_n)$. This can be thought of as an analogous of the disjunctive-conjunctive form that is commonly used in propositional logic.

▶ **Example 13.** The term $(x \oplus y) +_{\frac{1}{2}} (y +_{\frac{1}{3}} z)$ is not in n-p form, since $x \oplus y$ occurs inside $+_{\frac{1}{2}}$. However, by using the distributivity axiom (D), we have that $(x \oplus y) +_{\frac{1}{2}} (y +_{\frac{1}{3}} z) =_E (x +_{\frac{1}{2}} (y +_{\frac{1}{3}} z)) \oplus (y +_{\frac{1}{2}} (y +_{\frac{1}{3}} z))$ which is in n-p form.

The following proposition ensures that every term is equivalent to one in n-p form.

▶ **Proposition 14.** *For all $t \in T_\Sigma(X)$, there exists t' in n-p form such that $t =_E t'$.*

Proof. Intuitively, by virtue of the axiom (D) all the occurrences of $+_p$ can be pushed inside some $\oplus$. This can be proved formally by means of the following term rewriting system.

$$(t_1 \oplus t_2) +_p t_3 \rightsquigarrow (t_1 +_p t_3) \oplus (t_2 +_p t_3) \qquad t_1 +_p (t_2 \oplus t_3) \rightsquigarrow (t_1 +_p t_2) \oplus (t_1 +_p t_3)$$

If $t \in T_\Sigma(X)$ rewrites to $t' \in T_\Sigma(X)$, then $t =_E t'$ since the left rule is just the axiom (D), while the right can be derived using (C_p), (D) and (C_p) again. Using standard term rewriting techniques from [6] we can prove that the rewriting system terminates:

(1) Define the partial order $+_p > \oplus$ on Σ;
(2) Observe that the generated recursive path ordering on $T_\Sigma(X)$ is a simplification ordering (see e.g., Example A in Section 5 of [6]);
(3) Conclude by the First Termination Theorem.

Finally, we observe that a term t is in n-p form iff $t \not\rightsquigarrow$: Indeed, if t is in n-p form then there is no redex for the two rules above. On the other hand, if t is not in n-p form, then some $\oplus$ should occur inside a $+_p$ and then one of the rules applies.

Therefore, each term t can be rewritten into an E-equivalent term t' in n-p form. ◀

Given a term $t' \in T_\Sigma(X)$ in n-p form and $t_1, \dots, t_n \in T_{\Sigma_P}(X)$ such that $t' = i(t_1) \oplus \dots \oplus i(t_n)$, one would like $\{\iota^P([t_1]_{E_P}), \dots, \iota^P([t_n]_{E_P})\}$ to be the base for $\iota([t']_E)$. But this is not always the case since some $\iota^P([t_i]_{E_P})$ can be a convex combination of the other $\iota^P([t_j]_{E_P})$.

▶ **Example 15.** The term $(x +_{\frac{1}{2}} y) \oplus (x +_{\frac{2}{3}} (x \oplus y))$ is not in n-p form. By applying the rewriting procedure in the proof of Proposition 14 one obtains: $(x +_{\frac{1}{2}} y) \oplus (x +_{\frac{2}{3}} (x \oplus y)) =_E (x +_{\frac{1}{2}} y) \oplus (x +_{\frac{2}{3}} x) \oplus (x +_{\frac{2}{3}} y)$. Observe that this is equivalent to $(x +_{\frac{1}{2}} y) \oplus x \oplus (x +_{\frac{2}{3}} y)$. The convex set $\iota\big([(x +_{\frac{1}{2}} y) \oplus x \oplus (x +_{\frac{2}{3}} y)]_E\big)$ has base $\{\iota^P([x +_{\frac{1}{2}} y]_{E_P}), \iota_P([x]_{E_P})\} = \{\frac{1}{2}x + \frac{1}{2}y, \delta_x\}$. Indeed the distribution $\iota^P([x +_{\frac{2}{3}} y]_{E_P}) = \frac{2}{3}x + \frac{1}{3}y$ is a convex combination of $\{\frac{1}{2}x + \frac{1}{2}y, \delta_x\}$ as $\frac{2}{3}x + \frac{1}{3}y = \frac{2}{3}(\frac{1}{2}x + \frac{1}{2}y) + \frac{1}{3}x$.

The next two lemmas are necessary to show that, using the axioms in E, we can remove from t' those summands $i(t_i)$ such that $\iota^P([t_i]_{E_P})$ is a convex combination of the other $\iota^P([t_j]_{E_P})$. The first lemma is a well known observation (see e.g. [23, 33]), but we report its instructive proof; the second lemma follows easily from the first one and properties of convex algebras. We defer its proof to the Appendix.

▶ **Lemma 16** (Convexity law). *For all terms $t_1, t_2 \in T_\Sigma(X)$, for all $p \in (0, 1)$,*

$$t_1 \oplus t_2 =_E t_1 \oplus t_2 \oplus (t_1 +_p t_2).$$

Proof. We first prove that

$$\begin{aligned} t_1 \oplus t_2 &\overset{(I_p)}{=} (t_1 \oplus t_2) +_p (t_1 \oplus t_2) \\ &\overset{(D)}{=} ((t_1 \oplus t_2) +_p t_1) \oplus ((t_1 \oplus t_2) +_p t_2) \\ &\overset{(D)}{=} ((t_1 +_p t_1) \oplus (t_2 +_p t_1)) \oplus ((t_1 +_p t_2) \oplus (t_2 +_p t_2)) \\ &\overset{(I_p)}{=} t_1 \oplus (t_2 +_p t_1) \oplus (t_1 +_p t_2) \oplus t_2 \end{aligned}$$

Then, by applying first this equality and then idempotency, we derive the result:

$$\begin{aligned} t_1 \oplus t_2 \oplus (t_1 +_p t_2) &= t_1 \oplus (t_2 +_p t_1) \oplus (t_1 +_p t_2) \oplus t_2 \oplus (t_1 +_p t_2) \\ &\overset{(I_p)}{=} t_1 \oplus (t_2 +_p t_1) \oplus (t_1 +_p t_2) \oplus t_2 \end{aligned}$$

◀

▶ **Lemma 17.** *Let $t, t_1, \ldots, t_n \in T_{\Sigma_P}(X)$ such that $\iota^P([t]_{E_P}) \in \mathrm{conv}\{\iota^P([t_1]_{E_P}), \ldots, \iota^P([t_n]_{E_P})\}$. Then*

$$i(t_1) \oplus \ldots \oplus i(t_n) =_E i(t_1) \oplus \ldots \oplus i(t_n) \oplus i(t)$$

▶ **Proposition 18.** *For all terms $t \in T_\Sigma(X)$, there exist $t_1, \ldots, t_n \in T_{\Sigma_P}$ such that*

$$t =_E i(t_1) \oplus \ldots \oplus i(t_n)$$

and $\{\iota^P([t_1]_{E_P}), \ldots, \iota^P([t_n]_{E_P})\}$ is the base of $\iota([t]_E)$.

Proof. By Proposition 14, there exists a $t' \in T_\Sigma(X)$ in n-p form such that $t =_E t'$. Take $t'_1, \ldots, t'_m \in T_{\Sigma_P}$ such that $t' = i(t_1) \oplus \ldots \oplus i(t_m)$. By definition of ι, $\iota([t]_E) = \iota(i([t_1]_{E_P})) \oplus \cdots \oplus \iota(i([t_m]_{E_P}))$ which by Lemma 10 is $\{\iota^P([t_1]_{E_P})\} \oplus \cdots \oplus \{\iota^P([t_m]_{E_P})\}$. By definition of $\oplus$, this is just $\mathrm{conv}\{\iota^P([t_1]_{E_P}), \ldots, \iota^P([t_m]_{E_P})\}$. Therefore, to conclude that $\{\iota^P([t_1]_{E_P}), \ldots, \iota^P([t_m]_{E_P})\}$ is the base of $\iota([t]_E)$ we only need to show that none of the $\iota^P([t_i]_{E_P})$ is in the convex combination of the others $\iota^P([t_j]_{E_P})$. This is not true in general, but thanks to Lemma 17 all such t_i can be removed, while preserving E-equivalence. To be more precise, by associativity and commutativity of $\oplus$, we can assume that $\iota^P([t_1]_{E_P}), \ldots, \iota^P([t_n]_{E_P})$ form the base, while $\iota^P([t_{n+1}]_P), \ldots, \iota^P([t_m]_{E_P})$ are in $\mathrm{conv}\{\iota^P([t_1]_{E_P}), \ldots, \iota^P([t'_n]_{E_P})\}$. Then, by repeating $(m-n)$-times Lemma 17, we conclude that $t' =_E i(t_1) \oplus \ldots \oplus i(t_n)$. ◀

▶ **Proposition 19.** $\kappa \circ \iota = id_{T_{\Sigma,E}}$

Proof. We need to prove that for all terms $t \in T_\Sigma(X)$, $[t]_E = \kappa \circ \iota([t]_E)$. By Proposition 18, there exists $t_1, \ldots, t_n \in T_{\Sigma_P}(X)$ such that

$$t =_E i(t_1) \oplus \ldots \oplus i(t_n)$$

and $\{\iota^P([t_1]_{E_P}), \ldots, \iota^P([t_n]_{E_P})\}$ is the base for $\iota([t]_E)$.
By definition of κ, $\kappa(\iota([t]_E))$ is exactly $[i(\kappa^P \circ \iota^P[t_1]_{E_P}) \oplus \ldots \oplus i(\kappa^P \circ \iota^P[t_n]_{E_P})]_E = [t]_E$. ◀

This is enough to conclude the proof of Theorem 3. Indeed we have that $\iota \colon T_{\Sigma,E} \Rightarrow C$ is a monad map and that, by Propositions 11 and 19, it is an isomorphism.

References

1 Christel Baier and Joost-Pieter Katoen. *Principles of model checking*. MIT Press, 2008.

2 Filippo Bonchi, Alexandra Silva, and Ana Sokolova. The Power of Convex Algebras. In *CONCUR 2017*, volume 85, pages 23:1–23:18. LIPIcs, 2017. doi:10.4230/LIPIcs.CONCUR.2017.23.

3 Filippo Bonchi, Ana Sokolova, and Valeria Vignudelli. The theory of traces for systems with nondeterminism and probability. Extended version of paper in Proc.LICS'19, 2019. URL: http://arxiv.org/abs/1808.00923v3.

4 Pablo Samuel Castro, Prakash Panangaden, and Doina Precup. Equivalence relations in fully and partially observable markov decision processes. In *IJCAI*, pages 1653–1658, 2009.

5 Christian Dehnert, Sebastian Junges, Joost-Pieter Katoen, and Matthias Volk. A storm is coming: A modern probabilistic model checker. In *Proc. CAV 2017*, volume 10427 of *LNCS*, pages 592–600, 2017.

6 Nachum Dershowitz. Orderings for term-rewriting systems. *Theoretical computer science*, 17(3):279–301, 1982.

7 Ernst-Erich Doberkat. Eilenberg-Moore algebras for stochastic relations. *Inform. and Comput.*, 204(12):1756–1781, 2006. URL: http://dx.doi.org/10.1016/j.ic.2006.09.001, doi:10.1016/j.ic.2006.09.001.

8 Ernst-Erich Doberkat. Erratum and addendum: Eilenberg-Moore algebras for stochastic relations [mr2277336]. *Inform. and Comput.*, 206(12):1476–1484, 2008. URL: http://dx.doi.org/10.1016/j.ic.2008.08.002, doi:10.1016/j.ic.2008.08.002.

9 Jean Goubault-Larrecq. Prevision domains and convex powercones. In *FOSSACS 2008*, pages 318–333. LNCS 4962, 2008. doi:10.1007/978-3-540-78499-9_23.

10 Alexandre Goy and Daniela Petrisan. Combining probabilistic and non-deterministic choice via weak distributive laws. In Holger Hermanns, Lijun Zhang, Naoki Kobayashi, and Dale Miller, editors, *LICS '20: 35th Annual ACM/IEEE Symposium on Logic in Computer Science, Saarbrücken, Germany, July 8-11, 2020*, pages 454–464. ACM, 2020. doi:10.1145/3373718.3394795.

11 Hans A Hansson. Time and probability in formal design of distributed systems. *PhD thesis, Uppsala University*, 1991.

12 Holger Hermanns, Jan Krcál, and Jan Kretínský. Probabilistic bisimulation: Naturally on distributions. In *Proc. CONCUR'14*, volume 8704 of *LNCS*, pages 249–265, 2014.

13 Holger Hermanns, Augusto Parma, Roberto Segala, Björn Wachter, and Lijun Zhang. Probabilistic logical characterization. *Information and Computation*, 209(2):154–172, 2011.

14 Chris Heunen, Ohad Kammar, Sam Staton, and Hongseok Yang. A convenient category for higher-order probability theory. *CoRR*, abs/1701.02547, 2017. URL: http://arxiv.org/abs/1701.02547.

15 B. Jacobs. Convexity, duality and effects. In *Theoretical computer science*, volume 323 of *IFIP Adv. Inf. Commun. Technol.*, pages 1–19. Springer, Berlin, 2010. URL: http://dx.doi.org/10.1007/978-3-642-15240-5_1, doi:10.1007/978-3-642-15240-5_1.

16 Bart Jacobs. Coalgebraic trace semantics for combined possibilitistic and probabilistic systems. *Electr. Notes Theor. Comput. Sci.*, 203(5):131–152, 2008.

17 Leslie Pack Kaelbling, Michael L Littman, and Anthony R Cassandra. Planning and Acting in Partially Observable Stochastic Domains. *Artif. Intell.*, 1998.

18 Klaus Keimel and Gordon D. Plotkin. Mixed powerdomains for probability and nondeterminism. *Logical Methods in Computer Science*, 13(1), 2017. doi:10.23638/LMCS-13(1:2)2017.

19 Marta Z. Kwiatkowska, Gethin Norman, and David Parker. Prism: Probabilistic symbolic model checker. In *Computer Performance Evaluation / TOOLS*, pages 200–204. LNCS 2324, 2002.

20 Matteo Mio. Upper-expectation bisimilarity and łukasiewicz μ-calculus. In *Proc. FOSSACS'14*, volume 8412 of *LNCS*, pages 335–350, 2014.

21 Matteo Mio, Ralph Sarkis, and Valeria Vignudelli. Combining nondeterminism, probability, and termination: Equational and metric reasoning. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021, Rome, Italy, June 29 - July 2, 2021*, pages 1–14. IEEE, 2021. doi:10.1109/LICS52264.2021.9470717.

22 Matteo Mio and Valeria Vignudelli. Monads and quantitative equational theories for nondeterminism and probability. In Igor Konnov and Laura Kovács, editors, *31st International Conference on Concurrency Theory, CONCUR 2020, September 1-4, 2020, Vienna, Austria (Virtual Conference)*, volume 171 of *LIPIcs*, pages 28:1–28:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.CONCUR.2020.28.

23 M. Mislove, J. Ouaknine, and J. Worrell. Axioms for probability and nondeterminism. In *Proc. of the 10th Int. Workshop on Expressiveness in Concurrency (EXPRESS 2003)*, volume 96 of *ENTCS*, pages 7–28. Elsevier, 2003.

24 Michael W. Mislove. Nondeterminism and probabilistic choice: Obeying the laws. In *CONCUR 2000*, pages 350–364. LNCS 1877, 2000. doi:10.1007/3-540-44618-4_26.

25 Walter Rudin. *Functional Analysis*. McGraw-Hill, 1991.

26 Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 2009.

27 Roberto Segala and Nancy Lynch. Probabilistic simulations for probabilistic processes. *Nordic Journal of Computing*, 2(2):250–273, 1995.

28 Zbigniew Semadeni. *Monads and their Eilenberg-Moore algebras in functional analysis*. Queen's University, Kingston, Ont., 1973. Queen's Papers in Pure and Applied Mathematics, No. 33.

29 Sam Staton, Hongseok Yang, Frank Wood, Chris Heunen, and Ohad Kammar. Semantics for probabilistic programming: higher-order functions, continuous distributions, and soft constraints. In *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016*, pages 525–534, 2016. URL: http://doi.acm.org/10.1145/2933575.2935313, doi:10.1145/2933575.2935313.

30 T. Świrszcz. Monadic functors and convexity. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 22:39–42, 1974.

31 Regina Tix, Klaus Keimel, and Gordon D. Plotkin. Semantic domains for combining probability and non-determinism. *ENTCS*, 222:3–99, 2009. doi:10.1016/j.entcs.2009.01.002.

32 R. Tyllerr. *Convex Analysis*. Princeton University Press, 1972.

33 D. Varacca. *Probability, Nondeterminism and Concurrency: Two Denotational Models for Probabilistic Computation*. PhD thesis, Univ. Aarhus, 2003. BRICS Dissertation Series, DS-03-14.

34 D. Varacca and G. Winskel. Distributing probabililty over nondeterminism. *MSCS*, 16(1):87–113, 2006.

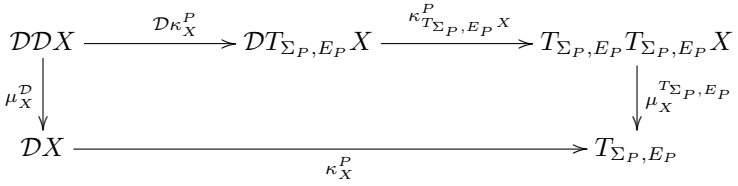
35 Moshe Y Vardi. Automatic verification of probabilistic concurrent finite state programs. In *Foundations of Computer Science, 1985., 26th Annual Symposium on*, pages 327–338. IEEE, 1985.

A Proof of Lemma 17.

▶ **Lemma 20.** *Let* $t, t_1, \ldots, t_n \in T_{\Sigma_P}(X)$ *such that* $\iota^P([t]_{E_P}) \in \mathrm{conv}\{\iota^P([t_1]_{E_P}), \ldots, \iota^P([t_n]_{E_P})\}$. *Then there exist* $p_1 \ldots p_{n-1} \in (0,1)$ *such that* $t =_{E_P} (\ldots(t_1 +_{p_1} t_2) +_{p_2} \ldots) +_{p_{n-1}} t_n$.

Proof. If $\iota^P([t]_{E_P}) \in \mathrm{conv}\{\iota^P([t_1]_{E_P}), \ldots, \iota([t_n]_{E_P})\}$, then $\iota^P([t]_{E_P}) = \mu^{\mathcal{D}}\big(\sum_i q_i\,(\iota^P([t_i]_{E_P}))\big)$ for some $\sum_i q_i = 1$. Since ι^P is a monad map, its inverse $\kappa^D \colon \mathcal{D} \Rightarrow T_{\Sigma_P, E_P}$ is also a monad

map and in particular it makes the following diagram commute.



Therefore, we have that

$$\begin{aligned}
[t]_{E_P} &= \kappa^P \circ \iota^P([t]_{E_P}) \\
&= \kappa^P \circ \mu^{\mathcal{D}}\Big(\sum_i q_i\,(\iota^P([t_i]_{E_P}))\Big) \\
&= \mu^{T_{\Sigma_P,E_P}} \circ \kappa^P_{T_{\Sigma_P,E_P}} \circ \mathcal{D}\kappa^P\Big(\sum_i q_i\,(\iota^P([t_i]_{E_P}))\Big) \\
&= \mu^{T_{\Sigma_P,E_P}} \circ \kappa^P_{T_{\Sigma_P,E_P}}\Big(\sum_i q_i\,(\kappa^P \circ \iota^P([t_i]_{E_P}))\Big) \\
&= \mu^{T_{\Sigma_P,E_P}} \circ \kappa^P_{T_{\Sigma_P,E_P}}\Big(\sum_i q_i\,[t_i]_{E_P}\Big)
\end{aligned}$$

Observe that $\sum_i q_i\,[t_i]_{E_P} \in \mathcal{D}T_{\Sigma_p,E_P}(X)$ and that $\kappa^P_{T_{\Sigma_p,E_P}X}$ maps it into an element of $T_{\Sigma_p,E_P}T_{\Sigma_p,E_P}(X)$, namely a term obtained by the operations $+_p$ and the constants $[t_i]_{E_P}$. Using the axioms in E_P any such term can always be written as $(\dots([t_1]_{E_P} +_{p_1} [t_2]_{E_P}) +_{p_2} \dots) +_{p_{n-1}} [t_n]_{E_P}$ for some $p_i \in (0,1)$. Then, the application of $\mu^{T_{\Sigma_P,E_P}}$ to $[(\dots([t_1]_{E_P} +_{p_1} [t_2]_{E_P}) +_{p_2} \dots) +_{p_{n-1}} [t_n]_{E_P}]_{E_P}$ gives $[(\dots(t_1 +_{p_1} t_2) +_{p_2} \dots) +_{p_{n-1}} t_n]_{E_P}$. Thus $t =_{E_P} (\dots(t_1 +_{p_1} t_2) +_{p_2} \dots) +_{p_{n-1}} t_n$. ◀

Proof of Lemma 17. By Lemma 20, we take $p_1, \dots, p_{n-1}$ such that

$$t =_{E_P} (\dots(t_1 +_{p_1} t_2) +_{p_2} \dots) +_{p_{n-1}} t_n. \tag{16}$$

By Lemma 16, $i(t_1) \oplus \dots \oplus i(t_n)$ is E-equivalent to $i(t_1) \oplus \dots \oplus i(t_n) \oplus i(t_1 +_{p_1} t_2)$. By applying Lemma 16 again, one obtains $i(t_1) \oplus \dots \oplus i(t_n) \oplus i(t_1 +_{p_1} t_2) \oplus i((t_1 +_{p_1} t_2) +_{p_2} t_3)$. We can then remove $i(t_1 +_{p_1} t_2)$ using Lemma 16, to obtain

$$i(t_1) \oplus \dots \oplus i(t_n) \oplus i((t_1 +_{p_1} t_2) +_{p_2} t_3).$$

By iterating this procedure, one obtains

$$i(t_1) \oplus \dots \oplus i(t_n) \oplus i((\dots(t_1 +_{p_1} t_2) +_{p_2} \dots) +_{p_{n-1}} t_n)$$

which, by (16), is $i(t_1) \oplus \dots \oplus i(t_n) \oplus i(t)$. ◀