Universität Dortmund
Fachbereich Informatik
Lehrstuhl für Software-Technologie

MEMO Nr. 149

Derandomizing probabilistic semantics through Eilenberg-Moore algebras for the Giry monad

Ernst-Erich Doberkat

July 2004

Internes Memorandum des
Lehrstuhls für Software-Technologie
Prof. Dr. Ernst-Erich Doberkat
Fachbereich Informatik
Universität Dortmund
Baroper Straße 301

D-44227 Dortmund

ISSN 0933-7725

Derandomizing probabilistic semantics through Eilenberg-Moore algebras for the Giry monad

Ernst-Erich Doberkat[1]
Chair for Software Technology
University of Dortmund
doberkat@acm.org

July 29, 2004

[1]Research funded in part by Deutsche Forschungsgemeinschaft, grant DO 263/8-1, *Algebraische Eigenschaften stochastischer Relationen*

Abstract

A simple language is defined, a probabilistic semantics and a partial correctness logic is proposed in terms of probabilistic relations. It is shown how a derandomized semantics can be constructed through the Eilenberg-Moore algebras. We investigate the category of Eilenberg-Moore algebras for the Giry monad associated with stochastic relations over Polish spaces with continuous maps as morphisms. The algebras are characterized through strongly convex structures on the base space.

Keywords: Stochastic relations, Giry monad, Eilenberg-Moore algebras, computation through monads, positive convexity, derandomizing semantics.

1 Introduction

Modeling a computation through a monad (as suggested e.g. by E. Moggi [18]), one represents state transitions or the transformation from inputs a to outputs b through a morphism $a \to \mathbf{T}b$ with $\mathbf{T}$ as the functor underlying the monad. Working in a probabilistic setting, a state from a base space X is in this way associated with a subprobability distribution $K(x)$ on X. Here $K : X \to \mathbf{S}(X)$ is a morphism for the probability monad, in which the functor assigns a space its probabilities. But we now have only a distribution of the outputs, not the outputs proper. What is needed for this is a map $h : \mathbf{S}(X) \to X$ that would transform a distribution into a state proper. This problem arises when studying the semantics of programming languages in terms of stochastic relations. States are modeled through these relations, and the program is modeled as a state transformer, thus mapping stochastic relations to stochastic relations again. We work with deterministic programs, and we would like to see how state transformations translate into effects on the underlying base space. For this, we introduce derandomization, mapping subprobabilities on the base space, hereby respecting the probabilistic structure. This means that point measures are mapped to their respective points, and that positive convex combinations of measures are mapped to the convex combinations of their images. But, alas, this requires a positive convex structure on X. We will show that those maps h that comply with the rest of the monad carry their own positive convex structures with them. Such pairs $\langle X, h\rangle$ are called an Eilenberg-Moore algebra (or simply an algebra) for this monad. Structurally, these algebras help to construct an adjunction for which the monad is just the given one [17, Theorem VI.2.1]. In fact, this adjunction and the one constructed through the Kleisli category form in some sense the extreme points in a category of all adjunctions from which the given monad can be recovered [17, Theorem VI.5.3]. Thus it is of algebraic interest to identify these algebras in general, and in particular to the probability functor. It has as a Kleisli construction stochastic relations and is in this sense quite similar to the powerset functor. For the latter functor the algebras are completely characterized, but the stochastic side of the analogy is not explored fully yet. This paper makes an attempt at providing characterizations for these algebras under the assumption of continuity. We work in the category of Polish spaces (these spaces are explained in Section 3) with continuous maps as morphisms. In this category the algebras for the Giry monad are identified, and the category of all algebras is investigated. The natural approach is to think of these algebras in terms of an equivalence relation which may be thought to identify probability distributions, and to investigate either these relations or the partitions associated with them. These characterizations lead to the identification of the algebras as the positive convex structures on their base space. A similar result has been known for probability measures on compact Hausdorff spaces [11, 27, 20], but a full characterization of the subprobabilistic case on Polish spaces seems to be new. With this characterization in mind, we show how the semantics of our example language can be derandomized consistently.

Thus the paper contributes to the theory of stochastic relations by providing a characterization of the category of Eilenberg-Moore algebras for the subprobability functor in the category of Polish spaces with continuous mappings as morphisms. It shows furthermore how stochastic relations apply to the semantics of programming languages, and it introduces the notion of derandomization, for which consistency with a partial correctness logic is established.

Organization To illustrate the problem, we have a look first at Ludwig, a simple language that permits expressing the question consistently. We give a partial correctness logic and a corresponding semantics in terms of stochastic relations, and prove consistency and completeness. This happens in Section 2. We define the objects we are dealing with in Section 3, in particular, the space of all subprobability measures on a Polish space is introduced together with the weak topology that renders it a Polish space. The Giry monad is also introduced. Section 5 is devoted to the characterization of the algebras for this monad through partitions, smooth equivalence relations, and positive convex structures (which are given a defining glance in Section 4). The category of algebras is shown to be isomorphic to these categories. Some examples are given there, too, indicating among others that the search for algebras in the – usually easily dealt with — finite case is somewhat hopeless. Section 6 returns to Ludwig and established consistency in terms of derandomizations; we show also that morphisms in the category of algebras permit the transport of derandomized semantics. Section 7 has a brief look at related work, and Section 8 proposes further investigations along the lines developed here.

Acknowledgement Georgios Lajios helped improving the clarity of the representation with critical remarks. Dieter Pumplün's suggestions to investigate convex structures, and his general advice on algebras are highly appreciated. The diagrams were typeset with P. Taylor's wonderful `diagrams` package.

2 A Simple Language

Perceiving a program as a state transforming device, we propose stochastic relations over the Polish base space X as states and give a semantics for a program in terms of state transformations. Since the programs are not expected to terminate always, we deal with relations $g : X \rightsquigarrow X$ that do not add up to 1, so that $1 - g(x)(X)$ is the "probability" for the program not to terminate. When a program works one wants to know how it operates not only in terms of probabilities associated with it but also in terms of its deterministic behavior (its work is deterministic, after all). Thus it is interesting to model the semantics of programs also in terms of derandomization.

Definition 1 *A* derandomization h *on the base space* X *assigns each subprobability* $\tau \in \mathbf{S}(X)$ *an element* $h(\tau) \in X$ *such that*

1. $h(\delta_x) = x$*, thus if the probability concentrates on a point, this is the de-randomization (*δ_x *is the Dirac measure on* x*),*
2. *whenever* $\tau_1, \ldots, \tau_n \in \mathbf{S}(X)$ *and* $c_1, \ldots, c_n \geq 0$ *with* $c_1 + \cdots + c_n \leq 1$*, then*
$$h(c_1 \cdot \tau_1 + \cdots + c_n \cdot \tau_n) = c_1 \cdot h(\tau_1) + \cdots + c_n \cdot h(\tau_n).$$
3. h *is continuous.*

Condition 1 seems to be fairly obvious, condition 2 implies that derandomization is geometrically smooth: if a probability is the positive convex combination of two other ones, its derandomization should be, too. Continuity in condition 3 means that if one subprobability is close to another one, their de-randomization is, too.

Remark 1 *Suppose for the sake of illustration that we model a nondeterministic computation using set theoretic relations. Given a* sup*-complete partial order, a decision* Δ *(borrowing the term from [2]) that extracts an element from a set of possible outcomes can be described through the following conditions:*

1. $\Delta(\{x\}) = x$ *for all* $x \in X$,
2. $\Delta(A) = \sup A$ *for all* $A \subseteq X$,
3. $\Delta(\bigcup \mathcal{D}) = \sup\{\Delta(A) \mid A \in \mathcal{D}\}$ *for all* $\mathcal{D} \subseteq \mathcal{P}(X)$

Condition 1 again is fairly obvious, condition 2 proposes selecting the optimal element w.r.t. the given partial order, and condition 3 indicates that selecting the optimal element from a collection is tantamount to optimizing choices already made for the components of the collection.
It will turn out that this notion is also governed by a monad, see Remark 3. —

There is, however, for the probabilistic case one slight —but not unimportant — catch: the base space X needs not have a positive convex structure at all, so forming positive convex combinations is not necessarily possible (note that we do not restrict ourselves to strict convex combinations, i.e., non-negative coefficients adding up to 1, since we work in the realm of subprobability measures).
Modeling computations through a monad as suggested by E. Moggi [18], we will show that each derandomization carries its own positive convex structure with it, and vice versa: once a map h is given, a corresponding positive convex structure can be set up, and from a positive convex structure a derandomization can be constructed. Condition 2 is replaced by a more amenable one, viz., $h \circ \mathbf{S}(h) = h \circ \mu_X$, where $\langle \mathbf{S}, \mu, \eta \rangle$ is the monad associated with the subprobability functor. In this way the operating theater is shifted to subprobability measures on which a natural positive convex structure exists. Thus a derandomization is an algebra for this functor, and we will need to characterize these algebras in terms of convex structures.
As a case in point, we will first give a simple language, a partial correctness semantics and a partial correctness logic for it (partiality creeping in through taking nontermination into account). Both, semantics and logic, will work on the basis of transforming stochastic relations. We demonstrate consistency and completeness, so that we are on safe grounds when dealing with the language's semantics. Let g_α be the initial state of program P and g_ω with $\mathcal{C}[\![P]\!]g_\alpha = g_\omega$ be its final state, then a derandomization h permits a deterministic interpretation of initial and final states in terms of the base space: viewed deterministically, the initial and the final state are represented through $h \circ g_\alpha$, and $h \circ g_\omega$, resp, making use of the convex structure on X imposed by h.
After defining the language and preparing for defining its semantics, we formulate the partial correctness semantics and logics, and demonstrate consistency and completeness. The reader is requested to take the probabilistic notions in this Section for granted; Section 3 will collect the constructions for the reader's convenience.

2.1 The Language: Ludwig

The language Ludwig[1] has assignments, conditional statements, loops and composition. Its syntax is given through:

$$\beta ::= \texttt{skip} \mid a := \mathbf{E} \mid \texttt{if}\ \alpha\ \texttt{then}\ \beta\ \texttt{else}\ \beta\ \texttt{fi} \mid \texttt{while}\ \alpha\ \texttt{do}\ \beta\ \texttt{od} \mid \beta;\beta$$

(we will abbreviate the conditional command by `if` α `then` β `fi`, if the alternative branch is `skip`). The variables a are taken from a set **Vars**, the expressions **E** from a set **Exprs**. Let X be a Polish space which will be fixed in the sequel. A Boolean α is modeled through a map $\alpha : X \to \{\mathsf{tt}, \mathsf{ff}\}$ from X to the truth values $\{\mathsf{tt}, \mathsf{ff}\}$. It is natural to assume measurability of this map, so that the set

$$\{\alpha = \mathsf{tt}\} := \{x \in X \mid \alpha(x) = \mathsf{tt}\}$$

is a Borel set (hence for $\{\alpha = \mathsf{ff}\}$). Because there are at most countably many of these sets for Ludwig, we can find a Polish topology on X which is finer than the given one having the same Borel sets such that these sets are clopen, see [26, Corollary 3.2.5]. We will assume that X is equipped with this topology. Associate with α the stochastic relations $\chi^{\natural}_{\{\alpha=\mathsf{tt}\}}$ and $\chi^{\natural}_{\{\alpha=\mathsf{ff}\}}$. The former is defined through

$$\chi^{\natural}_{\{\alpha=\mathsf{tt}\}}(x)(A) := \chi_{\{\alpha=\mathsf{tt}\}}(x) \cdot \delta_x(A),$$

where χ_S is the indicator function of set S, similarly for $\chi^{\natural}_{\{\alpha=\mathsf{ff}\}}$. Note that

$$\int_X f \ d\chi^{\natural}_{\{\alpha=\mathsf{tt}\}}(x) = \chi_{\{\alpha=\mathsf{tt}\}}(x) \cdot f(x).$$

Since $\{\alpha = \mathsf{tt}\}$ is a clopen set, $\chi^{\natural}_{\{\alpha=\mathsf{tt}\}}$ and $\chi^{\natural}_{\{\alpha=\mathsf{ff}\}}$ are both continuous maps from X to $\mathbf{S}(X)$. We assume that we have a map

$$\mathcal{E} : \mathbf{Vars} \to (\mathbf{Exprs} \to (X \to \mathbf{S}(X)))$$

at our disposal so that for each variable a and each expression **E** the stochastic relation $\mathcal{E}(a)(\mathbf{E})$ is continuous. The map $\mathcal{E}$ is intended to model the assignment of **E** to variable a by evaluating the environment, performing the computation which is associated with the expression **E** and assigning the result to a.

2.2 Some Preparations

Modeling the semantics of the `while` loop requires some preparations. Denote for the stochastic relations $K, L : X \rightsquigarrow X$ their product $K*L$ through

$$(K*L)(x)(A) := \int_X L(y)(A)\ K(x)(dy)$$

[1] Ludwig Wittgenstein writes in his *Philosophische Untersuchungen*: “Wer in ein fremdes Land kommt, wird manchmal die Sprache der Einheimischen durch hinweisende Erklärungen lernen, die sie ihm geben; und er wird die Deutung dieser Erklärungen oft *raten* müssen und manchmal richtig, manchmal falsch raten” [28, Nr. 32, p.29]

(see Section 3 which relates $K*L$ to the Kleisli product for stochastic relations K and L). Fix a stochastic relation β, and define for the Boolean α and for the relation $L : X \rightsquigarrow X$ inductively the sequence $(\tau_k^\alpha(\beta, L))_{k\geq 0}$ of relations through

$$\begin{aligned}\tau_0^\alpha(\beta, L) &:= \chi^\natural_{\{\alpha=\mathsf{ff}\}} * L,\\ \tau_{k+1}^\alpha(\beta, L) &:= \left(\chi^\natural_{\{\alpha=\mathsf{tt}\}} * \beta\right) * \tau_k^\alpha(\beta, L),\end{aligned}$$

and put

$$\tau^\alpha(\beta, L) := \sum_{k=0}^{\infty} \tau_k^\alpha(\beta, L).$$

Note that an explicit representation of the inductive step is

$$\tau_{k+1}^\alpha(\beta, L)(x)(A) := \chi_{\{\alpha=\mathsf{tt}\}}(x) \cdot \int_X \tau_k^\alpha(\beta, L)(y)(A)\ \beta(x)(dy),$$

whenever $x \in X, A \in \mathcal{B}(X)$. This implies

$$\int_X f\ d\tau_{k+1}^\alpha(\beta, L)(x) = \chi_{\{\alpha=\mathsf{tt}\}}(x) \cdot \int_X \int_X f(z)\ \tau_k^\alpha(\beta, L)(y)(dz)\ \beta(x)(dy),$$

whenever $f : X \to \mathbb{R}$ is bounded and measurable. Using this, one can show that τ_k^α may be explicitly written as

$$\tau_k^\alpha(\beta, L)(x)(A) =$$

$$\chi^\natural_{\{\alpha=\mathsf{tt}\}}(x) \cdot \overbrace{\int_X \cdots \int_X}^{k \text{ integrals}} L(x_k)\left(\{\alpha = \mathsf{tt}\}^{k-1} \times \{\alpha = \mathsf{ff}\} \times A\right) \beta(x_{k-1})(dx_k) \cdots$$
$$\cdots \beta(x_1)(dx_2)\ \beta(x)(dx_1)$$

Proposition 1 *$\tau_k^\alpha(\beta, L)$ is for each $k \geq 0$ a stochastic relation, so is $\tau^\alpha(\beta, L)$. If both β and L are continuous, then $\tau_k^\alpha(\beta, L)$ and $\tau^\alpha(\beta, L)$ are, too.*

Proof 1. We first show that $\tau_k^\alpha(\beta, L)$ is a stochastic relation by induction on k. The case $k = 0$ is trivial, and the inductive step notes first that $A \mapsto \tau_{k+1}^\alpha(\beta, L)(x)(A)$ is a sub-probability, in particular it is σ-additive (use e.g. the Monotone Convergence Theorem [3, Theorem 16.2]). A similar argument implies also that $A \mapsto \tau^\alpha(\beta, L)(x)(A)$ is a sub-probability. Using standard arguments again, measurability of $x \mapsto \tau_{k+1}^\alpha(\beta, L)(x)(A)$ as well as $x \mapsto \tau^\alpha(\beta, L)(x)(A)$ for each $A \in \mathcal{B}(X)$ is inferred.
2. Now assume that β and L are both continuous. We argue again by induction on k. Because $\{\alpha = \mathsf{tt}\}$ is clopen, the case $k = 0$ is established. For the inductive step, note that

$$\varphi_{k,f} : x \mapsto \chi_{\{\alpha=\mathsf{tt}\}}(x) \cdot \int_X f(z)\ \tau_k^\alpha(\beta, L)(x)(dz)$$

constitutes a continuous map, whenever f is continuous. This is so by the induction hypothesis and the clopenness of $\{\alpha = \mathsf{tt}\}$. But because β is continuous, and because

$$\int_X f\ d\tau_{k+1}^\alpha(\beta, L)(x) = \int_X \varphi_{k,f}\ d\beta(x)$$

holds, continuity is established.
3. Continuity of $\tau^\alpha(\beta, L)$ is proved through the Portmanteau Theorem (see Section 3) by showing that

$$\liminf_{n\to\infty} \tau^\alpha(\beta, L)(x_n)(G) \geq \tau^\alpha(\beta, L)(x_0)(G)$$

whenever $x_n \to x_0$ and $G \subseteq X$ is an open set. In fact, by Fatou's Lemma [3, Theorem 16.3] and the continuity of all $\tau_k^\alpha(\beta, L)$ we see

$$\begin{aligned}\liminf_{n\to\infty} \tau^\alpha(\beta, L)(x_n)(G) &\geq \sum_{k\geq 0} \liminf_{n\to\infty} \tau^\alpha(\beta, L)(x_n)(G) \\ &\geq \sum_{k\geq 0} \tau_k^\alpha(\beta, L)(x_0)(G) \\ &= \tau^\alpha(\beta, L)(x_0)(G),\end{aligned}$$

establishing the desired inequality. □
The properties of $\tau^\alpha(\beta, L)$ noted below will be most helpful in the sequel:

Proposition 2 *The transformation $\tau^\alpha(\beta, \cdot)$ has the following properties:*

1. *$\tau^\alpha(\beta, L)$ is a fixed point of $\tau^\alpha(\beta, \cdot)$,*
2. *Suppose that $L_1, L_2 : X \rightsquigarrow X$ are stochastic relations such that $\chi^\natural_{\{\alpha=\mathsf{ff}\}} * L_1 = \chi^\natural_{\{\alpha=\mathsf{ff}\}} * L_2$, then $\tau^\alpha(\beta, L_1) = \tau^\alpha(\beta, L_2)$.*

Proof 1. A straightforward computation shows

$$\begin{aligned}\tau_0^\alpha\left(\beta, \tau^\alpha(\beta, L)\right)(x) &= \chi_{\{\alpha=\mathsf{ff}\}}(x) \cdot \tau^\alpha(\beta, L)(x) \\ &= \chi_{\{\alpha=\mathsf{ff}\}}(x) \cdot \tau_0^\alpha(\beta, L)(x) \\ &= \chi_{\{\alpha=\mathsf{ff}\}}(x) \cdot L(x) \\ &= \tau_0^\alpha(\beta, L)(x),\end{aligned}$$

and inductively one sees that

$$\tau_k^\alpha\left(\beta, \tau^\alpha(\beta, L)\right)(x) = \tau_k^\alpha(\beta, L)$$

for each $k \geq 1$. This establishes the first part.
2. The second part demonstrates first that $\tau_0^\alpha(\beta, L_1) = \tau_0^\alpha(\beta, L_2)$ holds and then proceeds inductively again. □
The second part of Proposition 2 shows that the behavior of the fixed point is uniquely determined by the behavior of its argument on the set $\{\alpha = \mathsf{ff}\}$ which is comprised of all those elements of X on which α is false.

2.3 Partial Correctness Semantics

A state is a continuous stochastic relation, and the semantics will model state transformations. Thus we perceive a program as a device which produces a stochastic relation from another one, in this way performing a transformation between states. More formally, We will show how a semantic function $\mathcal{C}[\![\cdot]\!]$ operates on these stochastic relations. We associate with each program statement β a continuous stochastic relation, again denoted by β, subject to the rules outlined below.

1. Clearly, `skip` works as the identity:

$$\mathcal{C}[\![\texttt{skip}]\!]g := g$$

2. The sequential execution of statements corresponds to the Kleisli product of single executions:

$$\mathcal{C}[\![\beta_1;\beta_2]\!] := \mathcal{C}[\![\beta_1]\!]*\mathcal{C}[\![\beta_2]\!].$$

3. The assignment is given through the evaluation function:

$$\mathcal{C}[\![a := \mathbf{E}]\!]g := \mathcal{E}(a)(\mathbf{E})*g.$$

4. The conditional statement permits operating on the true branch separately from the false branch:

$$\mathcal{C}[\![\texttt{if } \alpha \texttt{ then } \beta \texttt{ else } \gamma \texttt{ fi}]\!]g := \chi^\natural_{\{\alpha=\mathsf{tt}\}}*\beta*g + \chi^\natural_{\{\alpha=\mathsf{ff}\}}*\gamma*g$$

Note that we obtain as a special case $\mathcal{C}[\![\texttt{if } \alpha \texttt{ then } \beta \texttt{ fi}]\!]g = \chi^\natural_{\{\alpha=\mathsf{tt}\}}*\beta*g + \chi^\natural_{\{\alpha=\mathsf{ff}\}}*g$.

5. Finally the transformation of the `while`- statement is modeled. Suppose that we are in state g, then the resulting state after executing the statement not at all will be

$$K_0 := \chi^\natural_{\{\alpha=\mathsf{ff}\}}*g,$$

executing the loop exactly $n+1$ times will result in the state

$$K_{n+1} = \chi^\natural_{\{\alpha=\mathsf{tt}\}}*(\beta*K_n) + \chi^\natural_{\{\alpha=\mathsf{ff}\}}*K_n$$

An easy inductive argument shows that

$$K_n = \tau^\alpha_n(\beta, g)$$

holds for each $n \geq 0$: For $n = 0$ we obtain this identity directly from the definition. The induction step proceeds in this way:

$$\begin{aligned} K_{n+1}(x)(A) &= \left(\chi^\natural_{\{\alpha=\mathsf{tt}\}}*(\beta*K_n) + \chi^\natural_{\{\alpha=\mathsf{ff}\}}*K_n\right)(x)(A) \\ &= \chi^\natural_{\{\alpha=\mathsf{tt}\}}(x)\cdot\int_X K_k(y)(A)\ \beta(x)(dy) + \chi^\natural_{\{\alpha=\mathsf{ff}\}}(x)\cdot g(x)(A) \\ &= \chi^\natural_{\{\alpha=\mathsf{tt}\}}(x)\cdot\left(\sum_{j=0}^{n}\int_X \tau^\alpha_j(\beta, L)(y)(A)\ \beta(x)(dy)\right) + \tau_0(\beta, L)(x)(A) \\ &= \sum_{j=1}^{n+1}\tau^\alpha_j(\beta, L)(x)(A) + \tau_0(\beta, L)(x)(A) \end{aligned}$$

Consequently we put

$$\mathcal{C}[\![\texttt{while } \alpha \texttt{ do } \beta \texttt{ od}]\!]g := \tau^\alpha(\beta, g).$$

An easy induction on the length of a program P yields

Proposition 3 *The semantic function $\mathcal{C}[\![P]\!]$ maps continuous stochastic relations to continuous stochastic relations.*

Remark 2 *For later use it is noted that $\mathcal{C}[\![P]\!]g$ can be represented as the action $\beta_P * g$ of some stochastic relation β_P on g. —*

This follows by induction on the length of program P from the definition of the semantic function.
The following simple example illustrates the concept.

Example 1 Let $X := [0,1]^n$ be the input space to heapsort, $H := \{x \in X \mid x \text{ is a heap}\}$ be the output of Floyd's algorithm for heap construction. Recall that $\langle x_1, \ldots, x_n\rangle \in X$ is a heap iff $x_{\lfloor i/2 \rfloor} \leq x_i$ for $2 \leq i \leq n$. Denote by λ Lebesgue measure on X, so λ is the uniform distribution on the input space. It is not difficult to show that $H \in \mathcal{B}(X)$, and that

$$\lambda(H) = F_n := \mid \{\pi \mid \pi \text{ is a permutation of } 1, ..., n \text{ with the heap property}\} \mid^{-1}$$

holds (see [6]). Now let $g_0(x) := \lambda$ for each $x \in X$, so that g_0 is a constant stochastic relation, and put $g_1(x)(A) := F_n \cdot \lambda(A \cap H)$, for each $x \in X$, hence g_1 is also a constant relation. Then [6] shows that

$$\mathcal{C}[\![P]\!]g_0 = g_1$$

where P is the formulation of Floyd's algorithm as a Ludwig program. —

2.4 A Partial Correctness Logic

We will give rules which permit statements of the form $f\ \{P\}\ g$ indicating that if execution of P started in state f, and if P terminates, then the computation will be in state g. Here states are continuous stochastic relations.

1. `skip` does not change much:
$$f\ \{\texttt{skip}\}\ f$$

2. Compositionality is preserved:
$$\frac{f\ \{\beta\}\ g \quad g\ \{\gamma\}\ h}{f\ \{\beta;\gamma\}\ h}$$

3. The assignment is modeled through the evaluation $\mathcal{E}$
$$\mathcal{E}(a)(\mathbf{E}) * g\ \{a := \mathbf{E}\}\ g$$

4. The predicate transformation through the conditional statement models the composition through both branches
$$\frac{f\ \{\beta\}\ g \quad h\ \{\gamma\}\ g}{\left(\chi^{\natural}_{\{\alpha=\texttt{tt}\}} * f + \chi^{\natural}_{\{\alpha=\texttt{ff}\}} * h\right)\ \{\texttt{if}\ \alpha\ \texttt{then}\ \beta\ \texttt{else}\ \gamma\ \texttt{fi}\}\ g}$$

5. Finally, the `while`−statement composes the transformations in each step and puts them together to the fixed point:
$$\tau^{\alpha}(\beta, g)\{\texttt{while}\ \alpha\ \texttt{do}\ \beta\ \texttt{od}\}\ g$$

As usual, a *triplet* $\langle f, P, g\rangle$ is just $f\ \{P\}\ g$, and a *proof* of triplet $\mathcal{T}_\ell$ is a sequence of triplets $\mathcal{T}_0, \ldots, \mathcal{T}_\ell$ where each triplet $\mathcal{T}_i$ follows either the previous one by one of the rules above, or is one of the axioms. Denote by $P, f \vdash g$ that there is a proof for triplet $\langle f, P, g\rangle$.

2.5 Consistency and Completeness

Proposition 4 *Given a program P and two continuous stochastic relations f and g, the following statements are equivalent*

1. $P, f \vdash g$,
2. $\mathcal{C}[\![P]\!]g = f$.

Proof 1. (Consistency) We show that if $P, f \vdash g$ then $\mathcal{C}[\![P]\!]g = f$ by induction on the length of program P. For the induction step it is enough to assume that the statement under consideration is the conditional statement such that the embedded statements are already shown to satisfy the requirement already.
Now assume that

$$f \ \{\texttt{if}\ \alpha\ \texttt{then}\ \beta\ \texttt{else}\ \gamma\ \texttt{fi}\}\ g,$$

then

$$f = \chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * h_1 + \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * h_2$$

such that both $h_1\ \{\beta\}\ g$ and $h_2\ \{\gamma\}\ g$. Consequently,

$$\begin{aligned} \mathcal{C}[\![\texttt{if}\ \alpha\ \texttt{then}\ \beta\ \texttt{else}\ \gamma\ \texttt{fi}]\!]g &= \chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * \beta * g + \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * \gamma * g \\ &= \chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * h_1 + \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * h_2 \\ &= f. \end{aligned}$$

2. (Completeness) We show that $\mathcal{C}[\![P]\!]f = g$ implies $P, g \vdash f$ This is also established by an inductive proof, and again it is enough to focus on the conditional statement. In fact, assume that $\mathcal{C}[\![\texttt{if}\ \alpha\ \texttt{then}\ \beta\ \texttt{else}\ \gamma\ \texttt{fi}]\!]g = g'$ holds, thus $g' = \chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * \beta * g + \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * \gamma * g$. The latter equation may be decomposed to

$$\begin{aligned} \chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * g' &= \chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * \beta * g \\ \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * g' &= \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * \gamma * g. \end{aligned}$$

We can find stochastic relations h_1 and h_2 such that

$$\beta * g = \chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * g' + \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * h_1$$

and

$$\gamma * g = \chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * h_2 + \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * g',$$

so that the induction hypothesis yields

$$\left(\chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * g' + \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * h_1\right)\ \{\beta\}\ g$$

and

$$\left(\chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * h_2 + \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * g'\right)\ \{\gamma\}\ g.$$

Applying the induction hypothesis again, we see

$$\left(\chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * \left(\chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * g' + \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * h_1\right) + \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * \left(\chi^{\natural}_{\{\alpha=\mathsf{tt}\}} * h_2 + \chi^{\natural}_{\{\alpha=\mathsf{ff}\}} * g'\right)\right) \\ \{\texttt{if}\ \alpha\ \texttt{then}\ \beta\ \texttt{else}\ \gamma\ \texttt{fi}\}g$$

which is plainly equivalent to $g'\ \{\texttt{if}\ \alpha\ \texttt{then}\ \beta\ \texttt{else}\ \gamma\ \texttt{fi}\}\ g$, as claimed. □

3 The Giry Monad

In this section the constructions underlying the Giry monad are collected. We remind the reader of Polish spaces, of the topology of weak convergence on the space of all subprobabilities on a Polish space, and finally of the monad investigated by Giry.

Let X be a Polish space, i.e., a separable metric space for which a complete metric exists, and denote by $\mathbf{S}(X)$ the set of all subprobability measures on the Borel sets $\mathcal{B}(X)$ of X. The *weak topology* on $\mathbf{S}(X)$ is the smallest topology which makes $\tau \mapsto \int_X f \, d\tau$ continuous, whenever $f \in \mathcal{C}(X) := \{g : X \to \mathbb{R} \mid g \text{ is bounded and continuous}\}$. It is well known that the discrete measures are dense, and that $\mathbf{S}(X)$ is a Polish space with this topology [22, Section II.6]. Let d be the metric on X, and put $d(x, A) := \inf\{d(x, y) \mid y \in X\}$ the distance of $x \in X$ to the subset $A \subseteq X$, then the Prohorov metric d_P on $\mathbf{S}(X)$ is defined through

$$d_P(\tau_1, \tau_2) := \inf\{\varepsilon > 0 \mid \forall A \in \mathcal{B}(X) : \tau_1(A) \leq \tau_2(A^\varepsilon) + \varepsilon \wedge \tau_1(A) \leq \tau_2(A^\varepsilon) + \varepsilon\}$$

with $A^\varepsilon := \{y \in X \mid d(y, A) < \varepsilon\}$ as the set of all elements of X having distance less that ε from A. This metric topologizes the topology of weak convergence, see [3, Theorem 6.8].

More explicitly, a sequence $(\tau_n)_{n\in\mathbb{N}}$ with $\tau_n \in \mathbf{S}(X)$ converges to $\tau_0 \in \mathbf{S}(X)$ in this topology (indicated by $\tau_n \rightharpoonup_w \tau_0$) iff

$$\forall f \in \mathcal{C}(X) : \int_X f \, d\tau_n \to \int_X f \, d\tau_0$$

holds. The famous Portmanteau Theorem [22, II.6.1] states that this is equivalent to the condition

$$\liminf_{n\to\infty} \tau_n(G) \geq \tau_0(G)$$

whenever $G \subseteq X$ is an open set. We will assume throughout that $\mathbf{S}(X)$ is endowed with the weak topology.

Denote by $\mathfrak{Pol}$ the category of Polish spaces with continuous maps as morphisms. $\mathbf{S}$ assigns to each Polish space X the space of subprobability measures on X; if $f : X \to Y$ is a morphism in $\mathfrak{Pol}$, its image $\mathbf{S}(f) : \mathbf{S}(X) \to \mathbf{S}(Y)$ is defined through

$$\mathbf{S}(f)(\tau)(B) := \tau\left(f^{-1}[B]\right),$$

where $\tau \in \mathbf{S}(X)$ and $B \in \mathcal{B}(Y)$ is a Borel set. By virtue of the *Change of Variable Formula*

$$\int_Y g \, d\mathbf{S}(f)(\tau) = \int_X g \circ f \, d\tau$$

it is easy to see that $\mathbf{S}(f)$ is continuous. Thus $\mathbf{S} : \mathfrak{Pol} \to \mathfrak{Pol}$ is functor.

Denote by $\mu_X : \mathbf{S}(\mathbf{S}(X)) \to \mathbf{S}(X)$ the map

$$\mu_X(M)(A) := \int_{\mathbf{S}(X)} \tau(A) \, M(d\tau)$$

which assigns to each measure M on the Borel sets of $\mathbf{S}(X)$ a subprobability measure $\mu_X(M)$ on the Borel sets of X. Thus $\mu_X(M)(A)$ averages over the subprobabilities for A using measure M. Standard arguments show that

$$\int_X f \, d\mu_X(M) = \int_{\mathbf{S}(X)} \left(\int_X f \, f\tau\right) \, M(d\tau)$$

for each measurable and bounded map $f : X \to \mathbb{R}$.
The map μ_X is a morphism in $\mathfrak{Pol}$, as the following Lemma shows.

Lemma 1 $\mu_X : \mathbf{S}(\mathbf{S}(X)) \to \mathbf{S}(X)$ *is continuous.*

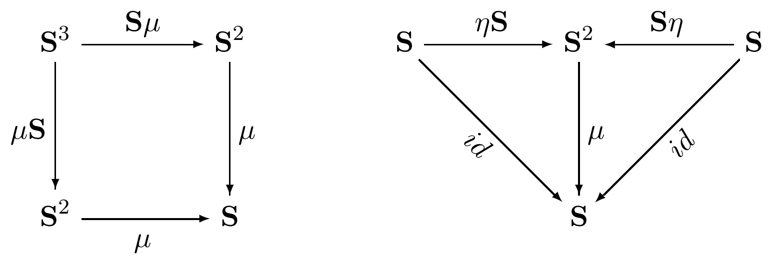
Proof Let $(M_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbf{S}(\mathbf{S}(X))$ with $M_n \rightharpoonup_w M_0$, then we get for $f \in \mathcal{C}(X)$ through the Change of Variable Formula, and because

$$\tau \mapsto \int_X f \; d\tau$$

is a member of $\mathcal{C}(\mathbf{S}(X))$, the following chain

$$\begin{aligned} \int_{\mathbf{S}(X)} f \; d\mu_X(M_n) &= \int_{\mathbf{S}(X)} \left(\int_X f \; d\tau \right) \; M_n(d\tau) \\ &\to \int_{\mathbf{S}(X)} \left(\int_X f \; d\tau \right) \; M_0(d\tau) \\ &= \int_{\mathbf{S}(X)} f \; d\mu_X(M_0). \end{aligned}$$

Thus $\mu_X(M_n) \rightharpoonup_w \mu_X(M_0)$ is established, as desired. □
The argumentation in [13] shows that $\mu : \mathbf{S}^2 \stackrel{\bullet}{\to} \mathbf{S}$ is a natural transformation. Together with $\eta_X : X \to \mathbf{S}(X)$, which assigns to each $x \in X$ the Dirac measure δ_x on x, and which is a natural transformation $\eta : \mathbb{1} \stackrel{\bullet}{\to} \mathbf{S}$, the triplet $\langle \mathbf{S}, \eta, \mu \rangle$ forms a monad [13]. It was originally proposed and investigated by Giry and will be referred to as the *Giry monad*. This means that these diagrams commute in the category of endofunctors of $\mathfrak{Pol}$ with natural transformations as morphisms:



Definition 2 *A stochastic relation $K : X \rightsquigarrow Y$ between the Polish spaces X and Y is a Kleisli morphism for the Giry monad.*

Equivalently, a stochastic $K : X \rightsquigarrow Y$ may be represented as a map $K : X \to \mathbf{S}(Y)$ with the following properties:

1. $x \mapsto K(x)$ is (weakly) continuous,
2. $B \mapsto K(x)(B)$ constitutes a subprobability measure on the Borel sets $\mathcal{B}(Y)$ of Y for each $x \in X$.

The composition $K \circ L$ of stochastic relations $K : X \rightsquigarrow Y$ and $L : Y \rightsquigarrow Z$ is the Kleisli product, thus

$$(L \circ K)(x)(C) = \int_Y L(y)(C) \; K(x)(dy)$$

holds for $x \in X$ and $C \in \mathcal{B}(Z)$. We use in Section 2 for the semantics of Ludwig the notation

$$K*L := L \circ K,$$

because this is intuitively more appealing for the purposes of the semantics.
If the probability measures $\mathbf{P}(X)$ on X are considered, then we get the monad $\langle \mathbf{P}, \eta, \mu \rangle$ which is actually the monad that was investigated by Giry. We will concentrate on the subprobability functor $\mathbf{S}$ with occasional sidelong glances to the probability functor. The former one is a bit more convenient to work with because $\mathbf{S}(X)$ is positive convex.

4 Positive Convex Structures

Suppose X can be embedded into a real vector space as a positive convex set. Thus $c_1 \cdot x_1 + c_2 \cdot x_2 \in X$ whenever $x_1, x_2 \in X$, and c_1, c_2 are positive convex coefficients. We call then scalar multiplication and addition a *positive convex structure* on X. Such a positive convex structure is continuous iff $c_n \to c_0, c'_n \to c'_0$ and $x_n \to x_0, x'_n \to x'_0$ together imply $c_n \cdot x_n + c'_n \cdot x'_n \to c_0 \cdot x_0 + c'_0 \cdot x'_0$, as n approaches infinity (this is just like a topological vector space, postulating continuity of addition and scalar multiplication).
Put for the rest of the paper

$$\Omega := \{\langle \alpha_1, \ldots, \alpha_n \rangle \mid n \in \mathbb{N}, \alpha_i \geq 0, \sum_{i=1}^{n} \alpha_i \leq 1\}.$$

We call the elements of Ω *positive convex tuples* or simply *positive convex.*
Formally, we define a positive convex structure following e.g. Pumplün [24].

Definition 3 *A* positive convex structure $\mathcal{P}$ *has for each* $\alpha = \langle \alpha_1, \ldots, \alpha_n \rangle \in \Omega$ *a continuous map* $\alpha_{\mathcal{P}} : X^n \to X$ *which we write as*

$$\alpha_{\mathcal{P}}(x_1, \ldots, x_n) := \sum_{1 \leq i \leq n}^{\mathcal{P}} \alpha_i \cdot x_i,$$

such that

1. $\sum_{1 \leq i \leq n}^{\mathcal{P}} \delta_{i,k} \cdot x_i = x_k$, *where* $\delta_{i,j}$ *is Kronecker's* δ,
2. *the identity*

$$\sum_{1 \leq i \leq n}^{\mathcal{P}} \alpha_i \cdot \left(\sum_{1 \leq k \leq m}^{\mathcal{P}} \beta_{i,k} \cdot x_k \right) = \sum_{1 \leq k \leq m}^{\mathcal{P}} \left(\sum_{1 \leq i \leq n}^{\mathcal{P}} \alpha_i \beta_{i,k} \right) \cdot x_k$$

   holds whenever $\langle \alpha_1, \ldots, \alpha_n \rangle, \langle \beta_{i,k}, \ldots, \beta_{i,k} \rangle \in \Omega, 1 \leq i \leq n$.

It can be shown that the usual rules for manipulating sums in vector spaces apply, e.g. $1 \cdot x = x, 0 \cdot x + c \cdot x' = c \cdot x'$, or associativity. Thus we will use freely the notation from vector spaces, omitting in particular the explicit reference to the structure whenever possible.

A morphism $\theta : X_1 \to X_2$ between continuous positive convex structures $\mathcal{P}_1$ and $\mathcal{P}_2$ on X_1 resp. X_2 is a continuous and *affine* map. Thus

$$\theta\left(\sum_{1\leq i\leq n}^{\mathcal{P}_1} \alpha_i \cdot x_i\right) = \sum_{1\leq i\leq n}^{\mathcal{P}_2} \alpha_i \cdot \theta(x_i)$$

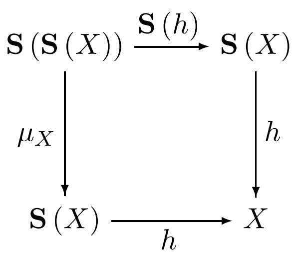
holds for $x_1, \ldots, x_n \in X$ and $\langle \alpha_1, \ldots, \alpha_n \rangle \in \Omega$. Positive convex structures with their morphisms form a subcategory $\mathfrak{StrConv}$ of $\mathfrak{Pol}$.

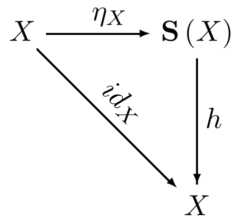
5 Characterizing the Algebras

The Eilenberg-Moore algebras are represented through partitions and through smooth equivalence relations, both on the respective space of subprobability measures. We first deal with partitions and investigate the partition induced by an algebra. This leads to a necessary and sufficient condition for a partition to be generated from an algebra which in turn can be used for characterizing the category of these algebras by introducing a suitable notion of morphisms for partitions. The second representation capitalizes on the fact that equivalence relations induced by continuous maps (as special cases of Borel measurable maps) have some rather convenient properties in terms of measurability. This is used for an alternative description of the category of all algebras.

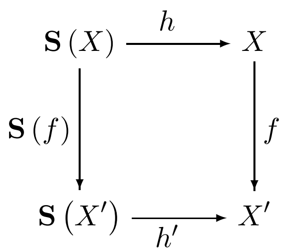
5.1 Algebras

An *Eilenberg-Moore algebra* $\langle X, h \rangle$ for the Giry monad is an object X in $\mathfrak{Pol}$ together with a morphism $h : \mathbf{S}(X) \to X$ such that the following diagrams commute





When talking about algebras, we refer always to Eilenberg-Moore algebras for the Giry monad, unless otherwise indicated. An *algebra morphism* $f : \langle X, h \rangle \to \langle X', h' \rangle$ between the algebras $\langle X, h \rangle$ and $\langle X', h' \rangle$ is a continuous map $f : X \to X'$ which makes the diagram



commute. Algebras together with their morphisms form a category $\mathfrak{Alg}$. This construction is discussed for monads in general in [17, Chapter IV.2].

Remark 3 *Looking aside, we mention briefly a well-known monad in the category* $\mathfrak{Set}$ *of sets with maps as morphisms. The functor* $\mathcal{P}$ *assigns each set* A *its power set* $\mathcal{P}(A)$*, and if* $f : A \to B$ *is a map,* $\mathcal{P}(f) : \mathcal{P}(A) \to \mathcal{P}(B)$ *assigns each subset* $A_0 \subseteq A$ *its image* $f[A_0]$*, thus* $\mathcal{P}(f)(A_0) = f[A_0]$*. Define the natural transformation* $\mu : \mathcal{P}^2 \overset{\bullet}{\to} \mathcal{P}$ *through*

$$\mu_A : \mathcal{P}(\mathcal{P}(A)) \ni M \mapsto \bigcup M \in \mathcal{P}(A),$$

and $\eta : \mathbb{1} \overset{\bullet}{\to} \mathcal{P}$ *through* $\eta_X : x \mapsto \{x\}$*, then the triplet* $\langle \mathcal{P}, \eta, \mu \rangle$ *forms a monad (the* Manes monad*). It is well known that the algebras for this monad may be identified with the complete* sup*-semi lattices [17, Exercise VI.2.1], cp. Remark 1. —*

For the rest of this paper each free occurrence of X refers to a Polish space.
We need some elementary properties for later reference. They are collected in the next Lemma.

Lemma 2 *1. Let* $f : A \to B$ *be a map between the Polish spaces* A *and* B*, and let* $\tau = \alpha_1 \cdot \delta_{a_1} + \alpha_2 \cdot \delta_{a_2}$ *be the linear combination of Dirac measures for* $a_1, a_2 \in A$ *with positive convex* α_1, α_2*. Then* $\mathbf{S}(f)(\tau) = \alpha_1 \cdot \delta_{f(a_1)} + \alpha_2 \cdot \delta_{f(a_2)}$*.*

2. Let τ_1, τ_2 *be subprobability measures on* X*, and let* $M = \alpha_1 \cdot \delta_{\tau_1} + \alpha_2 \cdot \delta_{\tau_2}$ *be the linear combination of the corresponding Dirac measures in* $\mathbf{S}(\mathbf{S}(X))$ *with positive convex coefficients* α_1, α_2*. Then* $\mu_X(M) = \alpha_1 \cdot \tau_1 + \alpha_2 \cdot \tau_2$*.*

Proof The first part follows directly from the observation that $\delta_x(f^{-1}[D]) = \delta_{f(x)}(D)$, and the second one is easily inferred from

$$\begin{aligned} \mu_X(\delta_\tau)(Q) &= \int_{\mathbf{S}(X)} \rho(Q)\ \delta_\tau(d\rho) \\ &= \tau(Q) \end{aligned}$$

for each Borel subset $Q \subseteq X$, and from the linearity of the integral. □

5.2 Positive Convex Partitions

Assume that the pair $\langle X, h \rangle$ is an algebra, and define for each $x \in X$

$$G_h(x) := \{\tau \in \mathbf{S}(X) \mid h(\tau) = x\} \left(= h^{-1}[\{x\}]\right).$$

Then $G_h(x) \neq \emptyset$ for all $x \in X$ due to h being onto. The algebra h will be characterized through properties of the set-valued map G_h. Define the *weak inverse* $\exists R$ for a set-valued map $R : X \to \mathcal{P}(Y) \setminus \{\emptyset\}$ with non-empty images through

$$\exists R(W) := \{x \in X \mid R(x) \cap W \neq \emptyset\}.$$

for $W \subseteq Y$. If Y is a topological space, if R takes closed values, and if $\exists R(W)$ is compact in X whenever $W \subseteq Y$ is compact, then R is called *k-upper-semicontinuous* (abbreviated as k.u.s.c.). If Y is compact, this is the usual notion of upper-semicontinuity (cf. [26, Section 5.1]).
The importance of being k.u.s.c. becomes clear at once from

Lemma 3 *Let* $f : A \to B$ *be a surjective map between the Polish spaces* A *and* B*, and put* $G_f(b) := f^{-1}[\{b\}]$ *for* $b \in B$*. Then* f *is continuous iff* G_f *is k.u.s.c.*

Proof A direct calculation for the weak inverse shows $\exists G_f(A_0) = f[A_0]$ for each subset $A_0 \subseteq A$. The assertion now follows from the well-known fact that a map between metric spaces is continuous iff it maps compact sets to compact sets. □
Applying this observation to the set-valued map G_h, we obtain:

Proposition 5 *The set-valued map $x \mapsto G_h(x)$ has the following properties:*

1. *$\delta_x \in G_h(x)$ holds for each $x \in X$.*
2. *$\mathcal{G}_h := \{G_h(x) \mid x \in X\}$ is a partition of $\mathbf{S}(X)$ into closed and positive convex sets.*
3. *$x \mapsto G_h(x)$ is k.u.s.c.*
4. *Let $\sim_h$ be the equivalence relation on $\mathbf{S}(X)$ induced by the partition $\mathcal{G}_h$. If $\tau_1 \sim_h \tau_1'$ and $\tau_2 \sim_h \tau_2'$, then $(\alpha_1 \cdot \tau_1 + \alpha_2 \cdot \tau_2) \sim_h (\alpha_1 \cdot \tau_1' + \alpha_2 \cdot \tau_2')$ for the positive convex coefficients α_1, α_2.*

Proof Because $\{x\}$ is closed, and h is continuous, $G_h(x) = h^{-1}[\{x\}]$ is a closed subset of $\mathbf{S}(X)$. Because h is onto, every G_h takes non-empty values; it is clear that $\{G_h(x) \mid x \in X\}$ forms a partition of $\mathbf{S}(X)$. Because h is continuous, G_h is k.u.s.c. by Lemma 3. Positive convexity will follow immediately from part 4.
Assume that $h(\tau_1) = h(\tau_1') = x_1$ and $h(\tau_2) = h(\tau_2') = x_2$, and observe that $h(\delta_x) = x$ holds for all $x \in X$. Using Lemma 2, we get:

$$\begin{aligned} h(\alpha_1 \cdot \tau_1 + \alpha_2 \cdot \tau_2) &= (h \circ \mu_X)(\alpha_1 \cdot \delta_{\tau_1} + \alpha_2 \cdot \delta_{\tau_2}) \\ &= (h \circ \mathbf{S}(h))(\alpha_1 \cdot \delta_{\tau_1} + \alpha_2 \cdot \delta_{\tau_2}) \\ &= h\left(\alpha_1 \cdot \delta_{h(\tau_1)} + \alpha_2 \cdot \delta_{h(\tau_2)}\right) \\ &= h(\alpha_1 \cdot \delta_{x_1} + \alpha_2 \cdot \delta_{x_2}) \end{aligned}$$

In a similar way, $h(\alpha_1 \cdot \tau_1' + \alpha_2 \cdot \tau_2') = h(\alpha_1 \cdot \delta_{x_1} + \alpha_2 \cdot \delta_{x_2})$ is obtained. This implies the assertion. □
Thus $\mathcal{G}_h$ is invariant under taking positive convex combinations. It is a positive convex partition in the sense of the following definition.

Definition 4 *An equivalence relation ρ on $\mathbf{S}(X)$ is said to be* positive convex *iff $\tau_1 \,\rho\, \tau_1'$ and $\tau_2 \,\rho\, \tau_2'$ together imply $(\alpha_1 \cdot \tau_1 + \alpha_2 \cdot \tau_2) \;\rho\; (\alpha_1 \cdot \tau_1 + \alpha_2 \cdot \tau_2)$ for positive convex coefficients α_1, α_2. A partition of $\mathbf{S}(X)$ is called* positive convex *iff its associated equivalence relation is.*

Note that the elements of a positive convex partition form positive convex sets. The converse to Proposition 5 characterizes algebras:

Proposition 6 *Assume $\mathcal{G} = \{G(x) \mid x \in X\}$ is a positive convex partition of $\mathbf{S}(X)$ into closed sets indexed by X such that $\delta_x \in G(x)$ for each $x \in X$, and such that $x \mapsto G(x)$ is k.u.s.c. Define $h : \mathbf{S}(X) \to X$ through $h(\tau) = x$ iff $\tau \in G(x)$. Then $\langle X, h\rangle$ is an algebra for the Giry monad.*

Proof 1. It is clear that h is well defined and surjective, and that $\exists G(F) = h[F]$ holds for each subset $F \subseteq \mathbf{S}(X)$. Thus $h[K]$ is compact whenever K is compact, because G is k.u.s.c. Thus h is continuous by Lemma 3.

2. An easy induction establishes that h respects positive convex combinations: if $h(\tau_i) = h(\tau_i')$ for $i = 1, \ldots, n$, and if $\alpha_1, \ldots, \alpha_n$ are positive convex coefficients, then

$$h(\sum_{i=1}^{n} \alpha_i \cdot \tau_i) = h(\sum_{i=1}^{n} \alpha_i \cdot \tau_i').$$

We claim that $(h \circ \mu_X)(M) = (h \circ \mathbf{S}(h))(M)$ holds for each *discrete* $M \in \mathbf{S}(\mathbf{S}(X))$. In fact, let

$$M = \sum_{i=1}^{n} \alpha_i \cdot \delta_{\tau_i}$$

be such a discrete measure, then Lemma 2 implies that

$$\mu_X(M) = \sum_{i=1}^{n} \alpha_i \cdot \tau_i,$$

thus

$$(h \circ \mu_X)(M) = h\left(\sum_{i=1}^{n} \alpha_i \cdot \tau_i\right) = h\left(\sum_{i=1}^{n} \alpha_i \cdot \delta_{h(\tau_i)}\right) = (h \circ \mathbf{S}(h))(M),$$

because we know also from Lemma 2 that

$$\mathbf{S}(h)(M) = \sum_{i=1}^{n} \alpha_i \cdot \delta_{h(\tau_i)}$$

holds.

3. Since the discrete measures are dense in the weak topology, we find for $M_0 \in \mathbf{S}(\mathbf{S}(X))$ a sequence $(M_n)_{n\in\mathbb{N}}$ of discrete measures M_n with $M_n \rightharpoonup_w M_0$. Consequently, we get from the continuity of both h and μ_X (Lemma 1) together with the continuity of $\mathbf{S}(h)$

$$(h \circ \mu_X)(M_0) = \lim_{n\to\infty}(h \circ \mu_X)(M_n) = \lim_{n\to\infty}(h \circ \mathbf{S}(h))(M_n) = (h \circ \mathbf{S}(h))(M_0).$$

This proves the claim. □

We have established

Proposition 7 *The algebras $\langle X, h\rangle$ for the Giry monad for Polish spaces X are exactly the positive convex k.u.s.c. partitions $\{G(x) \mid x \in X\}$ into closed subsets of $\mathbf{S}(X)$ such that $\delta_x \in G(x)$ for all $x \in X$ holds.*

We characterize the category $\mathfrak{Alg}$ of all algebras for the Giry monad. To this end we package the properties of partitions representing algebras into the notion of a G-partition. They will form the objects of category $\mathfrak{GPart}$.

Definition 5 *$\mathcal{G}$ is called a* G-partition *for X iff*

1. *$\mathcal{G} = \{G(x) \mid x \in X\}$ is a positive convex partition for $\mathbf{S}(X)$ into closed sets indexed by X,*
2. *$\delta_x \in G(x)$ holds for all $x \in X$,*
3. *the set-valued map $x \mapsto G(x)$ is k.u.s.c.*

Define the objects of category $\mathfrak{GPart}$ as pairs $\langle X, \mathcal{G}\rangle$ where X is a Polish space, and $\mathcal{G}$ is a G-partition for X. A morphism f between $\mathcal{G}$ and $\mathcal{G}'$ will map elements of $G(x)$ to $G'(f(x))$ through its associated map $\mathbf{S}(f)$. Thus an element $\tau \in G(x)$ will correspond to an element $\mathbf{S}(f)(\tau) \in G'(f(x))$.

Definition 6 *A morphism for $\mathfrak{GPart}$ $f : \langle X, \mathcal{G}\rangle \to \langle X', \mathcal{G}'\rangle$ is a continuous map $f : X \to X'$ such that $G(x) \subseteq \mathbf{S}(f)^{-1}[G'(f(x))]$ holds for each $x \in X$.*

Define the functor $F : \mathfrak{Alg} \to \mathfrak{GPart}$ by associating each algebra $\langle X, h\rangle$ its Giry partition $F(X, h)$ according to Proposition 7. Assume that $f : \langle X, h\rangle \to \langle X', h'\rangle$ is a morphism in $\mathfrak{Alg}$, and let $\mathcal{G} = \{G(x) \mid x \in X\}$ and $\mathcal{G}' = \{G'(x') \mid x' \in X'\}$ be the corresponding partitions. Then the properties of an algebra morphism yield

$$\begin{aligned} \tau \in \mathbf{S}(f)^{-1}\left[G'(f(x))\right] &\Leftrightarrow \mathbf{S}(f)(\tau) \in G'(f(x)) \\ &\Leftrightarrow \left(h' \circ \mathbf{S}(f)\right)(\tau) = f(x) \\ &\Leftrightarrow (f \circ h)(\tau) = f(x). \end{aligned}$$

Thus $\tau \in \mathbf{S}(f)^{-1}[G'(f(x))]$, provided $\tau \in G(x)$. Hence f is a morphism in $\mathfrak{GPart}$ between $F(X, h)$ and $F(X', h')$. Conversely, let $f : \langle X, \mathcal{G}\rangle \to \langle X', \mathcal{G}'\rangle$ be a morphism in $\mathfrak{GPart}$ with $\langle X, \mathcal{G}\rangle = F(X, h)$ and $\langle X', \mathcal{G}'\rangle = F(X, h')$. Then

$$\begin{aligned} h(\tau) = x &\Leftrightarrow \tau \in G(x) \\ &\Rightarrow \mathbf{S}(f)(\tau) \in G'(f(x)) \\ &\Leftrightarrow h'(\mathbf{S}(f)(\tau)) = f(x), \end{aligned}$$

thus $h' \circ \mathbf{S}(f) = f \circ h$ is inferred. Hence f constitutes a morphism in category $\mathfrak{Alg}$. Summarizing, we have shown

Proposition 8 *The category $\mathfrak{Alg}$ of algebras for the Giry monad is isomorphic to the category $\mathfrak{GPart}$ of G-partitions.*

5.3 Smooth Relations

The characterization of algebras so far encoded the crucial properties into a partition of $\mathbf{S}(X)$, thus indirectly into an equivalence relation on that space. We can move directly to a particular class of these relations when looking at an alternative characterization of the algebras through smooth equivalence relations.

Definition 7 *An equivalence relation ρ on a Polish space A is called* smooth *iff there exists a Polish space B and a Borel measurable map $f : A \to B$ such that*

$$a_1 \, \rho \, a_2 \Leftrightarrow f(a_1) = f(a_2)$$

holds, thus ρ is just the kernel of f.

Smooth equivalence relations are a helpful tool in the theory of Borel sets [26]. They have some interesting properties that have been capitalized upon in the theory of labeled Markov transition processes [4] and stochastic relations [8, 9].

Some basic notations and constructions first: Denote for an equivalence relation ρ on A by A/ρ the factor space, i.e., the set of all equivalence classes $[a]_\rho$, and by

$$\varepsilon_\rho : A \to A/\rho$$

the canonical projection. If A is a Polish space, then let $\mathcal{T}/\rho$ be the final topology on A/ρ with respect to the given topology and ε_ρ, i.e., the largest topology on A/ρ which makes ε_ρ continuous. Clearly a map $g : A/\rho \to B$ for a topological space B is continuous with respect to $\mathcal{T}/\rho$ iff $g \circ \varepsilon_\rho : A \to B$ is continuous w.r.t. the given topologies. We will need this observation in the proof of Proposition 9.
Now let $\langle X, h\rangle$ be an algebra for the Giry monad. Obviously

$$\tau_1 \, \rho_h \, \tau_2 \Leftrightarrow h(\tau_1) = h(\tau_2)$$

defines a smooth equivalence relation on the Polish space $\mathbf{S}(X)$. Its properties are summarized in

Proposition 9 *The equivalence relation ρ_h is positive convex, each equivalence class $[\tau]_{\rho_h}$ is closed and positive convex, and the factor space $\mathbf{S}(X)/\rho_h$ is homeomorphic to X when the former is endowed with the topology $\mathcal{T}/\rho_h$.*

Proof 1. Positive convexity of ρ_h follows from the properties of h exactly as in the proof of Proposition 5, from this, also positive convexity of the classes is inferred. Continuity of h implies that the classes are closed sets.
2. Define $\chi_h([\tau]_{\rho_h}) := h(\tau)$ for $\tau \in \mathbf{S}(X)$. Then $\chi_h : X/\rho_h \to X$ is well defined and a bijection. Let $G \subseteq X$ be an open set, then $\varepsilon_{\rho_h}^{-1}\left[\chi_h^{-1}[G]\right] = h^{-1}[G]$. Because $\mathcal{T}/\rho_h$ is the largest topology on $\mathbf{S}(X)/\rho_h$ that renders ε_{ρ_h} continuous, and because $h^{-1}[G] \subseteq \mathbf{S}(X)$ is open by assumption, we infer that $\chi_h^{-1}[G]$ is $\mathcal{T}/\rho_h$-open. Thus χ_h is continuous. On the other hand, if $(x_n)_{n\in\mathbb{N}}$ is a sequence in X converging to $x_0 \in X$, then $\delta_{x_n} \rightharpoonup_w \delta_{x_0}$ in $\mathbf{S}(X)$, thus $[\delta_{x_n}]_{\rho_h} \to [\delta_{x_0}]_{\rho_h}$ in $\mathcal{T}/\rho_h$ by construction. Consequently χ_h^{-1} is also continuous. □
Thus each algebra induces a G-triplet in the following sense

Definition 8 *A* G-triplet *$\langle X, \rho, \chi\rangle$ is a Polish space X with a smooth and positive convex equivalence relation ρ on $\mathbf{S}(X)$ such that $\chi : \mathbf{S}(X)/\rho \to X$ is a homeomorphism with $\chi([\delta_x]_\rho) = x$ for all $x \in X$. Here $\mathbf{S}(X)/\rho$ carries the final topology with respect to the weak topology on $\mathbf{S}(X)$ and ε_ρ.*

Now assume that a G-triplet $\langle X, \rho, \chi\rangle$ is given. Define $h(\tau) := \chi([\tau]_\rho)$ for $\tau \in \mathbf{S}(X)$. Then $\langle X, h\rangle$ is an algebra for the Giry monad: $h(\delta_x) = x$ follows from the assumption, and because $h = \chi \circ \varepsilon_\rho$, holds, the map h is continuous. An argument very similar to that used in the proof of Proposition 5 shows that $h \circ \mu_X = h \circ \mathbf{S}(h)$ holds; this is so since ρ is assumed to be positive convex.

Definition 9 *The continuous map $f : X \to X'$ between the Polish spaces X and X' constitutes a* G-triplet morphism *$f : \langle X, \rho, \chi\rangle \to \langle X', \rho', \chi'\rangle$ iff these conditions hold:*

1. *$\tau \, \rho \, \tau'$ implies $\mathbf{S}(f)(\tau) \, \rho' \, \mathbf{S}(f)\,\tau'$,*

2. the diagram

$$\begin{array}{ccc}
\mathbf{S}(X)/\rho & \xrightarrow{\mathbf{S}(f)_{\rho,\rho'}} & \mathbf{S}(X')/\rho' \\
\downarrow \chi & & \downarrow \chi' \\
X & \xrightarrow[f]{} & X'
\end{array}$$

commutes, where

$$\mathbf{S}(f)_{\rho,\rho'}\left([\tau]_\rho\right) := [\mathbf{S}(f)(\tau)]_{\rho'}$$

G-triplets with their morphisms form a category $\mathfrak{GTrip}$.

Lemma 4 *Each algebra morphism $f : \langle X, h\rangle \to \langle X', h'\rangle$ induces a G-triplet morphism $f : \langle X, \rho_h, \chi_h\rangle \to \langle X', \rho_{h'}, \chi_{h'}\rangle$.*

Proof 1. It is an easy calculation to show that $\tau\,\rho_h\,\tau'$ implies $\mathbf{S}(f)(\tau)\,\rho_{h'}\,\mathbf{S}(f)(\tau)$. This is so because f is a morphism for the algebras.
2. Since for each $\tau \in \mathbf{S}(X)$ there exists $x \in X$ such that $[\tau]_{\rho_h} = [\delta_x]_{\rho_h}$ (in fact, $h(\tau)$ would do, because $h(\tau) = h\left(\delta_{h(\tau)}\right)$, as shown above), it is enough to demonstrate that

$$\chi'_{h'}\left(\mathbf{S}(f)_{\rho_h,\rho'_{h'}}\left([\delta_x]_{\rho_h}\right)\right) = f(\chi_h([\delta_x]_{\rho_h}))$$

is true for each $x \in X$. Because $\mathbf{S}(f)(\delta_x) = \delta_{f(x)}$, a little computation shows that both sides of the above equation boil down to $f(x)$. □
The morphisms between G-triplets are just the morphisms between algebras (when we forget that these games play in different categories).

Proposition 10 *Let $f : \langle X, \rho, \chi\rangle \to \langle X', \rho', \chi'\rangle$ be a morphism between G-triplets, and let $\langle X, h\rangle$ resp. $\langle X', h'\rangle$ be the associated algebras. Then $f : \langle X, h\rangle \to \langle X', h'\rangle$ is an algebra morphism.*

Proof Given $\tau \in \mathbf{S}(X)$ we have to show that $(f \circ h)(\tau)$ equals $(h' \circ \mathbf{S}(f))(\tau)$. Since $h(\tau) = \chi([\tau]_\rho)$, we obtain

$$\begin{aligned}
(f \circ h)(\tau) &= f\left(\chi([\tau]_\rho)\right) \\
&= \chi'\left(\mathbf{S}(f)_{\rho,\rho'}([\tau]_\rho)\right) \\
&= \chi'\left([\mathbf{S}(f)(\tau)]_{\rho'}\right) \\
&= (h' \circ \mathbf{S}(f))(\tau)
\end{aligned}$$

□
Putting all these constructions with their properties together, we obtain

Proposition 11 *The category $\mathfrak{Alg}$ of algebras for the Giry monad is isomorphic to the category $\mathfrak{GTrip}$ of G-triplets.*

For the probabilistic case we obtain a similar result:

Corollary 1 *The category of algebras for the Giry monad for the probability functor is isomorphic to the full subcategory of G-triplets $\langle X, \rho, \chi\rangle$ with a smooth and convex equivalence relation such that $\chi : \mathbf{P}(X)/\rho \to X$ is a homeomorphism.*

We will show now that $\mathfrak{StrConv}$ is isomorphic to $\mathfrak{Alg}$.
Given an algebra $\langle X, h\rangle$, define for $x_1, \ldots, x_n \in X$ and the positive convex coefficients $\langle \alpha_1, \ldots, \alpha_n\rangle \in \Omega$

$$\sum_{i=1}^{n} \alpha_i \cdot x_i := h(\sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i}),$$

then this defines a positive convex structure on X. Let conversely such a positive convex structure be given. We show that we can define a G-triplet from it. Let

$$\mathcal{T}_X := \{\sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i} \mid n \in \mathbb{N}, x_1, \ldots, x_n \in X, \langle \alpha_1, \ldots, \alpha_n\rangle \in \Omega\},$$

then $\mathcal{T}_X$ is dense in $\mathbf{S}(X)$. Put

$$h_0\left(\sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i}\right) := \sum_{i=1}^{n} \alpha_i \cdot x_i,$$

then $h_0 : \mathcal{T}_X \to X$ is uniformly continuous, because

$$d\left(h_0(\sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i}), h_0(\sum_{j=1}^{m} d_j \cdot \delta_{y_j})\right) \leq d_P\left(\sum_{i=1}^{n} \alpha_i \cdot x_i, \sum_{j=1}^{m} d_j \cdot y_j\right).$$

Define ρ_0 as the kernel of h_0, then ρ_0 is a smooth equivalence relation on $\mathcal{T}_X$, and it is not difficult to see that the set of topological closures $\{\left([t]_{\rho_0}\right)^{\mathsf{cl}} \mid t \in \mathcal{T}_X\}$ forms a partition of $\mathbf{S}(X)$:

1. the closures of different equivalence classes are disjoint,
2. given $\tau \in \mathbf{S}(X)$, one can find a sequence $(t_n)_{n\in\mathbb{N}}$ in $\mathcal{T}_X$ with $t_n \rightharpoonup_w \tau$. Since X is Polish, in particular complete, the sequence $(h_0(t_n))_{n\in\mathbb{N}}$ converges to some t_0, and because h_0 is uniformly continuous, one concludes that $\tau \in \left([t_0]_{\rho_0}\right)^{\mathsf{cl}}$. Thus each member of $\mathbf{S}(X)$ is in some class.

This yields an equivalence relation ρ on $\mathbf{S}(X)$. Uniform continuity of h_0 gives a unique continuous extension h of h_0 to $\mathbf{S}(X)$, thus ρ equals the kernel of h, hence ρ is a smooth equivalence relation, and it is evidently positive convex. Defining on $\mathbf{S}(X)/\rho$ the metric

$$D([\tau_1]_\rho, [\tau_2]_\rho) := d(h(\tau_1), h(\tau_2)),$$

it is rather immediate that

1. the metric space $(\mathbf{S}(X)/\rho, D)$ is homeomorphic to X with metric d,

2. the topology induced by the metric is just the final topology with respect to the weak topology on $\mathbf{S}(X)$ and ε_ρ.

It is clear that each affine and continuous map between positive convex structures gives rise to a morphisms between the corresponding G-triplets, and vice versa.
Thus we have established:

Proposition 12 *The category of* $\mathfrak{Alg}$ *of algebras for the Giry monad is isomorphic to the category* $\mathfrak{StrConv}$ *of positive convex structures with continuous affine maps as morphisms.*

For the probability functor we obtain

Corollary 2 *The category of algebras for the Giry monad for the probability functor is isomorphic to the full subcategory of continuous convex structures.*

This characterization has been known for the probability functor in the case that X is a compact Hausdorff space [11, 2.14] (the attribution to Swirszcz's work [27] in [11] is slightly unclear). The methods for the proof are, however, rather different: the compact case makes essential use of the right adjoint of the probability functor as a functor between the respective categories of compact Hausdorff spaces and compact convex sets. This adjoint is not yet characterized fully in the present situation. Thus Corollary 2 generalizes the characterization to Polish spaces.

5.4 Examples

This Section illustrates the concept and proposes some examples by looking at some well-known situations, thus most of this Section is not really new, probably apart from the proposed point of view. We first show that the monad carries for each Polish space an instance of an algebra with it. Then we prove that in the finite case an algebra exists only in the case of a singleton set. Finally a geometrically oriented example is discussed by investigating the barycenter of a probability in a compact and convex subset of $\mathbb{R}^n$. In each case it turns out the that convex structure associated with the algebra is the natural one.

Example 2 The pair $\langle \mathbf{S}(X), \mu_X \rangle$ is always an algebra. We know from Lemma 1 that $\mu_X : \mathbf{S}(\mathbf{S}(X)) \to \mathbf{S}(X)$ is continuous. Because $\langle \mathbf{S}, \eta, \mu \rangle$ is a monad, the natural transformation $\mu : \mathbf{S}^2 \stackrel{\bullet}{\to} \mathbf{S}$ satisfies

$$\mu \circ \mathbf{S}\mu = \mu \circ \mu\mathbf{S}$$

in the category of functors with natural transformations as morphisms, see the diagram at the end of Section 3. Since $(\mathbf{S} \circ \mu)_X = \mathbf{S}(\mu_X)$ and $(\mu \circ \mathbf{S})_X = \mu_{\mathbf{S}(X)}$, this translates to

$$\mu_X \circ \mathbf{S}(\mu_X) = \mu_X \circ \mu_{\mathbf{S}(X)}.$$

Because the equation $\mu_X \circ \eta_X = id_{\mathbf{S}(X)}$ is easily established through a simple computation, the defining diagrams are commutative.
Since

$$\mu_X(\alpha_1 \cdot \tau_1 + \alpha_2 \cdot \tau_2) = \alpha_1 \cdot \mu_X(\tau_1) + \alpha_2 \cdot \mu_X(\tau_2),$$

the positive convex structure induced on $\mathbf{S}(X)$ by this algebra is the natural one. —

The finite case can easily be characterized: there are no algebras for $\{1, \ldots, n\}$ unless $n = 1$. This will be shown now. As a byproduct we obtain a simple geometric description as a necessary condition for the existence of algebras.
We need a wee bit elementary topology for this.

Definition 10 *A metric space A is called* connected *iff the decomposition $A = A_1 \cup A_2$ with disjoint open sets A_1, A_2 implies $A_1 = \emptyset$ or $A_2 = \emptyset$.*

Thus a connected space cannot be decomposed into two non-trivial open sets. The connected subspaces of the real line $\mathbb{R}$ are just the open, half-open or closed finite or infinite intervals. The rational numbers $\mathbb{Q}$ are not connected. A subset $\emptyset \neq A \subseteq \mathbb{N}$ of the natural numbers which carries the discrete topology (because we assume that it is a Polish space) is connected as a subspace iff $A = \{n\}$ for some $n \in \mathbb{N}$.
The following facts about connected spaces are well known, cp. [10, Chapter 6.1] (or any other standard reference to topology).

Lemma 5 *Let A be a metric space.*

1. *If A is connected, and $f : A \to B$ is a continuous and surjective map to another metric space B, then B is connected.*
2. *If two arbitrary points in A can be joined through a connected subspace of A, then A is connected.*

This has as a consequence

Corollary 3 *If $\langle X, h \rangle$ is an algebra for the Giry monad, then X is connected.*

Proof If $\tau_1, \tau_2 \in \mathbf{S}(X)$ are arbitrary probability measures on X, then the line segment $\{c \cdot \tau_1 + (1 - c) \cdot \tau_2 \mid 0 \leq c \leq 1\}$ is a connected subspace which joins τ_1 and τ_2. This is so because it is the image of the connected unit interval $[0, 1]$ under the continuous map $c \mapsto c \cdot \tau_1 + (1 - c) \cdot \tau_2$. Thus $\mathbf{S}(X)$ is connected by Lemma 5. Since h is onto, its image X is connected. □
Consequently it is hopeless to search for algebras for, say, the natural numbers or a non-trivial subset of it:

Corollary 4 *A subspace $A \subseteq \mathbb{N}$ has algebras for the Giry monad iff A is a singleton set.*

Proof It is clear that a singleton set has an algebra. Conversely, if A has an algebra, then A is connected by Lemma 3, and this can only be the case when A is a singleton. □
The next example deals with the unit interval:

Example 3 The map

$$h : \mathbf{S}([0, 1]) \ni \tau \mapsto \int_0^1 t \; \tau(dt) \in [0, 1]$$

defines an algebra $\langle [0, 1], h \rangle$. In fact, $h(\tau) \in [0, 1]$ because τ is a subprobability measure. It is clear that $h(\delta_x) = x$ holds, and — by the very definition of the weak topology — that $\tau \mapsto h(\tau)$ is continuous. Thus by Proposition 6 it remains to show that the partition induced by h is positive convex. This is a fairly simple calculation. Consequently, the partition induced by h is a G-partition, showing that h is indeed the morphism part of an algebra.
It is not difficult to see that the positive convex structure induced on $[0, 1]$ is the natural one.
—

The final example has a more geometric touch to it and deals only with the probabilistic case. We work with bounded and closed subsets of some Euclidean space and show that the construction of a barycenter yields an algebra. Fix $X \subseteq \mathbb{R}^n$ as a bounded, closed and convex subset of the Euclidean space $\mathbb{R}^n$ (for example, X could be a closed ball or a cube in $\mathbb{R}^n$). Denote for two vectors $x, x' \in \mathbb{R}^n$ by

$$x \star x' := \sum_{i=1}^{n} x_i \cdot x'_i$$

their inner product. Then $\lambda x.x \star x'$ constitutes a continuous linear map on $\mathbb{R}^n$ for fixed x'. In fact, each linear functional on $\mathbb{R}^n$ can be represented in this way.

Definition 11 *The vector $x^* \in \mathbb{R}^n$ is called a* barycenter *of the probability measure $\tau \in \mathbf{P}(X)$ iff*

$$x \star x^* = \int_X x \star y \ \tau(dy)$$

holds for each $x \in X$.

Because X is compact, the integrand is bounded on X, thus the integral is always finite. We collect some basic facts about barycenters and refer the reader to [12] for details.

Lemma 6 *The barycenter of $\tau \in \mathbf{P}(X)$ exists, it is uniquely determined, and it is an element of X.*

Proof Once we know that the barycenter exists, uniqueness follows from the well-known fact that the linear functionals on $\mathbb{R}^n$ separate points. Existence of the barycenter is established in [12, Theorem 461 E], its membership in X follows from [12, Theorem 461 H]. □
These preparations help in establishing that the barycenter constitutes an algebra:

Proposition 13 *Let $b(\tau)$ be the barycenter of $\tau \in \mathbf{P}(X)$. Then $\langle X, b\rangle$ is an algebra for the Giry monad.*

Proof 1. $b : \mathbf{P}(X) \to X$ is well defined by Lemma 6. From the uniqueness of the barycenter it is clear that $b(\delta_x) = x$ holds for each $x \in X$.
2. Assume that $(\tau_n)_{n\in\mathbb{N}}$ is a sequence in $\mathbf{P}(X)$ with $\tau_n \rightharpoonup_w \tau_0$. Put $x_n^* := b(\tau_n)$ as the barycenter of τ_n, then $(x_n^*)_{n\in\mathbb{N}}$ is a sequence in the compact set X, thus has a convergent subsequence (which we take w.l.g. as the sequence itself). Let x_0^* be its limit. Then we have for all $x \in X$:

$$x \star x_n^* = \int_X x \star y \ \tau_n(dy) \to \int_X x \star y \ \tau_0(dy) = x \star x_0^*$$

Hence b is continuous.
3. It remains to show that the partition induced by b is convex. This, however, follows immediately from the linearity of $y \mapsto \lambda x.x \star y$. □
Calculating the convex structure for b, we infer from affinity of the integral as a function of the measure and from

$$x \star b(\tau) = \int_X x \star y \ \tau(dy)$$

that $(0 \leq c \leq 1, \tau_i \in \mathbf{P}(X))$

$$b(c \cdot \tau_1 + (1-c) \cdot \tau_2) = c \cdot b(\tau_1) + (1-c) \cdot b(\tau_2)$$

that the convex structure induced by b is the natural one.
It should be mentioned that this example can be generalized considerably to metrizable topological vector spaces. The terminological effort is, however, somewhat heavy, and the example remains essentially the same. Thus we refrain from a more general discussion.
Although the characterization of algebras in terms of positive convex structures yields a somewhat uniform approach, it becomes clear from these examples that the specific instances of the algebras provide a rather colorful picture unified only through the common abstract treatment.

6 Returning to Ludwig

Returning to the language of Section 2, we interpret these results in terms of the algebras involved. Let $\mathcal{Q}$ be a positive convex structure on X, hence $\mathcal{Q}$ yields an algebra $h_{\mathcal{Q}}$ on X according to Proposition 12. A stochastic relation $g : X \rightsquigarrow X$ is interpreted in $\mathcal{Q}$ through $g^{\mathcal{Q}} := h_{\mathcal{Q}} \circ g$.
The derandomized semantics $\mathcal{C}_{\mathcal{Q}}[\![P]\!]$ of a program P given a positive convex structure $\mathcal{Q}$ on X is $h_{\mathcal{Q}} \circ \mathcal{C}[\![P]\!]$ (with associated algebra $h_{\mathcal{Q}}$ according to Proposition 12). Following Section 2.3, we associate with P's work a stochastic relation β_P, see Remark 2. If stochastic relation g represents the initial state, the derandomized final state is given through

$$\mathcal{C}_{\mathcal{Q}}[\![P]\!]g = g^{\mathcal{Q}} \circ \beta_P,$$

with the Kleisli product as the composition.
Consistency of logic and semantics translates into consistency for derandomized semantics; we see also that morphisms for strongly convex structures translate into derandomized interpretations.

Proposition 14 *Given a program P, the states $f, g : X \rightsquigarrow X$ with $P, g \vdash f$ and a positive convex structure $\mathcal{Q}$ on X. Then*

1. $\mathcal{C}_{\mathcal{Q}}[\![P]\!]f = g^{\mathcal{Q}}$,
2. *if $\mathcal{R}$ is a positive convex structure on X and $\theta : \mathcal{Q} \rightarrow \mathcal{R}$ a morphism, then*

$$\mathcal{C}_{\mathcal{R}}[\![P]\!]\left(\mathbf{S}(\theta) \circ g\right) = \theta \circ g^{\mathcal{Q}}.$$

Proof Immediate from consistency of the semantics, Proposition 4, and from the isomorphism between $\mathfrak{Alg}$ and $\mathfrak{StrConv}$, Proposition 12. □
Establishing the counterpart to completeness is open; given the continuous derandomizations $f_1, g_1 : X \rightarrow X$ with

$$\mathcal{C}_{\mathcal{Q}}[\![P]\!]f_1 = g_1$$

it would require finding stochastic relations $f, g : X \rightsquigarrow X$ with $f_1 = f^{\mathcal{Q}}, g_1 = g^{\mathcal{Q}}$ and $\mathcal{C}[\![P]\!]f = g$. This looks like a difficult selection problem (remember that f and g should be continuous), and conditions permitting such a selection would have to be identified.

7 Related Work

Derandomization is a well known technique in algorithm design to gain a less randomized algorithm from a randomized one by reducing the number of random bits. This term was borrowed from this field for describing what happens when one has a probabilistic semantics and wants to see what is happening on the deterministic level. Consequently, these techniques are related in spirit.

The monad on which the present investigation is based was originally proposed and investigated by M. Giry [13] in an approach to provide a categorical foundation of Probability Theory. The functor on which it is based assigns each measurable space all probabilities defined on its σ-algebra, it is somewhat similar to the functor assigning each set its power set on which the monad investigated by Manes is based. While the Kleisli construction for the latter one leads to relations based on sets, it leads for the former one to stochastic relations as a similar relational construction. This point of view was emphasized first by P. Panagaden in [21] when pointing out similarities between set based and probability based relations. It was extended further in [1]. In [7] this aspect is elaborated in depth by showing how a software architecture can be modeled using a monad as the basic computational model; the monad is shown to subsume both the Manes and the Giry monad as special cases. Stochastic relations turned out to be a fruitful field for investigations [1, 8, 9] in particular in such areas as labeled Markov transition systems and modeling stochastically algebraic aspects of modal logic. A development towards the semantics of probabilistic programs in terms of the Eilenberg-Moore algebras on probabilistic powerdomains is presented in [14]: Jones shows among others how programs can be understood as state transformers using upper and lower continuous functions with evaluations (for which an integration theory is developed); the underlying computational model is Moggi's λ_c calculus, as in the present paper. An early proposal for using probabilities for modelling in automata theory can be found as an illustrating example in [2, Nos. (6), (13)]. The situation discussed is the set of all probabilities with finite support, and it is mentioned without going into details that algebras, which are called *deciders* here, are generalized convex sets.

Using programs as transformers has been used for quite some time when the objects to be transformed may be cast as predicates in some logic, e.g. for Hoare triplets. The use of probabilities is not quite as common. Kozen uses in [15] a predicate transformer technique in his analysis of probabilistic programs; he demonstrated essentially how the randomizing constructs give rise to a transform on a Banach space of continuous functions. For deterministic programs, the present author showed e.g. in [6, 5] how the notion of programs as measure transforms can be used for the average case analysis of programs (see Example 1). Monniaux's work, e.g. [19], uses probabilistic methods for the abstract interpretation of programs. He proposes among others an adjoint semantics of nondeterministic, probabilistic programs, and relates this to abstract interpretations of non-probabilistic programs. The methods employed come specifically from the duality theory of linear operators on integrable functions.

The investigation of Eilenberg-Moore algebras for the probability functor using convexity arguments has been advocated e.g. in [23, 24, 20], pioneering work having been reported in [25, 27], see [11]. It is clearly intimately connected with the question of identifying the adjoints for this functor, which are not yet completely known in the category of Polish spaces with continuous or with measurable maps as morphisms.

8 Further Work

We characterize the algebras for the Giry monad which assigns each Polish space its space of probabilities. The morphisms in this category are continuous maps between Polish spaces. Continuity is technically crucial for the argumentation. This approach will not work for general Borel measurable maps serving as morphisms between Polish spaces (although these maps are fairly interesting from the point of view of applications), thus a more general characterization for these algebras is desirable.
Continuity plays also a crucial role in some of the examples that are discussed. Through the geometric argument of connectedness we could show that for the discrete case no algebras exist, except in the very trivial case of a one point space. This argument also does work only when the morphisms involved are continuous. So it is desirable to find algebras for the general case of Borel maps over finite domains (probably they do not exist usually there either: one would also like to know that). The last example hints at a connection between these algebras and barycenters for compact convex sets in topological vector spaces. It ends here where the fun begins there, viz., when looking at Choquet's theory of integral representations. There is room for further work exploring this avenue. The examples show that the world of algebras for this monad is quite colorfully polymorphous.
The most interesting question, however, addresses the expansion of the characterization given here for Borel measurable maps which are based on Polish spaces, or, going one crucial step further, on analytic ones. This goes hand in hand with the request for identifying adjoints for the probability functor (or its close cousin, the subprobability functor).
We have shown how derandomization can be carried out for a simple programming language in the scenario of Polish spaces with continuous maps. It would be interesting to see how this can be extended to questions pertaining to programming in the large. From a software architectural point of view, the approach that was shown to work for pipelines [7] should be widened to other architectures [16].

References

[1] S. Abramsky, R. Blute, and P. Panangaden. Nuclear and trace ideal in tensored *-categories. *Journal of Pure and Applied Algebra*, 143(1 – 3):3 – 47, 1999.

[2] M. A. Arbib and E. G. Manes. Fuzzy morphisms in automata theory. In E. G. Manes, editor, *Category Theory Applied to Computation and Control, Proc. First Int. Symp, San Francisco*, number 25 in Lecture Notes in Computer Science, pages 80 – 86, Berlin, 1975. Springer-Verlag.

[3] P. Billingsley. *Probability and Measure.* Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, New York, third edition, 1995.

[4] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation of labelled Markov-processes. *Information and Computation*, 179(2):163 – 193, 2002.

[5] E.-E. Doberkat. Continuous models that are equivalent to randomness for the analysis of many sorting algorithms. *Computing*, 31:11 – 31, 1983.

[6] E.-E. Doberkat. An analysis of Floyd's algorithm for heapconstruction. *Information and Control*, 61:114 – 131, 1984.

[7] E.-E. Doberkat. Pipelines: Modelling a software architecture through relations. *Acta Informatica*, 40:37 – 79, 2003.

[8] E.-E. Doberkat. Semi-pullbacks and bisimulations in categories of stochastic relations. In *Proc. ICALP'03*, volume 2719 of *Lecture Notes in Computer Science*, pages 996 – 1007, Berlin, 2003. Springer-Verlag.

[9] E.-E. Doberkat. Factoring stochastic relations. *Information Processing Letters*, 90(4):161 – 166, May 2004.

[10] R. Engelking. *General Topology*, volume 6 of *Sigma Series in Pure Mathematics*. Heldermann-Verlag, Berlin, revised and completed edition edition, 1989.

[11] V. V. Fedorchuk. Probability measures in topology. *Russian Math. Surveys*, pages 45 – 93, 1991. (Uspekhi Mat. Nauk 46:1 (1991), 41 – 80).

[12] D. H. Fremlin. *Measure Theory*, volume 4 (Topological Measure Spaces). Torres Fremlin, Colchester, November 2003.

[13] M. Giry. A categorical approach to probability theory. In *Categorical Aspects of Topology and Analysis*, volume 915 of *Lecture Notes in Mathematics*, pages 68 – 85, Berlin, 1981. Springer-Verlag.

[14] C. Jones. *Probabilistic Non-determinism*. PhD thesis, University of Edinburgh, 1989.

[15] D. E. Kozen. Semantics of probabilistic programs. *Journal of Computer and Systems Sciences*, 30(2):328 – 350, 1985.

[16] G. Lajios. *Zur kategoriellen Beschreibung von Schichtenarchitekturen*. PhD thesis, Chair for Software Technology, University of Dortmund, 2005 (in preparation).

[17] S. Mac Lane. *Categories for the Working Mathematician*. Number 5 in Graduate Texts in Mathematics. Springer-Verlag, Berlin, 2 edition, 1997.

[18] E. Moggi. Notions of computation and monads. *Information and Computation*, 93:55 – 92, 1991.

[19] D. Monniaux. Backwards abstract interpretation of probabilistic programs. In *Proc. ESOP'01*, number 2028 in Lecture Notes in Computer Science, pages 367 – 382. Springer-Verlag, 2001.

[20] O. R. Nykyforchyn. Semiconvex compacta. *Comment. Math. Univ. Carolinae*, 38(3):761 – 774, 1997.

[21] P. Panangaden. Probabilistic relations. In C. Baier, M. Huth, M. Kwiatkowska, and M. Ryan, editors, *Proc. PROBMIV*, pages 59 – 74, 1998. Also available from the School of Computer Science, McGill University, Montreal.

[22] K. R. Parthasarathy. *Probability Measures on Metric Spaces*. Academic Press, New York, 1967.

[23] D. Pumplün. Eilenberg-Moore algebras revisited. Technical Report 29, Fachbereich Mathematik und Informatik, FernUniversität Hagen, 1988.

[24] D. Pumplün. Positively convex modules and ordered normed linear spaces. *J. Convex Analysis*, 10(1):109 – 127, 2003.

[25] Z. Semadeni. Monads and their Eilenberg-Moore algebras in functional analysis. *Queen's Papers Pure Appl. Math.*, 33, 1973.

[26] S. M. Srivastava. *A Course on Borel Sets*. Number 180 in Graduate Texts in Mathematics. Springer-Verlag, Berlin, 1998.

[27] T. Swirszcz. Monadic functors and convexity. *Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys.*, 22:39 – 42, 1974.

[28] L. Wittgenstein. *Philosophische Untersuchungen*. Suhrkamp Verlag, Frankfurt am Main, 1967. Wissenschaftliche Sonderausgabe.

/120/ Volker Gruhn, Ursula Wellen
Autonomies in a Software Process Landscape
Januar 2002

/121/ Ernst-Erich Doberkat, Gregor Engels (Hrsg.)
Ergebnisbericht des Jahres 2001
des Projektes “MuSofT – Multimedia in der SoftwareTechnik”
Februrar 2002

/122/ Ernst-Erich Doberkat, Gregor Engels, Jan Hendrik Hausmann, Mark Lohmann, Christof Veltmann
Anforderungen an eine eLearning-Plattform — Innovation und Integration —
April 2002

/123/ Ernst-Erich Doberkat
Pipes and Filters: Modelling a Software Architecture Through Relations
Juni 2002

/124/ Volker Gruhn, Lothar Schöpe
Integration von Legacy-Systemen mit Eletronic Commerce Anwendungen
Juni 2002

/125/ Ernst-Erich Doberkat
A Remark on A. Edalat’s Paper *Semi-Pullbacks and Bisimulations in Categories of Markov-Processes*
Juli 2002

/126/ Alexander Fronk
Towards the algebraic analysis of hyperlink structures
August 2002

/127/ Markus Alvermann, Martin Ernst, Tamara Flatt, Urs Helmig, Thorsten Langer
Ingo Röpling, Clemens Schäfer, Nikolai Schreier, Olga Shtern
Ursula Wellen, Dirk Peters, Volker Gruhn
Project Group Chairware Final Report
August 2002

/128/ Timo Albert, Zahir Amiri, Dino Hasanbegovic, Narcisse Kemogne Kamdem,
Christian Kotthoff, Dennis Müller, Matthias Niggemeier, Andre Pavlenko, Stefan Pinschke,
Alireza Salemi, Bastian Schlich, Alexander Schmitz
Volker Gruhn, Lothar Schöpe, Ursula Wellen
Zwischenbericht der Projektgruppe Com42Bill (PG 411)
September 2002

/129/ Alexander Fronk
An Approach to Algebraic Semantics of Object-Oriented Languages
Oktober 2002

/130/ Ernst-Erich Doberkat
Semi-Pullbacks and Bisimulations in Categories of Stochastic Relations
November 2002

/131 Yalda Ariana, Oliver Effner, Marcel Gleis, Martin Krzysiak,
Jens Lauert, Thomas Louis, Carsten Röttgers, Kai Schwaighofer,
Martin Testrot, Uwe Ulrich, Xingguang Yuan
Prof. Dr. Volker Gruhn, Sami Beydeda
Endbericht der PG nightshift:
Dokumentation der verteilten Geschäftsprozesse im FBI und Umsetzung von Teilen dieser Prozesse im Rahmen eines FBI-Intranets basierend auf WAP- und Java-Technologie
Februar 2003

/132/ Ernst-Erich Doberkat, Eugenio G. Omodeo
ER Modelling from First Relational Principles
Februar 2003

/133/ Klaus Alfert, Ernst-Erich Doberkat, Gregor Engels (Hrsg.)
Ergebnisbericht des Jahres 2002 des Projektes “MuSofT – Multimedia in der SoftwareTechnik”
März 2003

/134/ Ernst-Erich Doberkat
Tracing Relations Probabilistically
März 2003

/135/ Timo Albert, Zahir Amiri, Dino Hasanbegovic, Narcisse Kemogne Kamdem,
Christian Kotthoff, Dennis Müller, Matthias Niggemeier,
Andre Pavlenko, Alireza Salemi,Bastian Schlich, Alexander Schmitz,
Volker Gruhn, Lothar Schöpe, Ursula Wellen
Endbericht der Projektgruppe Com42Bill (PG 411)
März 2003

/136/ Klaus Alfert
Vitruv: Specifying Temporal Aspects of Multimedia Presentations —
A Transformational Approach based on Intervals
April 2003

/137/ Klaus Alfert, Jörg Pleumann, Jens Schröder
A Framework for Lightweight Object-Oriented Design Tools
April 2003

/138/ K. Alfert, A. Fronk , Ch. Veltmann (Hrsg.)
Stefan Borggraefe, Leonore Brinker, Evgenij Golkov, Rafael Hosenberg, Bastian Krol, Daniel Mölle,
Markus Niehammer, Ulf Schellbach, Oliver Szymanski, Tobias Wolf, Yue Zhang
Endbericht der Projektgruppe 415: Konzeption und Implementierung eines digitalen und hypermedialen Automobilcockpits (HyCop)
Mai 2003

/139/ Volker Gruhn, Malte Hülder, Sami Beydeda (Hrsg.)
Endbericht der Projektgruppe 409: Entwicklung von ortsbasierten Diensten für UMTS-Mobilfunkgeräte (mCube)
Mai 2003

/140/ Ernst-Erich Doberkat
Congruences for Stochastic Relations
Juli 2003

/141/ Marion Kamphans, Sigrid Metz-Göckel, Anja Tigges
Wie Geschlechteraspekte in die digitalen Medien integriert werden können – das BMBF-Projekt „MuSofT“
September 2003

/142/ Ernst-Erich Doberkat
Semi-Pullbacks for Stochastic Relations over Analytic Spaces
Januar 2004

/143/ Volker Gruhn, Lothar Schöpe (Hrsg.)
1. Workshop des Verbundforschungsprojektes Mobile Spedition im Web – SpiW
Oktober 2002

/144/ Ernst-Erich Doberkat
Stochastic Relations Interpreting Modal Logic
Oktober 2003

/145/ Alexander Fronk, Ernst-Erich Doberkat, Johannes Bergemann, Ulrich-Walter Gans
Ein interdisziplinäres methodisches Vorgehen zur Gestaltung webbasierter Studieneinheiten für die Altertumswissenschaften
November 2003

/146/ Ernst-Erich Doberkat
Factoring Stochastic Relations
Januar 2004

/147/ Ernst-Erich Doberkat
Characterizing the Eilenberg-Moore Algebras for a Monad of Stochastic Relations
March 2004

/148/ Ernst-Erich Doberkat, Gregor Engels, Corina Kopka (Hrsg.) Abschlussbericht des Projektes “MuSofT – Multimedia in der SoftwareTechnik”
April 2004

/149/ Ernst-Erich Doberkat
Derandomizing probabilistic semantics through Eilenberg-Moore algebras for the Giry monad
July 2004