Probabilistic Domains

Reinhold Heckmann

FB 14 – Informatik, Prof. Wilhelm
Universität des Saarlandes, Postfach 151150
D-66041 Saarbrücken, Germany
e-mail: heckmann@cs.uni-sb.de

September 16, 1997

Abstract

We show the equivalence of several different axiomatizations of the notion of (abstract) *probabilistic domain* in the category of dcpo's and continuous functions. The axiomatization with the richest set of operations provides probabilistic selection among a finite number of possibilities with arbitrary probabilities, whereas the poorest one has binary choice with equal probabilities as the only operation. The remaining theories lie in between; one of them is the theory of binary choice by Graham [1].

1 Introduction

A probabilistic programming language could contain different kinds of language constructs to express probabilistic choice. In a rather poor language, there might be a construct $x \oplus y$, whose semantics is a choice between the two possibilities x and y with equal probabilities $1/2$. The 'possibilities' x and y can be statements in an imperative language or expressions in a functional language. A quite rich language could contain a construct $[p_1 : x_1, \ldots, p_n : x_n]$, where p_i are real numbers between 0 and 1 whose sum is 1. The semantics would be to select one of the possibilities x_i with probability p_i.

Graham [1], Jones [3], and Jones / Plotkin [4] consider an intermediate language with a construct $x \overset{p}{—} y$, which is written as $p \to x, y$ by Graham and $x +_p y$ by Jones, where p is a real number between 0 and 1. The semantics of this construct is to select x with probability p and y with probability $1 - p$.

The notion of an (abstract) *probabilistic domain* was introduced by Graham [1] and further elaborated by Jones [3, 4] to describe the denotational semantics of a probabilistic language with the construct $x \overset{p}{—} y$. A probabilistic domain is a dcpo together with a continuous operation satisfying several axioms, which is used to model the choice construct semantically. For reasons of simplicity, we denote the results of this semantic operation by $x \overset{p}{—} y$ as well, where x and y are no longer language constructs, but members of the underlying dcpo.

Jones and Plotkin define a *probabilistic power domain construction* which produces the free probabilistic domain over a given base domain. This construction is a strong monad in the sense of Moggi [5, 6]. Hence the denotational semantics of probabilistic languages with a construct $x \stackrel{p}{—} y$ can be written down schematically using this monad. Such a semantics can be found in [3].

In the paper at hand, we introduce some other notions of probabilistic domain with other semantic operations which describe a whole range of language constructs from the binary choice $x \oplus y$ over $x \stackrel{p}{—} y$ till $[p_1 : x_1, \ldots, p_n : x_n]$ with various intermediate steps. Every kind of probabilistic domain is described as a dcpo together with some continuous operations satisfying some axioms, i.e., as a model of some algebraic theory in the category $\mathcal{DCPO}$. Then, we prove that *all these theories are equivalent* in the sense that their categories of models are equivalent. This has several consequences. First, one sees that in the category $\mathcal{DCPO}$, all the probabilistic operations mentioned above are in fact equally expressive; even the multiple choice $[p_1 : x_1, \ldots, p_n : x_n]$ can be defined in terms of binary choice $x \oplus y$ (with the help of fixed point iteration). Next, the probabilistic power construction of [3, 4] not only produces the free probabilistic domain with $x \stackrel{p}{—} y$, but free probabilistic domains with all other operations as well because all the categories of probabilistic domains are equivalent. Thus, one power construction can be used to describe the semantics of a whole range of probabilistic languages.

Finding algebraic theories which are equivalent to the probabilistic theory of Graham and Jones also shows how some disadvantages of that theory can be avoided. Their theory refers to the unit interval $\mathbb{I} = [0..1]$ of the reals with the standard Hausdorff topology and thus goes beyond domain theory, and their axiom of associativity (in the version of Jones)

$$\text{if } pq \neq 1 \text{, then } \; (x \stackrel{p}{—} y) \stackrel{q}{—} z \; = \; x \stackrel{pq}{—} (y \stackrel{r}{—} z) \text{, where } r = \tfrac{q(1-p)}{1-pq}$$

is a conditional axiom and contains a complex fraction which reflects the recalculation of probabilities when moving the parentheses. The condition is needed to prevent the denominator of this fraction to become zero. The equivalent theories presented in our paper tend to avoid these problems; many of them refer to domains (dcpo's) only, or contain only simple unconditional axioms.

After introducing the mathematical background in Section 2, we present our richest theory, Multiple Choice with Divergence (MCD), in Section 3. Its basic operations are the choice among n possibilities $x_1, \ldots, x_n$ with probabilities $p_1, \ldots, p_n$, where $\sum_{i=1}^{n} p_i \leq 1$, for every finite $n \geq 0$. The difference $1 - \sum_{i=1}^{n} p_i$ is the probability to stay undecisive for ever, i.e., to diverge. The possibility of divergence would be odd in a real programming language, but we need it for the purpose of our equivalence proof.

The theory MCD has particularly neat axioms, which are powerful, simple, algorithmically intuitive, and also algebraically intuitive if the multiple choice is considered as a formal linear combination $\sum_{i=1}^{n} p_i \cdot x_i$. In her thesis [3], Jones derived most parts of MCD from her theory and used this to prove that her probabilistic power domains are free probabilistic domains over the argument domain. However, Jones did not present a complete description of MCD, nor did she work the other way round and prove the equivalence of MCD and her theory.

In Section 4, we restrict multiple choice to binary choice where the sum of the two probabilities is 1. The resulting operation is the original operation $x \stackrel{p}{—} y$ of Graham and Jones. We present four versions of the theory of Binary Choice: Theory BC-A (Binary Choice with Associativity) is the original theory in the version of Jones with the complex conditional associativity axiom. In theory BC-R, associativity is replaced by two simpler unconditional axioms, the Rectangle axiom and the axiom of linear combination. Theory BC-P has yet another set of axioms which contains a Product axiom. Theory BC-L has a large axiom set which contains all the three other axiom sets. Because of our main result that all these theories have the same models, a dcpo with binary choice satisfies all the properties listed in theory BC-L once it has been proved that it satisfies one of the three small axiom sets BC-A, BC-R, or BC-P.

In Section 5, we restrict the binary choice between x with probability p and y with probability $1-p$ to two special cases: a binary choice '$\oplus$' between x and y with equal probabilities 1/2, and a choice between x with probability p and $\perp$ with probability $1-p$, which can be seen as multiplication of x by a scalar p drawn from the unit interval (with the Scott topology) and consequently is written as $p \cdot x$. These operations are similar to the operations in a vector space or module, where '$\oplus$' plays the role of addition. Thus, we call the theories with these operations IM for $\mathbb{I}$-module. We present two versions of this theory: a theory IM-S with a small axiom set, and IM-L with a large axiom set. The axioms of both theories are simple equality statements. Because of our results, it suffices to verify the axioms of IM-S for a given $\mathbb{I}$-module structure, and then all the properties listed in IM-L hold as well.

In Section 6, we even drop multiplication and obtain a theory with binary choice with equal probabilities as the only operation. We call this theory MV (Mean Value algebra) since this kind of choice has the algebraic properties of mean value formation. The theory MV has the advantage of not mentioning real numbers at all, at the expense of a complex axiom, which states that the least fixed point of the function $\lambda x.\, a \oplus x$ is a. Computationally, this axiom reflects the fact that in a recursive program $x = a \oplus x$, the possibility a is chosen with probability $1/2 + 1/4 + 1/8 + \cdots = 1$, whence the program is equivalent to a. The other theories do not need this axiom explicitly since it follows from the structure of the unit interval.

The theory MV was already presented by the author in [2] as an example of a 'power theory', i.e., a theory with a binary operation modeling the binary choice operator of non-deterministic programming languages. In that paper, we did not mention any other probabilistic theory, but only conjectured the equivalence with BC-A, the theory of Jones.

In the course of presenting these more and more restricted theories, we also derive every theory (except for MCD, of course) from a less restricted theory presented earlier. This is done in the following order:

$$\begin{tikzcd}
\mathrm{MCD} \arrow[rr] & & \mathrm{BC\text{-}L} \arrow[rr] \arrow[dl] \arrow[dr] & & \mathrm{IM\text{-}L} \arrow[dd] & \\
& \mathrm{BC\text{-}A} \arrow[dr] & & \mathrm{BC\text{-}R} \arrow[dl] & & \\
& & \mathrm{BC\text{-}P} \arrow[rr] & & \mathrm{IM\text{-}S} \arrow[r] & \mathrm{MV}
\end{tikzcd}$$

To show equivalence of all theories, we finally have to derive the most powerful theory MCD from the most restricted theory MV, i.e., we have to define multiple choice with arbitrary probabilities in terms of binary choice with equal probabilities 1/2 (and joins of ascending sequences). This is done in Section 7.

2 Mathematical Background

We use the standard definitions of posets, least upper bounds or joins (denoted by $\bigsqcup A$), and directed sets. A *dcpo* is a poset where every directed set has a least upper bound. In this paper, we often call dcpo's *domains*. A function $f : \mathbf{X} \to \mathbf{Y}$ between dcpo's $\mathbf{X}$ and $\mathbf{Y}$ is *continuous* if for all directed subsets D of $\mathbf{X}$, $f(\bigsqcup D) = \bigsqcup f(D)$ holds. The category of dcpo's and continuous functions is called $\mathcal{DCPO}$.

A subset O of a dcpo $\mathbf{X}$ is *Scott open* if it is upper ($x \in O$, $x \sqsubseteq y \Rightarrow y \in O$), and for all directed sets D, $\bigsqcup D \in O$ implies $D \cap O \neq \emptyset$. The Scott open sets of a dcpo $\mathbf{X}$ form a topology on $\mathbf{X}$, the Scott topology. A function $f : \mathbf{X} \to \mathbf{Y}$ between dcpo's is continuous in the sense of the previous paragraph iff $f^{-1}V$ is Scott open in $\mathbf{X}$ for every Scott open set V of $\mathbf{Y}$. Thus, $\mathcal{DCPO}$ can be considered as a full subcategory of the category $\mathcal{TOP}$ of topological spaces, and it makes sense to speak of the continuity of a function from a space to a dcpo, as it will be done in the description of the BC theories.

The *unit interval* $\mathbb{I}$ is the set of real numbers r with $0 \leq r \leq 1$. We can either consider it as a dcpo ordered by '$\leq$' and equip it with the Scott topology, or topologize it by the standard Hausdorff topology. The two possibilities are connected by the following lemma:

Lemma 2.1 A function $f : \mathbb{I} \to Y$ from the unit interval to a dcpo Y is continuous w.r.t. the Scott topology on $\mathbb{I}$ if and only if it is monotonic w.r.t. the standard order of $\mathbb{I}$ and continuous w.r.t. the Hausdorff topology of $\mathbb{I}$.

3 Multiple Choice with Divergence

In this section, we present the probabilistic theory with the most powerful operations: the theory MCD of multiple choice with divergence. The basic operation of MCD is choice between n possibilities $x_1, \ldots, x_n$ with probabilities $p_1, \ldots, p_n$, whose sum is at most 1.

In the following table as well as in all subsequent ones, we assume that the probabilistic domain to be described is a *dcpo* called X.

Operations: Continuous functions $\nabla_n \times X^n \to X$ for every $n \geq 0$, where

$$\nabla_n = \{(p_1, \ldots, p_n) \in \mathbb{I}^n \mid p_1 + \cdots + p_n \leq 1\}$$

are dcpo's, i.e., equipped with the Scott topology. The result of applying the operation of degree n to $(p_1, \ldots, p_n)$ in ∇_n and $(x_1, \ldots, x_n)$ in X^n is written $[p_1 : x_1, \ldots, p_n : x_n]$.

Homomorphisms are continuous functions h with

$$h\,[p_1 : x_1, \ldots, p_n : x_n] \;=\; [p_1 : h\,x_1, \ldots, p_n : h\,x_n].$$

Axioms: The axioms are those of linear combinations $\sum_{i=1}^{n} p_i \cdot x_i$, namely:

Permutation P: For every permutation π of $\{1, \ldots, n\}$,

$$[p_{\pi 1} : x_{\pi 1}, \ldots, p_{\pi n} : x_{\pi n}] \;=\; [p_1 : x_1, \ldots, p_n : x_n]$$

1 law: $[1 : x] \;=\; x$

0 law: $[0 : x,\, p_1 : y_1, \ldots, p_n : y_n] \;=\; [p_1 : y_1, \ldots, p_n : y_n]$

Addition $+$: $[p : x,\, q : x,\, r_1 : y_1, \ldots, r_n : y_n] \;=\; [p + q : x,\, r_1 : y_1, \ldots, r_n : y_n]$

Substitution S: $[p : [q_1 : x_1, \ldots, q_k : x_k],\, r_1 : y_1, \ldots, r_n : y_n]$

$$= \; [p\,q_1 : x_1, \ldots, p\,q_k : x_k,\, r_1 : y_1, \ldots, r_n : y_n]$$

For simplicity, we have omitted the universal quantification, which for instance for substitution should be: for all $k \geq 0$ and $n \geq 0$, $x_1, \ldots x_k$, $y_1, \ldots y_n$ in X, $(q_1, \ldots, q_k)$ in ∇_k, and $(p, r_1, \ldots, r_n)$ in ∇_{n+1}. Thus, the axioms are not conditional, although they can be applied only to real numbers satisfying certain conditions.

4 Binary Choice without Divergence

In this section, we present our four versions of the theory of binary choice without divergence BC, and derive them from MCD. The binary choice $x \overset{p}{—} y$ introduced here corresponds to $[p : x,\, 1-p : y]$ from the previous section. Whereas MCD refers to ∇_n with the Scott topology, the BC theories refer to $\mathbb{I}$ with the Hausdorff topology.

Operations: A constant $0 : X$ and an operation $\beta : \mathbb{I} \times X \times X \to X$, which is continuous in each argument separately, if $\mathbb{I}$ is equipped with the standard Hausdorff topology. We write $x \overset{p}{—} y$ instead of $\beta(p, x, y)$.

Homomorphisms are continuous functions h with
$h\,(0) = 0$ and $h\,(x \overset{p}{—} y) = hx \overset{p}{—} hy$.

Axioms, version L (large axiom set):

$\perp$ Least element: $0 \sqsubseteq x$ for all x in X.

C Commutativity: $x \overset{p}{—} y = y \overset{1-p}{—} x$.

0 Zero law: $x \overset{0}{—} y = y$.

1 One law: $x \overset{1}{—} y = x$.

I Idempotence: $x \overset{p}{—} x = x$.

D Distributivity: $x \overset{p}{—} (y \overset{q}{—} z) = (x \overset{p}{—} y) \overset{q}{—} (x \overset{p}{—} z)$.

P Product law: $(x \overset{p}{—} y) \overset{q}{—} y = x \overset{pq}{—} y$.

L Linear combination:
$(x \overset{p}{—} y) \overset{r}{—} (x \overset{q}{—} y) = x \overset{s}{—} y$ where $s = rp + (1-r)q$.

R Rectangle law: $(x \overset{q}{—} y) \overset{p}{—} (u \overset{q}{—} v) = (x \overset{p}{—} u) \overset{q}{—} (y \overset{p}{—} v)$.

A Associativity:
If $pq \neq 1$, then $(x \overset{p}{—} y) \overset{q}{—} z = x \overset{pq}{—} (y \overset{r}{—} z)$, where $r = \frac{q(1-p)}{1-pq}$.

Axioms, version A (with associativity):
Axioms $\perp$, C, 1, I, and A from the list above.

Axioms, version R (with rectangle law):
Axioms $\perp$, 0, 1, L, and R from the list above.

Axioms, version P (with product law):
Axioms $\perp$, 0, 1, I, P, $L\frac{1}{2}$, and $R\frac{1}{2}$, where $L\frac{1}{2}$ is the instance of L with $r = 1/2$, and $R\frac{1}{2}$ is the instance of R with $p = q = 1/2$.

The wording 'β is continuous in each argument separately' means that the functions $\lambda p.\, \beta(p, x, y) : \mathbb{I} \to X$ are continuous for every fixed x and y, and analogously for the second and third argument. Within $\mathcal{DCPO}$, such a separate continuity would be equivalent to the continuity of β itself; this is why we did not postulate it explicitly in case of MCD.

Jones [3] has a more complex continuity requirement for β which is more restrictive in general than ours, but equivalent for the important case of a continuous domain X. We relaxed the continuity requirement because with the stronger version, equivalence of BC-A to the other probabilistic theories cannot be proved.

The rectangle law is named because of the following rectangular scheme:

$$\begin{array}{ccc} x & \overset{p}{\text{---}} & y \\ {}_q| & & |_q \\ u & \overset{p}{\text{---}} & v \end{array}$$

On the left hand side of R, choice is first performed within the two rows, then the results of the rows are combined. On the right hand side, choice is first done in the two columns, then the results of the columns are combined.

We presented theory BC in four versions because the equivalence of the respective sets of axioms is by no means obvious. For instance, we have no idea how to prove associativity from the axioms of BC-R directly, without reconstructing MCD as we have done in the paper at hand.

Theory BC-L comprises all interesting equalities we know of, BC-A is the theory of Graham and Jones, BC-R is the nicest theory in our view since it consists of simple unconditional axioms only, and BC-P consists of just the properties we need to derive IM-S in the next section (which in turn consists of just the properties we need to derive MV).

Derivation of BC-L from MCD

The constant 0 of BC is case $n = 0$ of MCD, i.e., $0 = [\,]$; and $\beta(p, x, y) = x \overset{p}{\text{---}} y$ is defined by $[p : x,\, 1 - p : y]$. Continuity of β in the second and third argument is immediate, and continuity in the first argument holds since $\varphi : \mathbb{I}_H \to \nabla_2$ with $\varphi(p) = (p, 1-p)$ is continuous, where $\mathbb{I}_H$ is the unit interval with the Hausdorff topology, and ∇_2 has the Scott topology.

Next, we show that 0 is the least element: $0 = [\,] = [0 : x]$ holds by the 0 law of MCD. By continuity and hence monotonicity of the operation of MCD, $[0 : x] \sqsubseteq [1 : x]$ follows. By the 1 law, the latter term equals x.

The equational axioms of BC-L can be shown by the following strategy: translate both sides into MCD by $x \overset{p}{\text{---}} y = [p : x,\, 1 - p : y]$, then flatten them using substitution, delete entries $0 : x$ using the 0 axiom of MCD, combine multiple entries with the same dcpo element using the + axiom, and then compare both sides of the equality. For instance, the proof of L looks as follows:

We have to show $(x \overset{p}{\text{---}} y) \overset{r}{\text{---}} (x \overset{q}{\text{---}} y) = x \overset{s}{\text{---}} y$ where $s = rp + (1 - r)q$. Translating the left hand side into MCD yields:

$$(x \overset{p}{\text{---}} y) \overset{r}{\text{---}} (x \overset{q}{\text{---}} y) = [r : [p : x,\, 1 - p : y],\, 1 - r : [q : x,\, 1 - q : y]]$$

With substitution and addition, we obtain $[s : x,\, s' : y]$, where $s = rp + (1 - r)q$ as required, and $s' = r(1 - p) + (1 - r)(1 - q)$, which is $1 - s$ as required (check $s + s' = 1$).

Derivation of BC-P from BC-R

We have to deduce I and P from 0, 1, L, and R.

I: $x \overset{p}{\text{---}} x \overset{1}{=} (x \overset{1}{\text{---}} x) \overset{p}{\text{---}} (x \overset{1}{\text{---}} x) \overset{L}{=} x \overset{s}{\text{---}} x$

where $s = p \cdot 1 + (1 - p) \cdot 1 = 1$. By 1, $x \overset{s}{\text{---}} x = x$ follows.

P: $\quad (x \overset{p}{\text{---}} y) \overset{q}{\text{---}} y \overset{0}{=} (x \overset{p}{\text{---}} y) \overset{q}{\text{---}} (x \overset{0}{\text{---}} y) \overset{L}{=} x \overset{s}{\text{---}} y$
where $s = q \cdot p + (1-q) \cdot 0 = pq$.

Derivation of BC-P from BC-A

We have to deduce 0, P, $L\frac{1}{2}$, and $R\frac{1}{2}$ from C, 1, I, and A.

Axiom 0: $\quad x \overset{0}{\text{---}} y \overset{C}{=} y \overset{1}{\text{---}} x \overset{1}{=} y$

For P, we have to distinguish two cases: if $pq = 1$, then $p = q = 1$, and $(x \overset{1}{\text{---}} y) \overset{1}{\text{---}} y = x = x \overset{1}{\text{---}} y$ holds by the 1 law. Otherwise, axiom A can be applied:

$$(x \overset{p}{\text{---}} y) \overset{q}{\text{---}} y \overset{A}{=} x \overset{pq}{\text{---}} (y \overset{r}{\text{---}} y) \overset{I}{=} x \overset{pq}{\text{---}} y$$

For $L\frac{1}{2}$, we have to show $(x \overset{p}{\text{---}} y) \overset{1/2}{\text{---}} (x \overset{q}{\text{---}} y) = x \overset{(p+q)/2}{\text{---}} y$. Applying A to the left hand side, we obtain $x \overset{p/2}{\text{---}} (y \overset{r}{\text{---}} (x \overset{q}{\text{---}} y))$ with $r = \frac{(1-p)/2}{1-p/2} = \frac{1-p}{2-p}$. Applying C yields $x \overset{p/2}{\text{---}} ((x \overset{q}{\text{---}} y) \overset{r'}{\text{---}} y)$ with $r' = 1 - r = \frac{1}{2-p}$. Applying P (which is already proved) yields $x \overset{p/2}{\text{---}} (x \overset{qr'}{\text{---}} y)$. Now, we apply C twice, then P, and finally C again, which gives $x \overset{s}{\text{---}} y$ with $s = 1 - (1 - \frac{q}{2-p})(1 - \frac{p}{2}) = 1 - \frac{2-p-q}{2-p} \cdot \frac{2-p}{2} = (p+q)/2$.

For $R\frac{1}{2}$, we have to show $(x \overset{1/2}{\text{---}} y) \overset{1/2}{\text{---}} (u \overset{1/2}{\text{---}} v) = (x \overset{1/2}{\text{---}} u) \overset{1/2}{\text{---}} (y \overset{1/2}{\text{---}} v)$. This is done by the following chain of equations:

$$(x \overset{1/2}{\text{---}} y) \overset{1/2}{\text{---}} (u \overset{1/2}{\text{---}} v) \overset{A}{=} x \overset{1/4}{\text{---}} (y \overset{1/3}{\text{---}} (u \overset{1/2}{\text{---}} v)) \overset{C}{=} x \overset{1/4}{\text{---}} ((u \overset{1/2}{\text{---}} v) \overset{2/3}{\text{---}} y)$$
$$\overset{A}{=} x \overset{1/4}{\text{---}} (u \overset{1/3}{\text{---}} (v \overset{1/2}{\text{---}} y)) \overset{A}{=} (x \overset{1/2}{\text{---}} u) \overset{1/2}{\text{---}} (v \overset{1/2}{\text{---}} y) \overset{C}{=} (x \overset{1/2}{\text{---}} u) \overset{1/2}{\text{---}} (y \overset{1/2}{\text{---}} v)$$

5 $\mathbb{I}$-Modules

In this section, we introduce the theory IM of $\mathbb{I}$-modules. Its 'addition' is binary choice with equal probabilities 1/2, and its 'multiplication' is choice between a point and 0. The axioms are very much like those of a module, hence the name.

Theory IM comes up in two versions: IM-L has a large set of axioms which includes all useful properties we know of, and IM-S has a small subset thereof, just enough to derive theory MV in the next section.

Again, we do not know how all axioms of IM-L can be proved from those of IM-S directly, without reconstructing MCD as done in this paper.

Operations: A constant $0 : X$, a continuous operation $\oplus : X \times X \to X$, and a continuous operation $\cdot : \mathbb{I} \times X \to X$, where $\mathbb{I}$ is considered as a dcpo (Scott topology).

Homomorphisms are continuous functions h with
$h(0) = 0$, $h(x \oplus y) = hx \oplus hy$, and $h(p \cdot x) = p \cdot hx$.

Axioms, version L (large axiom set):

$\perp$ Least element: $0 \sqsubseteq x$ for all x in X.

C Commutativity: $x \oplus y = y \oplus x$.

I Idempotence: $x \oplus x = x$.

R Rectangle law: $(x \oplus y) \oplus (u \oplus v) = (x \oplus u) \oplus (y \oplus v)$.

$R0$ Right zero: $p \cdot 0 = 0$.

RD Right distributivity: $p \cdot (x \oplus y) = p \cdot x \oplus p \cdot y$.

$L0$ Left zero: $0 \cdot x = 0$.

LD Left distributivity: $(p \oplus q) \cdot x = p \cdot x \oplus q \cdot x$, where $p \oplus q = (p+q)/2$ in $\mathbb{I}$.

1 One law: $1 \cdot x = x$.

PA Product associativity: $p \cdot (q \cdot x) = (p \cdot q) \cdot x$.

Axioms, version S (small axiom set):

Axioms R, $L0$, LD, and 1 from the list above.

Surprisingly, idempotence and commutativity of '$\oplus$' do not show up as axioms of IM-S. Nevertheless, they hold in every IM-S algebra by the main result of our paper.

Derivation of IM-S from BC-P

We define: $x \oplus y \ = \ x \xrightarrow{1/2} y$, and $p \cdot x \ = \ x \xrightarrow{p} 0$.

The operation $\oplus = \lambda(x, y).\, \beta(1/2, x, y) : X \times X \to X$ is continuous in its two arguments separately, since β is continuous in its second and third argument separately. Within $\mathcal{DCPO}$, this separate continuity is equivalent to the continuity of '$\oplus$' itself.

The function $\cdot = \lambda(p, x).\, \beta(p, x, 0) : \mathbb{I} \times X \to X$ is continuous in its second argument since β is. For continuity in the first argument, we apply Lemma 2.1, i.e., we have to show that for fixed x, the function $\lambda p.\, p \cdot x$ is monotonic.

Let $p \leq q$ in $\mathbb{I}$. We have to show $x \xrightarrow{p} 0 \sqsubseteq x \xrightarrow{q} 0$. Because of $p \leq q$, there is r in $\mathbb{I}$ such that $p = r \cdot q$. Applying the product law P of BC-P, we obtain $x \xrightarrow{p} 0 = (x \xrightarrow{r} 0) \xrightarrow{q} 0$. From $0 \sqsubseteq x$, $x \xrightarrow{r} 0 \sqsubseteq x \xrightarrow{r} x \overset{I}{=} x$ follows, whence $x \xrightarrow{p} 0 \sqsubseteq x \xrightarrow{q} 0$.

R: The rectangle law of IM is $R\frac{1}{2}$ of BC-P.

$L0$: $0 \cdot x \ = \ x \xrightarrow{0} 0 \ \overset{0}{=} \ 0$.

LD: Left distributivity of IM is the instance of $L\frac{1}{2}$ of BC-P with $y = 0$.

1: $1 \cdot x \ = \ x \xrightarrow{1} 0 \ \overset{1}{=} \ x$.

Derivation of IM-L from BC-L

When expressed in the language of BC, all the axioms of IM-L but RD become instances of axioms of BC-L. Right distributivity becomes

$$(x \xrightarrow{1/2} y) \xrightarrow{p} 0 \ = \ (x \xrightarrow{p} 0) \xrightarrow{1/2} (y \xrightarrow{p} 0),$$

with follows from distributivity D of BC-L with commutativity.

6 Mean Values

The probabilistic theory with the weakest operations is MV, the theory of mean values. It results from IM by dropping multiplication. Thus, MV does not mention real numbers explicitly.

Operations: A constant $0 : X$, and a continuous operation $\oplus : X \times X \to X$.

Homomorphisms are continuous functions h with

$h(0) = 0$ and $h(x \oplus y) = hx \oplus hy$.

Axioms: Commutativity: $x \oplus y = y \oplus x$,

Rectangle law: $(x \oplus y) \oplus (u \oplus v) = (x \oplus u) \oplus (y \oplus v)$, Least element: $0 \sqsubseteq x$,

Fixed point axiom: The least fixed point of $\lambda x.\, a \oplus x$ is a.

For the fixed point axiom, remember that the carrier X is a dcpo. The fixed point axiom may alternatively be written as idempotence $a \oplus a = a$, which means that a is a fixed point of $\lambda x.\, a \oplus x$, plus the conditional axiom $a \oplus b = b \Rightarrow a \sqsubseteq b$, which means that a is the least fixed point. Yet another formulation follows from making the fixed point iteration explicit: if $a_0 = 0$ and $a_{n+1} = a \oplus a_n$, then $\bigsqcup_{n \geq 0} a_n = a$. From this statement, a slightly stronger one can be easily deduced: if $b_0 \sqsubseteq a$ and $b_{n+1} = a \oplus b_n$, then $\bigsqcup_{n \geq 0} b_n = a$. It is this last version which is needed for the derivation of MCD from MV at the end of this paper.

Now let us derive the axioms of MV from those of IM-S. The rectangle law is immediate. The least element property of 0 holds since $0 = 0 \cdot x \sqsubseteq 1 \cdot x = x$. For the fixed point axiom, we have to consider the sequence defined by $a_0 = 0$ and $a_{n+1} = a \oplus a_n$. We claim $a_n = (1 - 2^{-n}) \cdot a$ for all n; the equation $\bigsqcup_{n \geq 0} a_n = a$ then follows from continuity of multiplication and the property $1 \cdot a = a$. The equality $a_0 = (1 - 2^{-0}) \cdot a$ holds since $0 \cdot a = 0$. For the inductive step, we have $a_{n+1} = a \oplus a_n = 1 \cdot a \oplus (1 - 2^{-n}) \cdot a = (1 + 1 - 2^{-n})/2 \cdot a = (1 - 2^{-(n+1)}) \cdot a$.

For commutativity, fix two members a and b of X. We start with some auxiliary statements. In the proof of the third, we may use idempotence because it follows from the fixed point axiom, which is already validated. (It could also be shown directly by $x \oplus x = 1 \cdot x \oplus 1 \cdot x = 1 \cdot x = x$.)

(1) $a \oplus p \cdot a = p \cdot a \oplus a$

Proof: $a \oplus p \cdot a = 1 \cdot a \oplus p \cdot a = (p + 1)/2 \cdot a$, and same for $p \cdot a \oplus a$.

(2) $a \oplus 0 = 0 \oplus a$

Proof: From (1) with $p = 0$.

(3) $a \oplus c = c \oplus a \;\Rightarrow\; a \oplus (b \oplus c) = (c \oplus b) \oplus a$

Proof: $a \oplus (b \oplus c) \stackrel{I}{=} (a \oplus a) \oplus (b \oplus c) \stackrel{R}{=} (a \oplus b) \oplus (a \oplus c) = (a \oplus b) \oplus (c \oplus a) \stackrel{R}{=} (a \oplus c) \oplus (b \oplus a) = (c \oplus a) \oplus (b \oplus a) \stackrel{R}{=} (c \oplus b) \oplus (a \oplus a) \stackrel{I}{=} (c \oplus b) \oplus a$.

Now we define $c_0 = 0$ and $c_{n+1} = b \oplus c_n$. From the proof of the fixed point axiom, we know $c_n = (1 - 2^{-n}) \cdot b$. By (1), $b \oplus c_n = c_n \oplus b$ follows. Thus, $c_{n+1} = c_n \oplus b$ also holds. By induction, we can show $a \oplus c_n = c_n \oplus a$ for all n; the start holds by (2), and the inductive step follows from (3). We already know $\bigsqcup_{n \geq 0} c_n = b$. Continuity of '$\oplus$' yields $a \oplus b = b \oplus a$ as required.

7 The Big Step: From MV to MCD

In the previous sections, we started from the theory MCD, which has a rich set of operations, and restricted its operations until only binary choice with equal probabilities was left. In this section, we go all the way back: we assume an MV algebra X as given, and prove that it is an MCD algebra as well. To this end, we have to reconstruct multiple choice with arbitrary probabilities from binary choice with probabilities 1/2. We first consider multiple binary choice, then multiple choice with dyadic rationals as probabilities, and finally apply directed joins and continuity arguments to reach all reals in the unit interval.

Multiple Binary Choice

Let X be an MV algebra. For every n, we define an operator $\bigoplus_n : X^{2^n} \to X$ which takes 2^n arguments from X.

- $\bigoplus_0(x) = x$.
- $\bigoplus_{n+1}(x_i \mid i = 1 \mathrel{..} 2^{n+1}) \; = \; \bigoplus_n(x_i \mid i = 1 \mathrel{..} 2^n) \oplus \bigoplus_n(x_i \mid i = 2^n + 1 \mathrel{..} 2^{n+1})$.

Thus, $\bigoplus_n(x_1, \ldots, x_{2^n})$ is obtained by evaluating a complete binary tree of depth n, whose inner nodes are marked by '$\oplus$' and whose leaves are x_1 through x_{2^n}. Complete binary trees can be pasted into each other:

Proposition 7.1
$\bigoplus_n(\bigoplus_m(x_{ij} \mid j = 1 \mathrel{..} 2^m) \mid i = 1 \mathrel{..} 2^n) \; = \; \bigoplus_{n+m}(x_{ij} \mid i = 1 \mathrel{..} 2^n,\, j = 1 \mathrel{..} 2^m)$
(The 2^n sequences $(x_{ij} \mid j = 1 \mathrel{..} 2^m)$ are concatenated.)

Proof: Induction by n. □

From commutativity and the rectangle law, we can conclude:

Proposition 7.2 The operands of $\bigoplus_n$ can be arbitrarily permuted.

Proof: Induction on n. Case $n = 0$ is obvious, and case $n = 1$ holds by commutativity. For $n > 1$, let $\xi = (x_i \mid i = 1 \mathrel{..} 2^n)$ be the sequence of arguments, and let α be its first quarter, β the second, γ the third, and δ the last.

It suffices to show that two adjacent operands x_k and x_{k+1} can be transposed. If $k \neq 2^{n-1}$, then the pair to be transposed is contained in the first half $\alpha\beta$ or in the second $\gamma\delta$. Because of $\bigoplus_n(\xi) = \bigoplus_{n-1}(\alpha\beta) \oplus \bigoplus_{n-1}(\gamma\delta)$, the pair can be transposed by induction.

The difficult case is $k = 2^{n-1}$, i.e., the last item of β has to be transposed with the first item of γ. By definition,

$$\bigoplus_n(\xi) \; = \; (\bigoplus_{n-2}(\alpha) \oplus \bigoplus_{n-2}(\beta)) \oplus (\bigoplus_{n-2}(\gamma) \oplus \bigoplus_{n-2}(\delta))$$

holds. Applying commutativity to the right subexpression and then the rectangle law to the whole expression yields

$$(\bigoplus_{n-2}(\alpha) \oplus \bigoplus_{n-2}(\delta)) \oplus (\bigoplus_{n-2}(\beta) \oplus \bigoplus_{n-2}(\gamma)) \; = \; \bigoplus_{n-1}(\alpha\delta) \oplus \bigoplus_{n-1}(\beta\gamma)$$

Then the induction hypothesis can be applied, and the expression can be transformed back. □

In the sequel, we shall use an abbreviation: if an argument of $\bigoplus_n$ occurs k times, we shall write $k \cdot x$, e.g., $\bigoplus_2(3 \cdot x,\, 1 \cdot y,\, 0 \cdot z) = \bigoplus_2(x, x, x, y)$.

Multiple Choice with Dyadic Rationals

Now, we construct values corresponding to multiple choice $[p_1 : x_1, \ldots, p_r : x_r]$ with $p_i = k_i/2^n$ where k_i and n are non-negative integers. For simplicity, the exponent of the denominator is written as an index, i.e., we define values

$$\langle k_1 : x_1, \ldots, k_r : x_r \rangle_n$$

where $r \geq 0$, x_i in X, n and k_i are non-negative integers with $k_1 + \cdots + k_r \leq 2^n$.

The definition of these values is in terms of $\bigoplus_n$, where the necessary number of arguments is obtained by filling up with 0.

$$\langle k_1 : x_1, \ldots, k_r : x_r \rangle_n \; = \; \bigoplus_n(k_1 \cdot x_1, \ldots, k_r \cdot x_r, (2^n - \sum_{i=1}^{r} k_i) \cdot 0)$$

In the sequel, we show that the dyadic choice expressions defined above satisfy close analogues of the axioms of MCD. For simplicity, we shall often use abbreviations such as $\langle k_i : x_i,\, l : y,\, m_j : z_j \rangle_n$
for $\langle k_1 : x_1, \ldots, k_r : x_r,\, l : y,\, m_1 : z_1, \ldots, m_s : z_s \rangle_n$.

Proposition 7.3

(1) The entries in a dyadic choice expression may be arbitrarily permuted.

(2) For k_i with $\sum_i k_i \le 2^n$: $\quad \langle 0 : x,\, m_i : y_i \rangle_n = \langle m_i : y_i \rangle_n$

(3) For k, l, and m_i with $k + l + \sum_i m_i \le 2^n$:

$$\langle k : x,\, l : x,\, m_i : y_i \rangle_n = \langle k + l : x,\, m_i : y_i \rangle_n$$

(4) For k_i and l_i with $\sum_i k_i \le 2^n$ and $\sum_i l_i \le 2^n$:

$$\langle k_i : x_i \rangle_n \oplus \langle l_i : x_i \rangle_n = \langle k_i + l_i : x_i \rangle_{n+1}$$

(5) For k_i with $\sum_i k_i \le 2^n$: $\langle k_i : x_i \rangle_n = \langle 2k_i : x_i \rangle_{n+1}$

(6) For k and s_i with $k + \sum_i s_i \le 2^n$ and l_j with $\sum_j l_j \le 2^m$:

$$\langle k : \langle l_j : x_j \rangle_m,\, s_i : y_i \rangle_n \;=\; \langle k\, l_j : x_j,\, 2^m s_i : y_i \rangle_{n+m}$$

(7) For k and l_i with $k + 1 + \sum_i l_i \le 2^n$:

$$\langle k : x,\, l_i : y_i \rangle_n \sqsubseteq \langle k + 1 : x,\, l_i : y_i \rangle_n$$

Proof: Most proofs are straightforward. We give a few hints:

(5) follows from (4) and idempotence.

(6) is shown by applying Prop. 7.1 and the fact $y_i = \bigoplus_m (2^m \cdot y_i)$.

(7) After the transformation into $\bigoplus_n$ expressions, the two sides only differ in that one of the 2^n operands is 0 on the left hand side, and x on the right hand side. The operation '$\oplus$' is continuous, whence monotonic, and $0 \sqsubseteq x$ holds in every MV algebra. □

Multiple Choice with Real Coefficients

In this subsection, we use the dyadic choice expressions of the previous subsection to define choice expressions with real numbers as coefficients.

The unit interval $\mathbb{I}$ is a *continuous domain*, and the dyadic rationals form a basis of this domain. The way-below relation on $\mathbb{I}$ is given by $p \ll q$ iff $p = 0$ or $p < q$. It has the property that for all directed sets D of $\mathbb{I}$, $p \ll \bigsqcup D$ holds iff there is d in D with $p \ll d$. Every p in $\mathbb{I}$ is a directed join of all dyadic rationals d with $d \ll p$.

For every p in $\mathbb{I}$, we define $p^{(n)}$ to be the greatest non-negative integer k with $k/2^n \ll p$. We show several properties of this notion:

Proposition 7.4

(1) $0^{(n)} = 0$ for all n.

(2) If $p \le q$, then $p^{(n)} \le q^{(n)}$.

(3) $2p^{(n)} \le p^{(n+1)}$.

(4) If $k \ne 0$ and $m \ge n$, then $(k/2^n)^{(m)} = 2^{m-n} \cdot k - 1$.

Proof:

(1) is obvious, and (2) holds since $p^{(n)}/2^n \ll p \le q$.

(3) $p^{(n)}/2^n \ll p$ implies $2p^{(n)}/2^{n+1} \ll p$, whence $2p^{(n)} \leq p^{(n+1)}$.

(4) The statement $l/2^m \ll k/2^n$ is equivalent to $l \ll 2^{m-n} \cdot k$. Since the right hand side is not 0, this is in turn equivalent to $l < 2^{m-n} \cdot k$, or $l \leq 2^{m-n} \cdot k - 1$. $\square$

We now define arbitrary choice as follows:

$$[p_1 : x_1,\, \ldots,\, p_r : x_r] = \bigsqcup_{n\geq 0} \langle p_1^{(n)} : x_1,\, \ldots,\, p_r^{(n)} : x_r \rangle_n$$

The dyadic choice expressions in this definition are well-defined since $\sum_i p_i \leq 1$ implies $\sum_i p_i^{(n)} \leq \sum_i p_i \cdot 2^n \leq 2^n$. The join is well-defined because it is directed. For, $\langle p_i^{(n)} : x_i \rangle_n$ equals $\langle 2p_i^{(n)} : x_i \rangle_{n+1}$ by Prop. 7.3 (5), which by Prop. 7.4 (3) and Prop. 7.3 (7) is below $\langle p_i^{(n+1)} : x_i \rangle_{n+1}$.

By Prop. 7.3 (1), the entries in a multiple choice expression may be arbitrarily permuted, i.e., the permutation axiom of MCD is satisfied. By Prop. 7.4 (1) and Prop. 7.3 (2), the expressions satisfy the 0 axiom of MCD. The proof of the remaining MCD axioms is postponed.

The expressions $[p_i : x_i]$ are continuous in every argument x_i as a directed join of continuous functions. Continuity in p_i is a bit more complex. Thanks to the permutation rule, it suffices to show continuity in the first argument. We claim: If D is a directed set in $\mathbb{I}$ with $\bigsqcup D = p$, then $\bigsqcup_{d\in D}[d : x,\, q_i : y_i] = [p : x,\, q_i : y_i]$.

The relation '$\sqsubseteq$' holds by monotonicity (Prop. 7.4 (2) and Prop. 7.3 (7)). For the opposite relation, $p^{(n)}/2^n \ll p = \bigsqcup D$ implies the existence of some d_n in D such that $p^{(n)}/2^n \ll d_n$, whence $p^{(n)} \leq d_n^{(n)}$. Then

$$\begin{array}{rcl}
[p : x,\, q_i : y_i] & = & \bigsqcup_n \langle p^{(n)} : x,\, q_i^{(n)} : y_i \rangle_n \\
& \sqsubseteq & \bigsqcup_n \langle d_n^{(n)} : x,\, q_i^{(n)} : y_i \rangle_n \\
& \sqsubseteq & \bigsqcup_n \bigsqcup_{d\in D} \langle d^{(n)} : x,\, q_i^{(n)} : y_i \rangle_n \\
& = & \bigsqcup_{d\in D} \bigsqcup_n \langle d^{(n)} : x,\, q_i^{(n)} : y_i \rangle_n \\
& = & \bigsqcup_{d\in D}[d : x,\, q_i : y_i]
\end{array}$$

This completes the proof of the continuity of the multiple choice operation.

Finally, we show that a multiple choice with coefficients which happen to be dyadic rationals coincides with the dyadic choice introduced earlier.

Proposition 7.5 $\quad [k_1/2^n : x_1,\, \ldots,\, k_r/2^n : x_r] \;=\; \langle k_1 : x_1,\, \ldots,\, k_r : x_r \rangle_n$

Proof: By applying Prop. 7.3 (2), we may assume without restriction that all k_i are strictly positive. Using Prop. 7.4 (4), we have to show

$$\bigsqcup_{m\geq n} \langle 2^{m-n} \cdot k_i - 1 : x_i \rangle_m \;=\; \langle k_i : x_i \rangle_n.$$

Renaming $m - n$ into m, this is equivalent to $\bigsqcup_{m\geq 0} \langle 2^m \cdot k_i - 1 : x_i \rangle_{n+m} \;=\; \langle k_i : x_i \rangle_n$.

We apply the last version of the fixed point axiom of MV (Section 6), which says: if $b_0 \sqsubseteq a$ and $b_{m+1} = b_m \oplus a$, then $\bigsqcup_{m\geq 0} b_m = a$. Of course, we set $a = \langle k_i : x_i \rangle_n$ and $b_m = \langle 2^m k_i - 1 : x_i \rangle_{n+m}$. The relation $b_0 \sqsubseteq a$ holds by Prop. 7.3 (7) and (5). Next, we compute $b_m \oplus a$.

$$\begin{array}{rcl}
b_m \oplus a & = & \langle 2^m k_i - 1 : x_i \rangle_{n+m} \oplus \langle k_i : x_i \rangle_n \\
& \overset{7.3\,(5)}{=} & \langle 2^m k_i - 1 : x_i \rangle_{n+m} \oplus \langle 2^m k_i : x_i \rangle_{n+m} \\
& \overset{7.3\,(4)}{=} & \langle 2^m k_i - 1 + 2^m k_i : x_i \rangle_{n+m+1} \;=\; b_{m+1}
\end{array}$$ $\square$

Now we are ready to prove the axioms of MCD. Permutation and the 0 law were already handled above. The 1 law holds, since $[1 : x]$ equals $\langle 1 : x\rangle_0$ by Prop. 7.5, which in turn equals x by definition. The addition law is left to the reader. For substitution, we have to show $[p : [r_j : x_j],\, q_i : y_i] = [p\, r_j : x_j,\, q_i : y_i]$.

Because of continuity of multiple choice, it suffices to prove this equation for dyadic rationals.

$$\begin{aligned}[p/2^n : [r_j/2^m : x_j],\, q_i/2^n : y_i] &\overset{7.5}{=} \langle p : \langle r_j : x_j\rangle_m,\, q_i : y_i\rangle_n \\ &\overset{7.3\,(6)}{=} \langle p\, r_j : x_j,\, 2^m q_i : y_i\rangle_{n+m} \\ &\overset{7.5}{=} [(p/2^n)(r_j/2^m) : x_j,\, q_i/2^n : y_i]\end{aligned}$$

This completes the derivation of MCD from MV.

Going Back and Forth

The derivation of MV from MCD and vice versa are inverse to each other. If we start with an MV algebra and construct multiple choice, then $[1/2 : x,\, 1/2 : y]$ equals $\langle 1 : x,\, 1 : y\rangle_1$ by Prop. 7.5, which in turn equals $\bigoplus_1(x,\, y) = x \oplus y$ by definition. This means that we get back the original MV algebra by restriction.

Conversely, if we start with a multiple choice operator, restrict it to the special case $x \oplus y = [1/2 : x,\, 1/2 : y]$, and then reconstruct multiple choice following the lines of this section, then we obtain the original multiple choice back. The proof of this fact is not particularly difficult and hence omitted.

Also, we never considered homomorphisms. The proofs that the homomorphisms of one theory are also homomorphisms of all other theories are straightforward and omitted.

Acknowledgements

This work was begun during my visit at the Theory and Formal Methods Section of Imperial College, London. This visit was made possible by a grant of the Deutsche Forschungsgemeinschaft. The work was continued at Universität des Saarlandes, Saarbrücken, and finished when I stayed with family Nomine as a guest in their home at Lübben, Brandenburg. I like to thank all the people at these different places for their stimulating environment, and all the people who commented upon a draft version of this paper.

References

[1] S.K. Graham. Closure properties of a probabilistic domain construction. In Michael G. Main, A. Melton, Michael Mislove, and D. Schmidt, editors, *Mathematical Foundations of Programming Language Semantics (MFPLS '87)*, pages 213–233. *Lecture Notes in Computer Science 298*, Springer-Verlag, 1988.

[2] R. Heckmann. Product operations in strong monads. In G.L. Burn, S.J. Gay, and M.D. Ryan, editors, *Proceedings of the First Imperial College, Department of Computing, Workshop on Theory and Formal Methods*, Workshops in Computing, pages 159–170. Springer-Verlag, 1993.

[3] C.J. Jones. *Probabilistic Non-Determinism.* PhD thesis, University of Edinburgh, 1990.

[4] C.J. Jones and G.D. Plotkin. A probabilistic powerdomain of evaluations. In *LICS '89*, pages 186–195. IEEE Computer Society Press, 1989.

[5] E. Moggi. Computational lambda-calculus and monads. In *4th LICS Conference*, pages 14–23. IEEE, 1989.

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