Probabilistic Non-determinism

Claire Jones

Doctor of Philosophy

University of Edinburgh

August 1989

(Graduation Date July 1990)

Abstract

Much of theoretical computer science is based on use of inductive complete partially ordered sets (or ipos). The aim of this thesis is to extend this successful theory to make it applicable to probabilistic computations. The method is to construct a "probabilistic powerdomain" on any ipo to represent the outcome of a probabilistic program which has outputs in the original ipo. In this thesis it is shown that evaluations (functions which assign a probability to open sets with various conditions) form such a powerdomain. Further, the powerdomain is a monadic functor on the category **Ipo**.

For restricted classes of ipos a powerdomain of probability distributions, or measures which only take values less than one, has been constructed (by Saheb-Djahromi). In the thesis we show that this powerdomain may be constructed for continuous ipos where it is isomorphic to that of evaluations.

The powerdomain of evaluations is shown to have a simple Stone type duality between it and sets of upper continuous functions. This is then used to give a Hoare style logic for an imperative probabilistic language, which is the dual of the probabilistic semantics.

Finally the powerdomain is used to give a denotational semantics of a probabilistic metalanguage which is an extension of Moggi's λ_c-calculus for the powerdomain monad. This semantics is then shown to be equivalent to an operational semantics.

Acknowledgements

I would like to acknowledge the invaluable help and advice I have received from Gordon Plotkin, who supervised this research. I would also like to acknowledge Eugenio Moggi for his work on monads and also for reading an early draft of this thesis, and Alistair Sinclair for showing me some of the literature on probabilistic computation.

I would also like to thank the many people who have helped and supported me in various ways: particularly Simon Tett for cups of tea; Dave Berry for his forbearance; Rod Burstall for giving me a job; Sarah Hughes for more cups of tea; Mads Dam, Charlotte Elgaard, Andreas Knoble, James Crook, Fiona Robertson and finally my parents.

Declaration

I hereby declare that this thesis has been composed by myself and that except where otherwise stated the work is my own. Some of the material from Chapters 4 and 5 was published as "A Probabilistic Powerdomain of Evaluations" by C. Jones and G. D. Plotkin in the Proceedings of the Fourth Annual Symposium on Logic in Computer Science 1989 (Asilomar, California).

Table of Contents

Chapter 1

Introduction and Related Work

In this thesis we give a theoretical framework for studying probabilistic computation. We want to be able to give semantics for probabilistic languages, say whether a program or algorithm produces a desired output, or perhaps calculate the probability of getting some result. We need to express program results which are "probabilistic", i.e. give some probability distribution over the set of possible results. Existing semantic methodologies and program logics cannot express these sorts of results without some modification.

The approach to probabilistic non-determinism taken in this thesis is to extend the domain theory style of semantics—sets with a complete partial order (cpos) and functions which preserve the order and least upper bounds. We use the idea of Smyth [37] of open sets as tests or properties, and a computation being something with certain properties. So we model a random computation as something which has a probability of passing each particular test or being in each open set. Functions on the open set lattice of a topological space such as these are known as evaluations and are well known from mathematics (e.g. Lawson [23], Birkhoff [5]).

Other authors studying probabilistic non-determinism have started from measure theory giving them the useful tool of integration, we develop a theory of integration over evaluations with similar properties. We use the recent work of Moggi who considered monadic functors (with other properties) giving computational models in the sense that if X is an object representing a datatype, and T a monad, then TX is the possible computations on this datatype. We also use the ideas of Abramsky [1] that Stone Duality Theorem can be interpreted as a duality between semantics and program logic. The duality theorem we use is based on an idea first expressed by Kozen in [22], that integration gives a duality between measures and functions.

1.1 Probabilistic Computation

By a probabilistic computation we essentially mean programs or algorithms which make non-deterministic choices governed by some probability distribution, e.g. assignment to a random variable or an expression construct which branches between two or more possible expressions according to some probability. Most high-level languages contain some facility for making probabilistic computations, usually in the form of a random number generator from which assignment can be made. In [21], Kozen investigated an operational semantics for this form of probabilistic computation.

However most theoretical work on probabilistic computation uses the "probabilistic choice" construct, as does this thesis. We will write $e \text{ or}_p e'$ meaning the expression e with probability p and the expression e' with probability $1-p$ (as used by Graham, Saheb-Djahromi, Frutos Escrig [14,35,9] and others).

These two forms of probabilistic computation are equivalent when we only allow assignment to a random variable from a probability distribution over a finite set. The computational power of of a language with the construct or_p is is known (Gill, [13]) to be the same as that allowing only the probability $1/2$ provided p is always computable, a reasonable restriction for actually implementing this as part of a programming language. However the theory is no more complicated that for letting p be any real number in $[0, 1]$.

Various investigation has been made into the relative power of randomized algorithms. Probabilistic Turing machines were first defined by de Leeuw *et al* [8]; they made probabilistic "coin-tossing" moves, taking one of two branches each with probability 1/2. They showed that only partial recursive function were computable by these machines. See Gill [13] for more information on the position of randomized machines in the complexity hierarchies.

Random computations can be directly used to model some processes, e.g. communicating systems with some probability of transmission failure, real-time systems where with a low probability two events may occur simultaneously and concurrency modelled by a probabilistic scheduler.

1.2 Randomized Algorithms

Study of languages with these sorts of constructs has become important recently with the development of fast randomized algorithms. Historically the main use of these algorithms was via Monte-Carlo methods which did not hold great interest for theoretical computer scientists.

The recent interest in randomized algorithms came from two revolutionary algorithms. Firstly Rabin [32] along with Solovay and Strassen [40] gave a very fast algorithm for prime numbers, exploiting some number theoretic results about primes. The randomization in their algorithm was picking random integers between 1 and n with equal probability of choosing any particular integer. It is only "probabilistically" correct in that it may fail (i.e. claim a number is prime if it isn't, the other type of error is not possible) but with a known probability of order 2^{-k} where k is a parameter of the program. So the probability of an error can be made as small as is desired, e.g. below the probability of a hardware error occurring while the program is run. In a survey paper Welsh [41] describes some other algorithms of this type (e.g. Berlekamp or Angluin and Valiant [4,2]), he claims that the existence of this type of algorithm for a problem corresponds to it being tractable.

Secondly Karp [20] found a fast approximate solution to the travelling salesman problem. This algorithm gives an approximate result, i.e. one which is good but not necessarily the best solution, this is sufficient for many practical purposes.

1.3 Outline

The next chapter gives most of the background material for the thesis; there are two main parts, one about ipos and the other about category theory and duality. Chapter 3 describes the measure theory with special reference to ipos, and then develops a similar theory for evaluations.

In Chapter 4 the probabilistic powerdomain functor of evaluations is defined and shown to be a monad and some other results are given. Chapter 5 considers the powerdomain for the special case of continuous evaluations, giving a characterisation of evaluations as the sup of a directed set of linear combinations of point evaluations.

In Chapter 6 a Stone-type duality theorem between the ipos of evaluations on an ipo and the complete Heyting algebras of upper-continuous functions from the ipo to $[0,1]$ is given. In Chapter 7 the duality theorem is used to give a logic of partial correctness and a logic of total correctness for a simple probabilistic imperative language where the terms of the logic are upper continuous functions from an ipo of states to $[0,1]$. Chapter 8 gives denotational and operational semantics of a functional metalanguage extending Moggi's λ_c calculus and proves their equivalence.

1.4 Related Work

In this section we will describe work related to this thesis. Further reference to some of these authors will be made where appropriate.

Firstly we will consider the most relevant work, that of Saheb-Djahromi, Plotkin, Graham and Frutos Escrig who all worked with the notion of cpos as domains and gave probabilistic powerdomains consisting of sets of measures on the Borel sets of a cpo.

We first look at Saheb-Djahromi's paper [35]. In this he defined the partial order on measures $\mu \sqsubseteq \eta$ iff for all open sets O, $\mu(O) \leq \eta(O)$. He worked with ω-algebraic cpos with bottom and showed that for these cpos, the partial order is complete. He also showed that finite discrete distributions (those which are finite linear combinations of point-mass measures) form a basis for the powerdomain. He then showed that the partial order is related in a natural way to the Egli-Milner ordering used in the study of powerdomains for non-determinism, that various semantic functions are continuous on the powerdomain and that continuous functions on cpos can be extended to continuous functions on the powerdomains of those cpos.

In a later paper [34] he applied these results to the language LCF. He extended the language by adding a probabilistic choice operation and gave it an operational and denotational semantics. The operational semantics is given by a Markov chain over a countable (infinite) subset of the (uncountably many) terms of the language. In his denotational semantics, function types are represented by functions from one set of measures to another, without any linearity condition.

In [9], Frutos Escrig considers probabilistic powerdomains over SFP domains. He defines two more "finite" probabilistic powerdomains which can be embedded in Saheb-Djahromi's one by using trees and information systems to try to get a functor on SFP domains. The information systems powerdomain gives him such a functor which he then uses to give the semantics of a concurrent while language. He then relates this semantics to one given in terms of the original powerdomain and notes that the only terms which arise in the new semantics are one which arise from the embedding.

Graham [14] and Plotkin [30] showed that RSFP domains are preserved by Saheb-Djahromi's powerdomain. Graham showed that domain equations involving the powerdomain can be solved by a universal domain approach, he also considered issues of computability. He and Plotkin both tried to get some axiomatisation of a probabilistic powerdomain as a cpo with some notion of probabilistic sum.

The work in this thesis extends that described above in that it defines a powerdomain on all ipos, not just the RSFP and ω-continuous ones, which is also a functor.

We now consider the work of another two authors who considered the semantics of probabilistic languages in a very different setting.

In [21], Kozen gives semantics in terms of linear continuous operators on L-spaces—partially ordered Banach spaces. This arises from considering a measurable space X in which each of n variables take values, so that the states of a probabilistic program are measures on X^n. He then observes that spaces like this naturally have the structure of a real Banach space and also a conditionally complete lattice. Here the ordering is $\mu \sqsubseteq \eta$ iff $\mu(A) \leq \eta(A)$ for all A measurable. The semantics of the while loop is given in terms of least fixed points of affine, isotone transformations. He also shows how partially ordered domain used in deterministic semantics can be naturally embedded in partially ordered Banach spaces. He claims that in his approach, computationally meaningful functions are assumed to be norm continuous rather than order continuous as in Scott-Strachy style semantics.

Yamada [43] takes a similar approach to Kozen. He concentrates on topological spaces which avoid the measure theoretic problems such as products of Borel measures not being Borel measures and uses only regular measures. He gives several axiomatisations of his randomized domains and points out a weakness in a characterisation given by Kozen. His domains form a cartesian closed category and he uses band projections as an equivalent to retractions to get a universal domain.

Kozen went on to consider logics for making assertions about probabilistic programs in [22]. He notes that for probabilistic programs it is necessary to use an arithmetic operator $+$ to combine terms rather than just the usual logical operators. As terms he uses measurable functions using a different duality to the one we will give later in this thesis. He describes his logic as a probabilistic analogue of propositional dynamic logic (PPDL).

A probabilistic extension to dynamic logic was also investigated by Feldman [11]. He uses as terms, expressions in propositional dynamic logic, that is programs and logic terms along with certain terms in real number theory and frequencies—unnormalised measures. Thus he can very naturally express properties like a program having a probability p of satisfying some property. He notes that his logic is very similar in expressive power to Kozen's but decidable with fewer restrictions than his.

Another scheme for reasoning with probabilities is given by Nilsson [27]. He gives a semantical generalisation of logic in which truth values are replaced by probabilities and gives an entailment relation. This relation reduces to modus ponens when all the probabilities are either zero or one. The entailment relation involves multiplication of matrices which get large for many sentences, so approximation methods have to be used.

In [10] Fagin *et al* make a similar attempt using linear programming results to show that decidability is NP-complete, no worse then the deterministic case.

Also they investigate the non-measurable case by using Dempster-Shafer belief functions. Their logic is equivalent to Kozen's PPDL without the program terms.

The logic given in the thesis is most similar to PPDL, in that it uses functions as terms, and $+$ as a combinator. But we avoid the use of the combinator $*$ from dynamic logic, using instead just limits of increasing sequences of functions. The other work described above concentrate on adding probabilities to logic which gives systems which are easier to interpret but harder in which to express the necessary combinators.

Chapter 2

Preliminary Material

In this chapter I review the material on which the rest of this thesis depends. It divides simply into two sections, one on domain theory, the other on category theory and duality.

For an introduction to category theory see e.g. Mac Lane [36]. For more on the Stone dualities see Johnstone [17] and for its application to semantics see Abramsky [1].

2.1 Domains

In this section we briefly review some basic definitions in domain theory, establish some notation and give some useful results.

For more general information see Gunter and Scott [15] or Plotkin [29] although the definitions follow those in Johnstone [17] most closely.

2.1.1 Notation

We will denote a partial function, that is one which need not be defined on all the elements of its domain, by $f: X \rightharpoonup Y$. For expressions involving partial functions we will write $e\downarrow$ to mean that the expression e denotes a value, i.e. $f(x)\downarrow$ means that $f(x)$ exists. We can now introduce the notation $e_1 \simeq e_2$ to mean $e_1\downarrow$ iff $e_2\downarrow$ and if $e_1\downarrow$ then $e_1 = e_2$ and when we have some partial order $\sqsubseteq$, $e_1 \sqsubseteq\!\!\!\!\sim e_2$ to mean $e_1\downarrow$ implies $e_1 \sqsubseteq e_2$.

2.1.2 The Ipo

In this section we define inductive partial orders (ipos), and give the definitions of continuity of functions between ipos and the Scott topology of an ipo. We then describe some constructions on ipos.

We use the usual definition of a partially ordered set as a set with a transitive, reflexive relation $\sqsubseteq$ where $x \sqsubseteq y$ and $y \sqsubseteq x$ implies $x = y$.

A subset X, of a partially ordered set, is *directed* if it is non-empty and any pair of elements of X have an upper bound in X.

A *inductive partial order* (ipo) is a partially ordered set such that every directed subset has a least upper bound (or lub). These sets are also called directed complete (as in [17]). We will usually denote this partial order by the symbol $\sqsubseteq$ and the lub of a directed set X by $\bigsqcup X$.

A partial function between ipos is *continuous* iff it preserves the partial order and the least upper bounds of directed sets, i.e. f is continuous if whenever $x \sqsubseteq y$, then $f(x) \sqsubseteq\!\!\!\!\sim f(y)$, so for any directed set X, the set $f(X) = \{f(x) \mid x \in X\}$ is directed if it is non-empty, and $f(\bigsqcup X) \simeq \bigsqcup f(X)$ where $\bigsqcup f(X)$ is undefined if $f(X)$ is empty.

The *Scott topology* of an ipo is the set of subsets which are upper closed and inaccessible by lubs of directed sets, i.e. O is open iff for any $x \in O$ and y such that $x \sqsubseteq y$ then $y \in O$ and for all directed sets X, $\bigsqcup X \in O$ implies there exists $x \in X$ such that $x \in O$. It is easy to show that this defines a topology.

It is well-known that a function is continuous w.r.t. the Scott topology iff it is continuous in the sense described above (i.e. preserves directed lubs).

For any point x in an ipo, the subset $O_x = \{y \mid y \not\sqsubseteq x\}$ is open.

The *product* of ipos $P_1, \ldots, P_n$ is the product of the underlying sets with the partial order defined by

$$\langle x_1, \ldots, x_n \rangle \sqsubseteq \langle y_1, \ldots, y_n \rangle \iff \forall i, x_i \sqsubseteq y_i$$

and is denoted by $\prod_{i=1}^n P_i$. The *projection functions* $\pi_i : \prod_{i=1}^n P_i \to P_i$ defined by $\pi_i(\langle x_1, \ldots, x_n \rangle) = x_i$ are easily seen to be continuous.

The *sum* of ipos $P_1, \ldots, P_n$ is defined as the disjoint union of the base sets, i.e. $\bigcup_{i=1}^n \{i\} \times P_i$ and partial order

$$(i, x) \sqsubseteq (j, y) \text{ iff } i = j \text{ and (for } i = j\text{) } x \sqsubseteq y$$

and denoted $\sum_{i=1}^n P_i$. Associated with this are the injection maps $\mathrm{in}_i : P_i \to \sum_{i=1}^n P_i$ defined by $\mathrm{in}_i(x) = (i, x)$ which are clearly continuous.

Finally the set of continuous (total) functions $f : P \to Q$ form an ipo with the partial order $f \sqsubseteq g$ iff for all $x \in P$, $f(x) \sqsubseteq g(x)$, this ipo is denoted $P \to Q$. Similarly the set of partial continuous functions $f : P \rightharpoonup Q$ with partial order $f \sqsubseteq g$ iff for all $x \in P$, $f(x) \precsim g(x)$ and this ipo is denoted $P \rightharpoonup Q$.

We will use **Ipo** to denote the category of ipos and continuous functions between them.

2.1.3 Algebraic and Continuous Domains

In theoretical computer science it is often natural to impose extra conditions on ipos. In this section and the next we give the definitions of some special types of ipos.

An element of an ipo, x, is said to be *finite* iff whenever $x \sqsubseteq \bigsqcup X$, for some directed set X, there exists y in X such that $x \sqsubseteq y$. It is easy to show that the set $J_a = \{x \mid a \sqsubseteq x\}$ is Scott open iff a is finite.

An ipo is *algebraic* iff its finite elements form a *basis*, that is every x in P is the directed limit of finite elements. Equivalently, the set of finite elements below any x is directed and has lub x. If an ipo is algebraic, then the sets J_a where a ranges over the finite elements of P form a (sub)basis for the Scott topology (any open set is a union of such sets).

An ipo is ω-*algebraic* iff it contains countably many finite elements which form a basis as above. Then the Scott topology has a countable (sub)basis of the sets J_a where a is finite so every open set is a countable union of sets of the form J_a.

The relation *well-below*, written $\ll$, is defined in terms of $\sqsubseteq$ by $x \ll y$ iff for all directed sets X, $y \sqsubseteq \bigsqcup X$, implies there exists d in X such that $x \sqsubseteq d$.

The following are simple consequences of this definition.

$$\begin{array}{rcl} x \sqsubseteq y \ll z \sqsubseteq w & \Rightarrow & x \ll w \\ x \ll y & \Rightarrow & x \sqsubseteq y \\ x \ll x & \Longleftrightarrow & x \text{ is finite} \end{array}$$

We call a directed set X, a *well-below directed set* if any d in X is well-below $\bigsqcup X$.

An ipo P is *continuous* if for any x in P, the set of elements well-below x is directed and has lub x.

In a continuous ipo, well-below is dense in the sense that if $x \ll z$ then there exists some y with $x \ll y \ll z$. From this it is easy to prove that if P is continuous then for all $c \in P$ the set

$$V_c = \{x \mid c \ll x\}$$

is open in the Scott topology. Also the Scott topology has a subbasis given by the sets V_x over all x in X since for any open set O

$$O = \bigcup_{x \in O} V_x$$

An ipo is *ω-continuous* if it has a countable set B of elements (a basis) such that for any $x \in X$, the set $\{b \mid b \in B, b \ll x\}$ is directed and has lub x, i.e. every element is the lub of a well-below sequence of points in the basis. Note even elements of the basis must be such a limit. The Scott topology of an ω-continuous ipo has a countable subbasis, namely the sets V_b for b ranging over the basis. It is well-known that an ipo is ω-continuous iff it is continuous and has a countable subbasis.

If x is any element of an ω-continuous ipo P and has a (countable) well-below set b_n then

$$\bigcap_n V_{b_n} = \{y \mid x \sqsubseteq y\}$$

since if $y \in P$ and $x \sqsubseteq y$ then y is in V_{b_n} for all n (as $b_n \ll x \sqsubseteq y$) and if for all n, y is in V_{b_n}, so $b_n \ll y$, then $b_n \sqsubseteq y$ hence $\bigsqcup_n b_n \sqsubseteq y$ and $\bigsqcup_n b_n = x$. The fact that we can write $\{y \mid x \sqsubseteq y\}$ (which we will denote A_x) as a countable intersection of open sets will be useful later.

Now we give an easy lemma to show that in a finite product of ipos, x is well below y if and only if for all the projections π_i, $\pi_i(x) \ll \pi_i(y)$. The proof is given for the product of two ipos; the extension to a finite product is obvious.

Lemma 2.1 *In the product of two ipos P and Q, $\langle x_1, x_2 \rangle \ll \langle y_1, y_2 \rangle$ iff $x_1 \ll y_1$ and $x_2 \ll y_2$*

Proof Suppose $x_i \ll y_i$ for $i = 1, 2$. Suppose X is a directed subset of $P \times Q$ such that $\langle y_1, y_2 \rangle \sqsubseteq \bigsqcup X$. Then $\pi_i(X)$ is clearly directed and $y_i \sqsubseteq \bigsqcup \pi_i(X)$ for $i = 1$ and $i = 2$, so since $x_1 \ll y_1$ and $x_2 \ll y_2$ then we have d_1, d_2 such that $x_i \sqsubseteq \pi_i(d_i)$, and as X is directed, we let d be an upper bound of d_1 and d_2 then $\langle x_1, x_2 \rangle \sqsubseteq d$ since $x_i \sqsubseteq \pi_i(d_i) \sqsubseteq \pi_i(d)$ for $i = 1$ and $i = 2$ so $\langle x_1, x_2 \rangle \ll \langle y_1, y_2 \rangle$.

Conversely, suppose that $\langle x_1, x_2 \rangle \ll \langle y_1, y_2 \rangle$. If we have a directed subset X of P such that $y_1 \sqsubseteq \bigsqcup X$, then the (clearly) directed set $X \times \{y_2\}$ satisfies $\langle y_1, y_2 \rangle \sqsubseteq \bigsqcup X \times \{y_2\}$ so there exists $\langle d, y_2 \rangle$ in $X \times \{y_2\}$ such that $\langle x_1, x_2 \rangle \sqsubseteq \langle d, y_2 \rangle$ which implies $x_1 \sqsubseteq d$, thus $x_1 \ll y_1$. Similarly we can also show $x_2 \ll y_2$. ▮

From this lemma we can show that if P and Q are two continuous (or ω-continuous) ipos, then their product is also continuous (or ω-continuous). In the continuous case the set of points well-below $\langle x, y \rangle$ is simply the product of the set of points well-below x and y respectively; it is clearly directed with lub $\langle x, y \rangle$. In the ω-continuous case the product has a basis which is the product of the bases of P and Q.

Since the Scott topology of a continuous ipo is given by unions of sets of the form V_x for any x, the Scott topology on $X \times Y$ is given by unions of the sets $V_{\langle x,y \rangle}$ for x in X and y in Y. But by the lemma above it is clear that $V_{\langle x,y \rangle} = V_x \times V_y$ (since $\langle x, y \rangle \ll \langle a, b \rangle$ iff $x \ll a$ and $y \ll b$ i.e. $a \in V_x$ and $b \in V_y$). So any open set in $X \times Y$ is given by a union of products of sets open in X and Y respectively, this implies that the Scott topology on $X \times Y$ is the same as the product topology (which is the topology generated by products of open sets). This is not true for products of arbitrary ipos, where the Scott topology is in general larger than the product topology.

Finally we show that if we take a directed set X and for each x in X there is some well-below directed set for x, X_x, then $\bigcup_{x \in X} X_x$ is a well-below directed set for $\bigsqcup X$. First we need to show the union is directed, but if d_1, d_2 are in this union, then $d_1 \in X_{x_1}$ and $d_2 \in X_{x_2}$ for some x_1, x_2 in X, then as X is directed we can find an upper bound for x_1, x_2 in X, say x. Then for $i = 1$ and $i = 2$, $d_i \ll x$, hence as $\bigsqcup X_x = x$, there exist e_1, e_2 in X_x with $d_i \sqsubseteq e_i$, then the joint upper bound to e_1, e_2 in X_x is an upper bound for d_1 and d_2 and is in the union, hence it is directed. Clearly if $d \in \bigcup_{x \in X} X_x$ then $d \ll x$ for some x in X and so $d \ll \bigsqcup X$, so $\bigsqcup(\bigcup_{x \in X} X_x) \sqsubseteq \bigsqcup X$ and if e is any upper bound for $\bigcup X_x$, for all x in X, $\bigsqcup X_x \sqsubseteq e$, i.e. $x \sqsubseteq e$, i.e. $\bigsqcup X \sqsubseteq e$. So $\bigsqcup X$ is the lub of $\bigcup_{x \in X} X_x$.

2.1.4 RSFP Domains

We will first define SFP domains (finitely ω-algebraic) which are colimits in the category of embeddings of finite ipos, then we will give an equivalent characterisation in terms of functions $P \rightharpoonup P$. Next we will define RSFP domains (finitely ω-algebraic) as retracts of SFP domains and then give the analogous definition in terms of functions for RSFP.

We first define an *embedding* from an ipo P to an ipo Q to be a pair of partial continuous functions $f: P \rightharpoonup Q$ and $f^R: Q \rightharpoonup P$ such that $f^R \circ f = \mathrm{id}_P$ and $f \circ f^R \sqsubseteq \mathrm{id}_Q$. We write this as $P \overset{f}{\lhd} Q$. Such functions are also called projection-pairs . We will denote the category with objects as ipos and projection-pairs as morphisms by $\mathbf{Ipo}^E$. In [31] it is shown how colimits of sequences $\langle P_n, f_n \rangle$ where f_n is an embedding $P_n \overset{f_n}{\lhd} P_{n+1}$ can be formed (as colimits of this diagram in the category $\mathbf{Ipo}^E$) by taking the colimit to be the set $\{x: \omega \rightharpoonup \bigcup_n P_n \mid (\forall n, x_n\downarrow \Rightarrow x_n \in P_n \wedge x_n \simeq f_n(x_{n+1})) \text{ and } \exists n, x_n\downarrow\}$ partially ordered by $x \sqsubseteq y$ iff $\forall n$, $x_n \precsim y_n$.

An ipo is *finite* if it contains only finitely many elements.

An ipo is said to be *SFP* if it is the limit in the sense defined above of a sequence of finite ipos, that is $D = \lim_n \langle D_n, f_n \rangle$ where D_n is finite and f_n is an embedding, $D_n \overset{f_n}{\lhd} D_{n+1}$.

This definition is equivalent to the existence of a sequence of partial functions $f_n{:}D \rightharpoonup D$ such the range of each f_n is finite, $f_n^2 = f_n$, $f_n \sqsubseteq f_{n+1}$ and finally $\bigsqcup f_n = \mathrm{id}_D$.

A *retract* of an ipo P, is an ipo Q with partial functions $e: Q \rightharpoonup P$ and $r: P \rightharpoonup Q$ such that $e \circ r = \mathrm{id}_Q$. An ipo is *RSFP* iff it is a retract of an SFP ipo.

This definition can also be expressed in terms of the existence of certain functions $E \rightharpoonup E$ as follows: E is RSFP iff $\exists g_n : E \rightharpoonup E$ with $g_n(E)$ finite, $g_n \sqsubseteq g_{n+1}$ and $\bigsqcup g_n = \mathrm{id}_E$.

Plotkin has showed that RSFP ipos are closed under the operation of taking colimits of sequences of embeddings, that is the limit of a sequence of RSFP ipos is also RSFP. See e.g. Kamimura and Tang [19] for further reference.

2.1.5 Remarks

The definitions given in this section may not be the most familiar to the reader. The main differences are in the use of directed sets rather than ω-chains and the fact that ipos do not necessarily have a least element. Countable chains are a special case of a directed set, so these definitions are stronger. In some of the results, this strength is necessary to the proof. However if we are only considering a subclass of ipos with some countable condition (e.g. SFP, ω-algebraic, ω-continuous) than (as is well-known) the definitions with directed sets and with countable chains are equivalent. The algebraicity conditions will not be mentioned much in this thesis, although they were important to some of the related work.

2.2 Category Theory

This section contains the main definitions of the categorical concepts that will be needed later and a proof of a theorem of Manes which shows that given a construction like that of the Kleisli category of a monad, if the construction satisfies the equations necessary for it to be a category, then this construction gives a monadic functor.

A *category* consists of objects and morphisms from one object to another such that we can compose morphisms, if we have two morphisms $f : A \to B$ and

$g: B \to C$, their composition is a morphism $g \circ f: A \to C$, and this composition is associative, i.e. $h \circ (g \circ f) = (h \circ g) \circ f$ and for every object A, there is an identity morphism id_A such that $f \circ \mathrm{id}_A = f$ and $\mathrm{id}_A \circ g = g$.

A *functor* $T: \mathbf{C} \to \mathbf{D}$ is a structure preserving map on categories. Thus it consists of a function from objects of $\mathbf{C}$ to objects of $\mathbf{D}$ and one from morphisms of $\mathbf{C}$ to morphisms of $\mathbf{D}$ such that if $f: A \to B$ then $T(f): T(A) \to T(B)$ and $T(g \circ f) = T(g) \circ T(f)$ whenever $g \circ f$ exists and finally $T(\mathrm{id}_A) = \mathrm{id}_{T(A)}$ for all objects A.

A *natural transformation* $\alpha: S \to T$ between two functors $S, T: \mathbf{C} \to \mathbf{D}$ consists of a function from objects of $\mathbf{C}$ to morphisms of $\mathbf{D}$, written $A \mapsto \alpha_A$ such that $\alpha_A: S(A) \to T(A)$ and for all morphisms $f: A \to B$ in $\mathbf{C}$ we have $T(f) \circ \alpha_A = \alpha_B \circ S(f)$.

A well-known definition is that of the functor category, where objects are functors and morphisms are natural transformations. The identity natural transformation on a functor T is given by $A \mapsto \mathrm{id}_{T(A)}$ and composition of natural transformations by $(\beta \circ \alpha)_A = \beta_A \circ \alpha_A$.

Given two functors $F: \mathbf{C} \to \mathbf{D}$ and $G: \mathbf{D} \to \mathbf{C}$, we say F is *left adjoint* to G (equivalently G is right adjoint to F) if there is a bijection, natural in A and B between morphisms $f: A \to GB$ in $\mathbf{C}$ and $\bar{f}: FA \to B$. This is equivalent to saying G has a left adjoint provided for each object A of $\mathbf{C}$ we can find an object FA in $\mathbf{D}$ and a morphism $\eta_A: A \to GFA$ which is universal in that for all $f: A \to GB$ there is a unique $\bar{f}: FA \to B$ satisfying $f = G(\bar{f}) \circ \eta_A$.

A *monad* on a category $\mathbf{C}$ is a functor $T: C \to C$ with two natural transformations $\eta: \mathrm{id}_C \to T$ and $\mu: T^2 \to T$ which satisfy the following commutative diagrams.

$$
\begin{tikzcd}
T \arrow[r, "T\eta"] \arrow[dr, "\mathrm{id}_T"'] & TT \arrow[d, "\mu"] & T \arrow[l, "\eta_T"'] \arrow[dl, "\mathrm{id}_T"] \\
 & T &
\end{tikzcd}
$$

and

$$
\begin{tikzcd}
TTT \arrow[r, "T\mu"] \arrow[d, "\mu_T"] & TT \arrow[d, "\mu"] \\
TT \arrow[r, "\mu"] & T
\end{tikzcd}
$$

These diagrams are interpreted in the functor category described above.

Monads arise naturally as the composition of a functor and its left adjoint. The next definitions are motivated by trying to resolve a arbitrary monad into a pair of adjoint functors whose composition is the original monad.

Given a monad T on a category $\mathbf{C}$ we can define a T-algebra to be a pair (A, α) where A is an object in $\mathbf{C}$ and $\alpha\colon TA \to A$ a morphism in $\mathbf{C}$ that satisfies the diagrams below.

$$
\begin{tikzcd}
A \arrow[r, "\eta_A"] \arrow[dr, "\mathrm{id}_A"'] & TA \arrow[d, "\alpha"] \\
 & A
\end{tikzcd}
\qquad
\begin{tikzcd}
TTA \arrow[r, "\mu_A"] \arrow[d, "T\alpha"] & TA \arrow[d, "\alpha"] \\
TA \arrow[r, "\alpha"] & A
\end{tikzcd}
$$

We then define a morphism of T-algebras $f:(A,\alpha) \to (B,\beta)$ as a morphism $f: A \to B$ of $\mathbf{C}$ which satisfies the diagram

$$\begin{tikzcd}
TA \arrow[r, "Tf"] \arrow[d, "\alpha"] & TB \arrow[d, "\beta"] \\
A \arrow[r, "f"] & B
\end{tikzcd}$$

The category which has T-algebras as objects T-algebra morphisms as morphisms is denoted by $\mathbf{C}^T$. The functor which takes X to (TX, μ_X) has as a right adjoint the obvious forgetful functor and the composition of these two functors gives back the original monadic functor T. We can restate this as saying that for any $f: X \to A$, where A is a T-algebra, has a unique $\bar{f}: TX \to A$ which is a T-algebra morphism (given by $\alpha \circ Tf$) and such that $\eta_A \circ \bar{f} = f$.

Another resolution of a monadic functor into a pair of adjoint functors is given by the Kleisli category. The Kleisli category of a monad (see [36, page 143]), $\mathbf{C}_T$, is defined by associating to each object in the original category a corresponding object in $\mathbf{C}_T$ and to each morphism $f: X \to TY$ in $\mathbf{C}$, a morphism $\hat{f}: X \to Y$ in $\mathbf{C}_T$. The identity morphisms in $\mathbf{C}_T$ are given by η_X and composition of $\hat{f}$ and $\hat{g}$ by $\mu_Z \circ Tg \circ f$. The functor which takes f to $\eta_Y \circ f$ has a right adjoint and the composition of these functors is again the original monadic function T.

In showing that the Kleisli category is a category it is necessary to check that the composition given above is associative and that the identity morphism is indeed a right and left identity. These equations can be conveniently expressed in terms of an operation m on objects, an operation $\dagger$ which acts on morphisms $f: X \to m(Y)$ to give a morphism $f^\dagger: m(X) \to m(Y)$ and a set of morphisms $i_X: X \to m(X)$ for each object X, in which form they become,

$$(i_X)^\dagger \quad = \quad \mathrm{id}_{m(X)} \tag{2.1}$$

$$
\begin{aligned}
f^{\dagger} \circ i_X &= f &(2.2)\\
(g^{\dagger} \circ f)^{\dagger} &= g^{\dagger} \circ f^{\dagger} &(2.3)
\end{aligned}
$$

As is well-known, given any monadic functor T, these equations are satisfied by $m(X) = TX$, $f^{\dagger} = \mu_Y \circ Tf$ and $i_X = \eta_X$. We now prove a theorem which states that given operations which satisfy the three equations above, then m extends to a monadic functor. It was previously shown by Manes in [24].

Theorem 2.2 *An operation m on the objects of a category is part of a a monadic functor iff there exists an operation $\dagger$ which takes a morphism $f\colon X \to m(Y)$ to $f^{\dagger}\colon m(X) \to m(Y)$ and morphisms $i_X\colon X \to m(X)$ which obey equations*

$$
\begin{aligned}
(i_X)^{\dagger} &= id_{m(X)}\\
f^{\dagger} \circ i_X &= f\\
(g^{\dagger} \circ f)^{\dagger} &= g^{\dagger} \circ f^{\dagger}
\end{aligned}
$$

as above.

Proof Note: X,Y,Z range over all objects in the category and f,g over all morphisms.

First we show that a monadic functor T gives rise to these operations. Define

$$i_X = \eta_X$$

and

$$f^{\dagger} = \mu_Y \circ Tf.$$

Then the equations above become

$$
\begin{aligned}
\mu_X \circ T\eta_X &= \mathrm{id}_{TX}\\
\mu_Y \circ Tf \circ \eta_X &= f\\
\mu_Z \circ T(\mu_Z \circ Tg \circ f) &= \mu_Z \circ Tg \circ \mu_Y \circ Tf
\end{aligned}
$$

However we may write the commutative diagrams satisfied by a monad on page 25 as

$$\begin{aligned}
\mu_X \circ T\eta_X &= \mathrm{id}_{TX} \\
\mu_X \circ \eta_{TX} &= \mathrm{id}_{TX} \\
\mu_X \circ T\mu_X &= \mu_X \circ \mu_{TX}
\end{aligned}$$

This gives the first equation trivially. For the second note that since η is a natural transformation, we know

$$Tf \circ \eta_X = \eta_{TY} \circ f$$

then substituting this into the L.H.S of equation 2 above and applying the second monad equation we get

$$\mu_Y \circ \eta_{TY} \circ f = \mathrm{id}_{TX} \circ f = f$$

as required. Finally to prove the third equation is satisfied, note that

$$\mu_Z \circ T(\mu_Z \circ T(g) \circ f) = \mu_Z \circ T\mu_Z \circ T^2 g \circ Tf$$

and by substituting the third monad equation $\mu_X \circ T\mu_X = \mu_X \circ \mu_{TX}$ we get

$$L.H.S. = \mu_Z \circ \mu_{TZ} \circ T^2 g \circ Tf$$

but since μ is a natural transformation we know

$$\mu_{TZ} \circ T^2 g = Tg \circ \mu_Y$$

so substituting again we get

$$L.H.S. = \mu_X \circ Tg \circ \mu_Y \circ Tf = R.H.S.$$

Now we shall show that given $\dagger$ and i_X we obtain a monadic functor. We define

$$m(f) = (i_Y \circ f)^{\dagger}$$

then to show this gives a functor simply note

$$m(\mathrm{id}_X) = (i_X \circ \mathrm{id}_X)^\dagger = i_X^\dagger = \mathrm{id}_{m(X)}$$

by Equation 2.1 and also

$$m(g) \circ m(f) = (i_Z \circ g)^\dagger \circ (i_Y \circ f)^\dagger = ((i_Z \circ g)^\dagger \circ i_Y \circ f)^\dagger = (i_Z \circ g \circ f)^\dagger = m(g \circ f)$$

by using equations 2.2 and 2.3 above.

Now we define the natural transformations η and μ as follows

$$\eta_X = i_X$$

and

$$\mu_X = \mathrm{id}^\dagger_{m(X)}.$$

We need to show that $f \circ i_Y = m(f) \circ i_X$ to show η is natural and that $\mu_Y \circ m^2(f) = m(f) \circ \mu_X$ to show that μ is natural. For the first just note that

$$m(f) \circ i_X = (f \circ i_Y)^\dagger \circ i_X = f \circ i_Y$$

by Equation 2.2. Substituting the definitions into the second equation we obtain

$$(f \circ i_Y)^\dagger \circ (\mathrm{id}_{m(X)})^\dagger = (\mathrm{id}_{m(Y)})^\dagger \circ (i_{m(Y)} \circ (i_Y \circ f)^\dagger)^\dagger$$

and using the equations we obtain

$$L.H.S = ((f \circ i_Y)^\dagger \circ \mathrm{id}_{m(X)})^\dagger = ((f \circ i_Y)^\dagger)^\dagger$$

and

$$R.H.S. = ((\mathrm{id}_{m(Y)})^\dagger \circ i_{m(Y)} \circ (i_Y \circ f)^\dagger)^\dagger = (\mathrm{id}_{m(Y)} \circ (i_Y \circ f)^\dagger)^\dagger = ((f \circ i_Y)^\dagger)^\dagger.$$

Now finally we have to show these natural transformations satisfy the diagrams i.e. obey the following equations,

$$\begin{aligned} \mu_X \circ m(\eta_X) &= \mathrm{id}_{m(X)} \\ \mu_X \circ \eta_{m(X)} &= \mathrm{id}_{m(X)} \\ \mu_X \circ m(\mu_X) &= \mu_X \circ \mu_{m(X)} \end{aligned}$$

Substituting in the first equation we get

$$(\mathrm{id}_{m(X)})^{\dagger} \circ (i_{m(X)} \circ i_X)^{\dagger} = \mathrm{id}_{m(X)}$$

but using the equations we see that

$$L.H.S. = ((\mathrm{id}_{m(X)})^{\dagger} \circ i_{m(X)} \circ i_X)^{\dagger} = (\mathrm{id}_{m(X)} \circ i_X)^{\dagger} = \mathrm{id}_{m(X)}$$

then for the second equation we get

$$(\mathrm{id}_{m(X)})^{\dagger} \circ i_{m(X)} = \mathrm{id}_{m(X)}$$

which is just Equation 2.2 with $f = \mathrm{id}_{m(X)}$. Finally for the third equation substituting we get

$$(\mathrm{id}_{m(X)})^{\dagger} \circ (i_{m(X)} \circ (\mathrm{id}_{m(X)})^{\dagger})^{\dagger} = (\mathrm{id}_{m(X)})^{\dagger} \circ (\mathrm{id}_{m^2(X)})^{\dagger}$$

and then

$$L.H.S. = ((\mathrm{id}_{m(X)})^{\dagger} \circ i_{m(X)} \circ (\mathrm{id}_{m(X)})^{\dagger})^{\dagger} = ((\mathrm{id}_{m(X)} \circ (\mathrm{id}_{m(X)})^{\dagger})^{\dagger} = ((\mathrm{id}_{m(X)})^{\dagger})^{\dagger}$$

and

$$R.H.S. = ((\mathrm{id}_{m(X)})^{\dagger} \circ \mathrm{id}_{m^2(X)})^{\dagger} = ((\mathrm{id}_{m(X)})^{\dagger})^{\dagger}$$

hence we have a monadic functor as required. ∎

2.3 Stone Duality

This section will describe some of the Stone duality type theorems which arose from the Stone representation theorem for Boolean algebras. Starting from this theorem, we show how it extends to a duality of categories which naturally arises as a restriction of an adjunction.

More information on the duality theorems to be described can be found in Johnstone [17].

The Stone representation Theorem for Boolean algebras says that every Boolean algebra can be represented by a field of sets. The proof uses the construction from a Boolean algebra of a topological space consisting of the set of ultrafilters over the algebra, or equivalently the set of Boolean algebra homomorphisms $f\colon B \to 2$ to the two-point lattice $2 = \{0, 1\}$ with a topology given by basic open sets of the form $\{f \mid f(a) = 1\}$ where $a \in B$. Then the clopen subsets of this topological space form a field of sets which represents the original Boolean algebra. This solves the original problem.

The topological spaces that arise in this way can be shown to be the totally disconnected, compact Hausdorff spaces, called Stone spaces [17]. Given any Stone space, taking its clopen field of sets as a Boolean algebra and than taking the space of ultrafilters you also get back to the same space. We can also show that we can get a contravariant functor from the category of Stone spaces to Boolean algebras and vice versa, which gives a contravariant equivalence or duality of categories between **Bool** and **Stone**.

This duality turns out to be a prototype of many other dualities. To get the most general one, we start by looking at the adjunction between the category **Top** of topological spaces and the category **Loc** which is the opposite of **Frm**, the category of complete Heyting algebras (cHa). The lattice of open sets of a topological space is easily seen to be a cHa, and the continuous functions preserve intersections and unions, hence joins and meets under inverse image. This gives a functor **Top** $\to$ **Loc**. To get the adjoint functor, we take the set of frame morphisms of any cHa A to the two point space. This set is equivalent to the set of completely prime filters on A. We give the set a topology in a similar manner to before, open sets being given by

$$U_a = \{f\colon A \to 2 \mid f(a) = 1\}$$

this can be shown to be an adjunction (see Johnstone [17]). The largest equivalence is given by restriction to sober spaces and spatial locales; several other dualities can be obtained by restricting further.

2.4 Interpretation of Stone Duality Theorems

For more details on the ideas discussed here see Abramsky [1].

These duality theorems give the possibility of translating any theorem or proof about the objects on one side into an equivalent theorem or proof about the other. Thus a theorem about certain types of topological spaces has its equivalent theorem about the lattices of open sets.

Dually we could say the denotation of a program is a morphism in the category **Ipo** or a morphism in the dual category of distributive lattices which can be thought of as a predicate transformer. The duality enables us to link the denotational semantics and logic for programs making each one determine the other.

Chapter 3

Measure Theory and Evaluations

This chapter contains a summary of the measure theory which will be used later, together with some simple but non-standard deductions relating mainly to the application of measure theory to ipos with the Scott topology. Then evaluations, which are set functions similar to measures, are defined and a theory of integration and products similar to that for measures is developed.

For further reference on measures see Bollobas [6]; this gives the definitions used in this chapter, also Rudin [33]. For a fuller account see Halmos [16]. Some simple analysis and (even simpler) algebra is assumed. In particular reference is made to the extended real line $[-\infty, +\infty]$, with the topological and algebraic structure as defined in Rudin [33]. For more on evaluations see Pettis [28].

3.1 Motivation

A measure on a set gives the generalised notion of area or size of the subsets of a space. This has the basic axiom of additivity, the area of $A \cup B$ should be equal to the area of A plus the area of B, assuming A and B don't overlap. In fact we mostly use countable additivity to get nice properties of limits, i.e. the area of the union of a countable set of disjoint subsets of a set is equal to the sum of the areas of the subsets.

This idea is related to integration since given areas one can define the integral of a function as the "area under the curve", and given integration, the area of A can be defined as the size of $\int \chi_A$[1]. The well-known Riemann integral from elementary analysis lacks the limit properties that the Lebesgue integral possesses as well as lacking the generality of application possessed by a general theory of measure. The usual route is to first define and develop measures and then to define integration in terms of measures.

Initially mathematicians hoped to define a measure on *all* subsets of $\Re$ which corresponded to our intuitive idea of length—i.e. $[a, b]$ having area $b - a$. Unfortunately it was shown (using the axiom of choice) that no countably additive, translation invariant function could be defined on all subsets of $\Re$. So measure theory was defined so as to give a "size" for only certain certain subsets of a set. We naturally require some closure properties of the collection of "measurable" sets, these are defined in the following section.

---

[1] Here and elsewhere we assume the definition of $\chi_A \colon X \to [0, 1]$ where $A \in X$ to be

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise.} \end{cases}$$

3.2 Fields of subsets

Suppose Σ is a collection of subsets of a set X. Then Σ is a *field* if X is in Σ and Σ is closed under finite unions, and complements (hence also finite intersections). Similarly Σ is a *σ-field* if it contains X and is closed under countable unions and complements (therefore also countable intersections).

It can easily be shown that given some collection of subsets there is a (unique) least field, or σ-field, which contains that collection of sets. Such an object is called is called the field (or σ-field) generated by this collection. It is the intersection of all fields or σ-fields containing the generating set; this intersection is easily seen to have the required properties. In the case of the field generated by some set of subsets, we can define it inductively, but in the case of σ-fields this is not possible (except by using transfinite induction). Thus the usual method of proving that everything in a σ-field generated from Σ has some property is to show that the set of things with this property form a σ-field containing Σ. An example of this is given later.

The *Borel sets* on a topological space are usually defined to be the σ-field generated by all the open sets in the topology.[2]

Suppose P is an ipo with the Scott topology. We will be dealing with the Borel sets of P; they will be denoted $\mathcal{B}(P)$. Observe that $\mathcal{B}(P)$ contains the Scott open sets, Scott closed sets and (recalling the definition in Section 2.1.3) if P is ω-continuous so is A_x, since it is a countable intersection of Scott open sets.

Call a set a *crescent* if it can be expressed as the intersection of two other sets, one closed and the other open. We will consider the field on P which consists of all

[2]Some authors (notably Halmos) define the Borel sets as the σ-ring generated by compact subsets, however the open sets definition is more convenient for our purposes.

sets of P that can be expressed as a finite disjoint union of non-empty crescents. This field we will denote by $\mathcal{F}(P)$, hence

$$A \in \mathcal{F}(P) \text{ iff } A = \bigcup_{i=1}^{n} O_i \cap C_i \tag{3.1}$$

for suitable O_i open and C_i closed in P and such that if $(O_i \cap C_i) \cap (O_j \cap C_j) \neq \emptyset$ then $i = j$

Theorem 3.1 *$\mathcal{F}(P)$ is a field; it is identical to the field generated by the open sets, and the σ-field generated by $\mathcal{F}(P)$ is $\mathcal{B}(P)$.*

Proof The field generated by the open sets consists of all the sets defined by finite expressions involving open sets and the operations $\cup$, $\cap$ and complements. We shall see it is identical to the set of subsets given by Equation 3.1 above; this will also show that $\mathcal{F}(P)$ is a field.

Consider any such expression. By the de Morgan laws and distributivity we can rewrite the expression into one of the form $\bigcup_{i=1}^{n} O_i \cap C_i$ where O_i is open and C_i closed for any $1 \leq i \leq n$ but with the $O_i \cap C_i$ not necessarily pairwise disjoint. But clearly

$$\bigcup_{i=1}^{n} O_i \cap C_i = \bigcup_{i=1}^{n} \left((O_i \cap C_i) \setminus (\bigcup_{j<i} (C_j \cap O_j)) \right)$$

and the terms in the second union are pairwise disjoint unions of crescents (since $(O_i \cap C_i) \cap (\bigcap_{j<i} \tilde{}(C_j \cap O_j)) = O_i \cap C_i \cap (\bigcap_{j<i} (\tilde{}O_j \cap \tilde{}C_j) \cup (\tilde{}O_j \cap C_j) \cup (O_j \cap \tilde{}C_j))$ where the inside unions are disjoint). So the set that any such expression represents can be written as a disjoint union of crescents. So every expression represents a set in $\mathcal{F}(P)$, and since every set in $\mathcal{F}(P)$ is given by such an expression, $\mathcal{F}(P)$ must be the field generated by the open sets of P.

Now consider the σ-field generated by $\mathcal{F}(P)$. It contains all open sets so it must contain $\mathcal{B}(P)$ since $\mathcal{B}(P)$ is the *least* σ-field containing all open sets. Similarly $\mathcal{B}(P)$ also contains $\mathcal{F}(P)$ hence the minimal σ-field containing $\mathcal{F}(P)$ must be contained in it. So the two σ-fields are equal, i.e. $\mathcal{B}(P)$ is also generated by $\mathcal{F}(P)$. ∎

Finally we shall give a standard lemma, known as the Monotone Class Lemma. We define a *monotone class* as a collection of sets closed under nested unions and intersections, that is $\mathcal{G}$ is a monotone class if whenever $A_1 \subseteq A_2 \subseteq \dots$ and A_i is in $\mathcal{G}$, then $\bigcup A_i$ is also in $\mathcal{G}$ and if $B_1 \supseteq B_2 \supseteq \dots$ and B_i is in $\mathcal{G}$, then $\bigcap B_i$ is in $\mathcal{G}$.

Lemma 3.2 (Monotone Class Lemma) *If $\mathcal{G}$ is a field and a monotone class, then it is a σ-field.*

For proof see e.g. [6] page 10. ∎

3.3 Measures

Suppose that Σ is some collection of subsets of a set X. We say $f: \Sigma \to [0, \infty)$ is *countably additive* if whenever we have $A, A_1, A_2, \dots \in \Sigma$ such that

$$\bigcup_{i=1}^{\infty} A_i = A \text{ and } A_i \cap A_j \neq \emptyset \Rightarrow i = j$$

then

$$\sum_{i=1}^{\infty} f(A_i) = f(A).$$

A *measure* is a countably additive set function $\mu: \Sigma \to [0, \infty)$[3] where Σ is a σ-field on some set X. In fact we often say μ is a measure on X where the σ-field is obvious.

For any space X (in fact on any σ-field over a set X) we can define a measure called the "point measure at x" or "unit mass at x" as follows :-

$$\eta_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

[3] Some authors allow a measure to take the value $+\infty$

It it clear that this definition does give a measure.

We can define a finite linear sum of measures as follows

$$(\sum_{i=1}^{n} a_i\mu_i)(A) = \sum_{i=1}^{n} a_i\mu_i(A)$$

where $a_i \geq 0$, and clearly the resulting set function is also a measure.

Thus for any σ-field on a set X we can define any arbitrary linear combination of point measures. Later we will show that for certain types of ipos, these measures generate *all* the possible measures with a simple limit operation.

3.4 Extensions of Measures

Theorem 3.3 *If μ is defined on a field $\mathcal{F}$ and μ is countably additive then μ extends uniquely to a measure $\bar{\mu}$ defined on the σ-field generated by $\mathcal{F}$.*

For the proof see Halmos [16]. ∎

Thus if we can define a countably additive set function on $\mathcal{F}(X)$, we can extend it to one on $\mathcal{B}(X)$, i.e. a measure on $\mathcal{B}(X)$.

3.5 Integration

Integration with respect to a measure μ over a set X is defined for certain real-valued functions on X. We will define integration for functions $f: X \to [0, \infty]$, denoting integration w.r.t. μ by

$$\int f \, d\,\mu.$$

We allow functions to take the value ∞ since then limits become easier to express.

If we have an (positive real-valued) expression e with a free variable x in X and some measure μ on X, we denote the integral of the function $\lambda x.e$ w.r.t. μ by

$$\int_{x \in X} e \, d\mu.$$

A function $f: X \to [0, \infty]$ is said to be *measurable* (w.r.t. some measure on X) if for all open subsets of the real line O, $f^{-1}(O)$ is a measurable set. The collection of measurable functions is closed under various operations, including (positive) linear combinations, pointwise sups and pointwise limits.

There are various ways to define integration, we give an outline of one method.

We define measurable *simple functions* to be measurable functions whose range is a finite set of points in $[0, \infty)$ (i.e. not including infinity); these can clearly be uniquely written in the form

$$s = \sum_{i=1}^{n} a_i \chi_{A_i}$$

where $a_1, \ldots, a_n$ is the range of s and A_i is $s^{-1}(\{a_i\})$. Then we define integration on these functions by

$$\int \sum_{i=1}^{n} a_i \chi_{A_i} \, d\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

Then for positive measurable functions we define

$$\int f \, d\mu = \sup\{\int s \, d\mu \mid s \leq f, s \text{ a simple function}\}$$

If f is itself a step function it is clear that this definition is consistent with the one given above. Note the value of the integral may be $+\infty$, since the set over which we take the sup may be unbounded.

Integration can be shown to have many useful properties. For detailed proofs of the equations to follow see Rudin [33] for example.

$$\int \left(\sum_{i=1}^{n} a_i f_i \right) d\mu = \sum_{i=1}^{n} a_i \int f_i \, d\mu \qquad \text{(Linearity)}$$

$$f(x) \leq g(x) \Rightarrow \int f \, d\mu \leq \int g \, d\mu \qquad \text{(Monotonicity)}$$

$$\forall x. f_n \leq f_{n+1} \Rightarrow \sup_{n \to \infty} \int f_n \, d\mu = \int \sup_{n \to \infty} f_n(x) \, d\mu \qquad \text{(Monotone convergence)}$$

3.6 Product Measures

If Σ and Ψ are two σ-fields on spaces X and Y, then the *product σ-field* is defined to be the σ-field generated by the field of finite, disjoint unions of "rectangles"—sets of the form $A \times B$ where $A \in \Sigma$ and $B \in \Psi$.

We can define the *product measure* of two measures μ and ν defined on Σ and Ψ as follows: $\mu \times \nu$ will be defined on $\Sigma \times \Psi$ and on rectangles it will have the value $(\mu \times \nu)(A \times B) = \mu(A).\mu(B)$. This definition gives a countably additive function on the field of (finite, disjoint unions) of rectangles, hence we can extend it uniquely to give a measure by using Theorem 3.3.

The usual development of this subject is to prove the interchange of integrals theorems.

Theorem 3.4 (Fubini) *If μ and ν are measures on X and Y and f is a measurable function on $X \times Y$, then the functions $\lambda x. \int_{y \in Y} f(x,y) \, d\eta$ and $\lambda y. \int_{x \in X} f(x,y) \, d\mu$ are integrable and*

$$\int_{x \in X} \int_{y \in Y} f(x,y) \, d\eta \, d\mu = \int_{(x,y) \in X \times Y} f(x,y) \, d(\mu \times \eta) = \int_{y \in Y} \int_{x \in X} f(x,y) \, d\mu \, d\eta.$$

For proof see e.g. [33], page 150. ∎

Unfortunately, if we take the Borel sets of two topological spaces X and Y, then the product σ-field of the Borel sets of X and Y is not generally the same as the Borel sets of the product $X \times Y$ with the product topology but is generally smaller. So the product measure of two measures defined on the Borel sets is not a measure defined on the Borel sets of the product but on a smaller σ-field. For ipos the situation is even more complicated, since the Scott topology is not even the same as the product topology. In the next section we will show how it is possible to define a more appropriate product of measures where the spaces concerned are ipos.

3.7 Continuous Measures

In this section we will consider measures defined on ipos in particular. We will impose a further continuity conditions on the measures which will allow us to define a more satisfactory notion of product. We need this extra condition to deal with the fact that ipos are directed complete; if we had instead taken chain completeness, the new definition of product would be possible without the continuity.

We will define a *continuous measure* on an ipo P, to be a measure μ on the Borel sets of P, such that for any directed[4] set of open sets O_i,

$$\mu(\bigcup_{i \in I} O_i) = \sup_{i \in I} \mu(O_i).$$

If the topology of P is countable, e.g. if P is ω-continuous, then all measures are continuous.

With this condition we can prove a version of the Monotone Convergence Theorem, which holds for a directed collection of upper continuous functions.

[4]Here we mean directed with respect to the partial order given by $\subseteq$

Theorem 3.5 (Directed Monotone Convergence) *For any continuous measure μ on an ipo P, if X is a directed set of upper continuous functions (see Section 3.8), with pointwise sup f, then*

$$\sup_{g \in X} \int g \, d\mu = \int f \, d\mu.$$

Proof We note that for any upper continuous function f,

$$f = \sup_n \sum_{i=1}^{n.2^n} 2^{-n} \chi_{f^{-1}(i.2^{-n}, \infty]}$$

and by Monotone Convergence,

$$\int f \, d\mu = \sup_n \int \sum_{i=1}^{n.2^n} 2^{-n} \chi_{f^{-1}(i.2^{-n}, \infty]} \, d\mu = \sum_{i=1}^{n.2^n} 2^{-n} \mu(f^{-1}(i.2^{-n}, \infty]).$$

But if f is the pointwise sup of X, then for any a, $f^{-1}(a, \infty] = \bigcup_{g \in X} g^{-1}(a, \infty]$ and if X is directed, then so is the union on the right hand side of this expression. So by continuity of μ,

$$\mu(f^{-1}(a, \infty]) = \sup_{g \in X} \mu(g^{-1}(a, \infty])$$

so

$$\sup_{g \in X} \int g \, d\mu = \sup_{g \in X} \sup_n \sum_{i=1}^{2^n} 2^{-n} \mu(g^{-1}\left(i.2^{-n}, \infty\right]) = \int f \, d\mu$$

by swapping the sups and sums. ∎

We will use this theorem to help define a product of measures which is defined on the Borel field of the Scott topology of the product of two ipos.

Suppose we have continuous measures μ and η (on P and Q respectively) on the Borel sets of the Scott topology of the product $P \times Q$; then we define a new product, $\mu \otimes \eta$, which extends the usual product $\mu \times \eta$ by the equation

$$\mu \otimes \eta\,(R) = \int_{x \in P} \int_{y \in Q} \chi_R \, d\eta \, d\mu. \tag{3.3}$$

We shall first show that if R is a Borel set from the Scott topology of $P \times Q$, the interated integrals are defined and then that the definition gives a measure.

In general R is not measurable in the usual product measure so we cannot use Fubini's Theorem to show that this definition is symmetric; later we will show that it is symmetric when at least one of P and Q is ω-continuous.

Lemma 3.6 *For any R in the Borel field of the Scott topology on $P \times Q$, the iterated integral on the left of Equation 3.3 is well-defined, and the set function it defines is a continuous measure.*

Proof We introduce the notation $R_x = \{y \mid (x, y) \in R\}$, so R_x is a subset of Q, and observe that the inner integral of Equation 3.3 is equal to $\eta(R_x)$.

Let F be the collection of sets R, such that for any x, R_x is in $\mathcal{B}(Q)$. We shall show that F is a σ-field, and that it contains the Scott open subsets of $P \times Q$, hence it contains the Borel field of the Scott topology on $P \times Q$.

If R is Scott open, then for any x, R_x is clearly open since it is upper-closed and inaccessible by lubs of directed sets. Hence R_x is in $\mathcal{B}(Q)$ so F contains the Scott open sets.

To show that F is closed under intersection, countable unions and intersections we merely observe that:

$$(\tilde{\ } R)_x = \tilde{\ }(R_x)$$

and

$$(\bigcup R^i)_x = \bigcup R^i_x$$

and

$$(\bigcap R^i)_x = \bigcap R^i_x$$

so since $\mathcal{B}(Q)$ is closed under complements and countable intersections and unions so is F.

Now we will repeat this process to show that the function $\lambda x.\eta(R_x)$ which we will denote by f_R is measurable w.r.t. $\mathcal{B}(P)$. Let F' be the collection of sets R

such that f_R is defined and measurable and we will show that F' is a σ-field and that it contains the open sets.

Suppose R is open in the Scott topology of $P \times Q$. Clearly $x_1 \sqsubseteq x_2$ implies $R_{x_1} \subseteq R_{x_2}$, hence f_R is increasing. Further if $x = \sqcup X$ and X is directed, then $\bigcup_{z \in X} R_z = R_x$ since $y \in R_x$ implies $(x, y) \in R$ hence some $z \in X$ satisfies $(z, y) \in R$ since R is open. So by the continuity of μ, f_R is upper continuous, hence measurable.

To show that F' is a σ-field we note that

$$f_{O \cap C} = f_O - f_{O \setminus C}$$

for O open and C closed, and if A and B are disjoint then

$$f_{A \cup B} = f_A + f_B$$

so f_R is certainly measurable on $\mathcal{F}(P \times Q)$, the field generated by the Scott open sets. Then we observe that if $A_1 \subseteq A_2 \subseteq \ldots$ then

$$f_{\bigcup A_i} = \sup_i f_{A_i}$$

and if $B_1 \supseteq B_2 \supseteq \ldots$ then

$$f_{\bigcap B_i} = \inf f_{B_i}$$

and we know that measurability is closed under sups and infs, hence F' is a monotone class. Finally by the Monotone Class Lemma 3.2 we see that F' is a σ-field as required. So again, F' includes the Borel sets of the Scott topology on $P \times Q$, so for all Borel sets in this topology $\mu \otimes \eta(R)$ is well-defined.

It is trivial to see that $\mu \otimes \eta$ is a measure since if $\bigcup A_i = A$ and the A_i are pairwise disjoint, then

$$\sum_{i=1}^{\infty} f_{A_i} = f_A$$

and so

$$\sum_{i=1}^{\infty} \mu \otimes \eta(A_i) = \sum_{i=1}^{\infty} \int f_{A_i}\, d\mu = \int f_A\, d\mu = \mu \otimes \eta(A)$$

as required.

Finally to show the measure $\mu \otimes \eta$ is continuous, we suppose that $\bigcup_{i \in I} O_i$ is a directed union of open sets. Then f_{O_i} is a directed collection of positive upper continuous functions bounded by 1 with sup $f_{\bigcup_{i \in I} O_i}$, and so by the Directed Monotone Convergence theorem 3.5 for continuous measures,

$$\sup_{i \in I} \mu \otimes \eta(O_i) = \sup_{i \in I} \int f_{O_i}\, d\mu = \int f_{\bigcup_{i \in I} O_i}\, d\mu = \mu \otimes \eta(\bigcup_{i \in I} O_i)$$

so $\mu \otimes \eta$ is a continuous measure. ∎

It is easy to see that, like the usual product, the product $\otimes$ is continuous in each variable and bilinear.

Fubini's Theorem shows immediately that $\otimes$ extends the usual measure product. For sets which are measurable by the usual product, that is Borel sets of the product topology, $\otimes$ is symmetric since it agrees with $\mu \times \eta$ which is symmetric.

In the special case where one of the ipos is ω-continuous we can show that the two products are actually the same. We have already shown in Section 2.1.3 that the product topology is the same as the Scott topology when one ipo is continuous, now we shall see that the fact that the topology on one ipo is countably generated is enough to prove that the product of the Borel sets is the Borel sets of the product topology.

Theorem 3.7 *If P and Q are ipos and P is ω-continuous then $\mathcal{B}(P \times Q) = \mathcal{B}(P) \times \mathcal{B}(Q)$.*

Proof It is generally true for any topological spaces that $\mathcal{B}(P)\times\mathcal{B}(Q) \subseteq \mathcal{B}(P\times Q)$. We can prove this by recalling that $\mathcal{B}(P) \times \mathcal{B}(Q)$ is the least σ-field containing the measurable rectangles $A \times B$ where $A \in \mathcal{B}(P)$ and $B \in \mathcal{B}(Q)$. So we need to show that $A \times B$ is in $\mathcal{B}(P \times Q)$. To show this we define

$$K = \{A \mid (A \times Q) \in \mathcal{B}(P \times Q)\}.$$

It is clear that if O is open in P, then $O \in K$. It is also clear that K is a σ-field—it is closed under countable unions etc. since $\mathcal{B}(P \times Q)$ is. Hence if $A \in \mathcal{B}(P)$, then $A \in K$. Defining a similar set L with $B \in L \iff (P \times B) \in \mathcal{B}(P \times Q)$ we see that if $B \in \mathcal{B}(Q)$, then $B \in L$. But $(A \times Q) \cap (P \times B) = (A \times B)$ so $A \times B \in \mathcal{B}(P \times Q)$.

Note that the product topology on $P \times Q$ is given by countable unions of sets of the form $V_b \times O$ for b in the basis of P and O open in Q.

Now to show that $\mathcal{B}(P \times Q) \subseteq \mathcal{B}(P) \times \mathcal{B}(Q)$ we need to show that $\mathcal{B}(P) \times \mathcal{B}(Q)$ contains all the open sets in the product topology $P \times Q$. This is not generally true as we may take arbitrary unions of open rectangles in the product topology but only countable ones in forming the product of σ-fields. But in this case the open sets in the product topology are countable unions of sets of the form $V_b \times O$ hence in $\mathcal{B}(P) \times \mathcal{B}(Q)$.

Hence $\mathcal{B}(P) \times \mathcal{B}(Q) = \mathcal{B}(P \times Q)$ as required. ∎

So provided that P is ω-continuous, for any continuous measures μ and η,

$$\mu \times \eta = \mu \otimes \eta$$

since we know from the theorem above and Section 2.1.3 that the two products are defined on the same σ-field and by Fubini's theorem for any R in this σ-field,

$$(\mu \times \eta)(R) = \int_{(x,y) \in P \times Q} \chi_R \, d(\mu \times \eta) = \int_{x \in P} \int_{y \in Q} \chi_R \, d\eta \, d\mu = (\mu \otimes \eta)(R).$$

3.8 Upper Continuous Functions

A function f from X to $[0, \infty]$ is *upper-continuous* (u.c.) if for any $a \geq 0$, the set $f^{-1}(a, \infty]$ is open in X.

Note that an open step function (χ_O for any open set O) is always upper continuous since the inverse image is always either X or O, both of which are open. We will now show the set of upper continuous (u.c.) functions is closed under linear combinations, infinite sups, finite lubs and that it forms a complete Heyting algebra (cHa) with these pointwise operations. In fact we do this for upper continuous functions $f: X \to [0,1]$, i.e. functions bounded by 1 since this is the cHa we will consider in later chapters.

Theorem 3.8 *If X is a topological space then the set of upper continuous function $f: X \to [0,1]$ is a cHa.*

Proof We define the sups and infs pointwise that is

$$(f \wedge g)(x) = \min(f(x), g(x))$$

and

$$(\bigvee_{\gamma \in \Gamma} f_\gamma)(x) = \sup_{\gamma \in \Gamma} f_\gamma(x).$$

To see that these definitions give upper continuous functions we simply note

$$(f \wedge g)^{-1}(a, 1] = f^{-1}(a, 1] \cap g^{-1}(a, 1]$$

so if f and g are upper continuous, then the inverse image of an upper open set under $f \wedge g$ is the intersection of two open sets and thus open. Similarly for any $f_\gamma, \gamma \in \Gamma$, the sup is upper-continuous because

$$(\bigvee_{\gamma \in \Gamma} f_\gamma)^{-1}(a, 1] = \bigcup_{\gamma \in \Gamma} (f_\gamma)^{-1}(a, 1]$$

i.e. the inverse image is the union of open sets.

To see that the sups commute correctly is easy since they are all pointwise. We only need to show that for any value x,

$$\bigvee_{\gamma \in \Gamma} (f_\gamma(x) \wedge g(x)) = (\bigvee_{\gamma \in \Gamma} f_\gamma(x)) \wedge g(x)$$

i.e.

$$\sup_{\gamma \in \Gamma} \min(a(\gamma), b) = \min(\sup_{\gamma \in \Gamma} a(\gamma), b)$$

and this is well known to be true by elementary analysis. ∎

To show linear combinations of upper continuous functions are also upper continuous we observe that for multiples of functions

$$(r.f)^{-1}(a,\infty] = f^{-1}(a/r,\infty]$$

and for sums

$$(f+g)^{-1}(a,\infty] = \bigcup_{r\leq a}(f^{-1}(r,\infty] \cap g^{-1}(a-r,\infty])$$

so any linear combination of upper continuous functions is clearly upper continuous.

Finally every upper continuous function f can be written as a sup of increasing linear combinations of step functions, since

$$f = \lim_{n\to\infty} \sum_{i=1}^{n.2^n} 2^{-n}\chi_{f^{-1}(i.2^{-n},\infty]}.$$

If f is bounded then this limit is actually uniform, that is for all $\epsilon > 0$, there exists an N such that for all x and for all $n \geq N$,

$$f(x) - \sum_{i=1}^{n.2^n} 2^{-n}\chi_{f^{-1}(i.2^{-n},\infty]}(x) < \epsilon$$

viz N sufficiently large so that $2^{-N} < \epsilon$ and f is bounded by N.

3.9 Evaluations and Integration

In this section we will see that an simpler class of set functions than measures, called evaluations, can be used to define a theory of integration. The motivation for this comes from noting that we can define the integral of an upper continuous function w.r.t. a measure using only the values of the measure on open sets, by expressing the function as a pointwise limit of linear combinations of step functions. The integral we shall define will be linear and will satisfy a slightly restricted monotone convergence theorem, but will only be defined on upper continuous functions.

Evaluations are usually defined on a lattice. To facilitate the analogy between measures and evaluations we will define evaluations on the lattice of open sets of a topological space, inspired by Lawson's work [23].

We define an *evaluation* on a topological space Y to be a map ν from $\Omega(Y)$ to $[0, \infty)$ which satisfies

$$\nu(U) + \nu(V) = \nu(U \cup V) + \nu(U \cap V) \tag{3.4}$$

and

$$\nu(\emptyset) = 0 \tag{3.5}$$

(modularity) and

$$U \subseteq V \Rightarrow \nu(U) \leq \nu(V) \tag{3.6}$$

(monotonicity) and a *continuous evaluation* to be an evaluation for which if X is a directed subset of $\Omega(Y)$ (directed w.r.t. the partial order on $\Omega(Y)$ given by $\subseteq$) then

$$\sup\{\nu(O) \mid O \in X\} = \nu(\bigcup_{O \in X} O) \tag{3.7}$$

In [28], Pettis investigates the general problem of when an evaluation may be uniquely extended to a measure. He proves the theorem that any evaluation on

Y can be uniquely extended to a finitely additive set function on the field, $\mathcal{F}(Y)$, generated by the open sets in Y. As previously discussed, this field consists of all sets R that can be written as a disjoint union

$$R = \bigcup_{i=1}^{n} (O_i \cap C_i)$$

for closed sets C_i and open sets O_i. The unique extension of an evaluation ν must clearly give R the value

$$\sum_{i=1}^{n} \nu(O_i) - \nu(O_i \setminus C_i).$$

We will now define integration with respect to a continuous evaluation ν. In particular we want to show that integration is linear and that a version of the monotone convergence theorem holds. To simplify the definitions we shall only consider bounded upper continuous functions.

We first consider integration of the simple upper continuous functions—those which take a finite number of values. Let s be a simple upper continuous function and suppose it takes the values $\alpha_1, \ldots, \alpha_n$. Then as before we can write s uniquely as a linear combination of characteristic functions

$$s = \sum_{i=1}^{n} \alpha_i \chi_{s^{-1}(\{\alpha_i\})}$$

moreover $s^{-1}(\{\alpha_i\})$ is in the field generated by the open sets $\mathcal{F}(Y)$, since it is equal to

$$s^{-1}(\alpha_i - \epsilon, \infty] \setminus s^{-1}(\alpha_i, \infty]$$

for sufficiently small ϵ.

So we define

$$\int s \, d\nu = \sum_{i=1}^{n} \alpha_i \nu(s^{-1}(\{\alpha_i\}))$$

using the unique extension of ν to $\mathcal{F}(Y)$.

We will now give some properties of this definition.

Lemma 3.9 *Let B_j for $j = 1, \ldots, m$ be such that $\bigcup_{i=1}^m B_j = Y$ and $B_j \cap B_k \neq \emptyset$ implies $j = k$ (so B_j form a partition of Y) and B_j is in $\mathcal{F}(Y)$ for all j. Then for any simple u.c. function s taking values α_i and any evaluation ν,*

$$\int s\, d\nu = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \nu(s^{-1}(\alpha_i) \cap B_j).$$

Proof We defined

$$\int s\, d\nu = \sum_{i=1}^{n} \alpha_i \nu(s^{-1}(\alpha_i))$$

so it remains to show that $\sum_{j=1}^m \nu(s^{-1}(\alpha_i) \cap B_j) = \nu(s^{-1}(\alpha_i))$. But this follows from the finite additivity of the extension of ν, since if $A = s^{-1}(\alpha_i)$, as B_j partition Y, $\bigcup_{j=1}^m (B_j \cap A) = A$ and they are pairwise disjoint. ∎

We use the lemma above to show that the definition above is monotonic.

Lemma 3.10 *If s and t are two simple upper continuous functions and $s \leq t$ (pointwise) then $\int s\, d\nu \leq \int t\, d\nu$.*

Proof Suppose the values taken by s and t are $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_m$ respectively. We note that the sets $s^{-1}(\{\alpha_i\})$ and the sets $t^{-1}(\{\beta_j\})$ form two partitions of sets in $\mathcal{F}(Y)$ as required in Lemma 3.9 above.

By applying the lemma to s with partition $t^{-1}(\beta_j)$ and to t with partition $s^{-1}(\alpha_i)$ we see that

$$\int s\, d\nu = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \nu(s^{-1}(\alpha_i) \cap t^{-1}(\beta_j))$$

and

$$\int t\, d\nu = \sum_{j=1}^{m} \sum_{i=1}^{n} \beta_i \nu(t^{-1}(\beta_j) \cap s^{-1}(\alpha_i))$$

but since $t \leq s$, if $s^{-1}(\alpha_i) \cap t^{-1}(\beta_j) \neq \emptyset$, then $\beta_j \leq \alpha_i$ so

$$\int t\, d\nu \leq \int s\, d\nu$$

as required. ∎

Now we shall show that the definition is linear.

Lemma 3.11 *If s and t are simple upper continuous functions, then so is $s+t$ and*

$$\int s + t \, d\nu = \int s \, d\nu + \int t \, d\nu.$$

Proof Suppose as before that s and t take values α_i and β_j. Then $s+t$ takes the values $\alpha_i + \beta_j$ on sets $s^{-1}(\alpha_i) \cap t^{-1}(\beta_j)$, so clearly $s+t$ is simple and we have already seen that the sum of upper continuous functions is upper continuous. This representation of $s+t$ is not necessarily the unique one since there may be distinct pairs (i_1, j_1), (i_2, j_2) such that $\alpha_{i_1} + \beta_{j_1} = \alpha_{i_2} + \beta_{j_2}$. But if we suppose that $s+t$ takes the distinct values γ_k, then for each k, there is a non-empty finite set of pairs (i,j) such that $\alpha_i + \beta_j = \gamma_k$ and $(s+t)^{-1}(\gamma_k) = \bigcup_{(i,j) \text{s.t. } \alpha_i + \beta_j = \gamma_k} s^{-1}(\alpha_i) \cap t^{-1}(\beta_j)$. So by this and finite additivity since this union is disjoint,

$$\int (s+t) \, d\nu = \sum_i \sum_j (\alpha_i + \beta_j).\nu(s^{-1}(\alpha_i) \cap t^{-1}(\beta_j)).$$

But then by collecting the terms by coefficient and using finite additivity,

$$\int (s+t) \, d\nu == \sum_i \alpha_i \nu(s^{-1}(\alpha_i) \cap (\bigcup t^{-1}(\beta_j))) + \sum_j \beta_j \nu(t^{-1}(\beta_j) \cap (\bigcup s^{-1}(\alpha_i)))$$

and since clearly $\bigcup t^{-1}(\beta_j) = Y$ and similarly $\bigcup s^{-1}(\alpha_i) = Y$,

$$\int (s+t) \, d\nu = \int s \, d\nu + \int t \, d\nu$$

as required. ∎

So we define for any bounded upper continuous function f

$$\int f \, d\nu = \sup\{ \int s \, d\nu \mid s \leq f, s \text{ simple} \}.$$

Clearly if f happens to be simple, this definition agrees with the previous one. Also if $f \leq g$ then clearly $\int f \leq \int g$, so the definition is monotonic.

We now want to prove bounded monotone convergence and linearity. To do these we start by proving a lemma relating to uniform convergence.

Lemma 3.12 *Suppose that two functions are "close" everywhere e.g. suppose for some $\epsilon > 0$ we have $|f(x) - g(x)| < \epsilon$ for all x in X. Then*

$$|\int f\,d\nu - \int g\,d\nu| \leq \epsilon.\nu(X).$$

Proof If s is a simple function with $s \leq f$, then the simple function $t(x) = \max(s(x) - \epsilon, 0)$ clearly satisfies $t \leq g$ and $\int t\,d\nu \geq \int s\,d\nu - \epsilon.\nu(X)$. So

$$\sup_{s\leq f}(\int s\,d\nu) - \epsilon.\nu(X) \leq \sup_{t\leq g}(\int t\,d\nu)$$

thus

$$\int f\,d\nu - \epsilon.\nu(X) \leq \int g\,d\nu$$

and we can clearly interchange the role of f and g to show that

$$|\int f\,d\nu - \int g\,d\nu| \leq \epsilon.\nu(X)$$

as required. ∎

So if f_n is a sequence of bounded upper continuous functions which tend to some upper continuous f uniformly it is clear that

$$\lim_n \int f_n\,d\nu = \int f\,d\nu.$$

But for any bounded upper continuous f, recall that we may express f as a uniform limit of simple functions namely

$$f = \sup_{n\to\infty} \sum_{i=1}^{n2^n} 2^{-n} \chi_{f^{-1}(i.2^{-n},\infty]}$$

thus we know that

$$\int f\,d\nu = \sup \sum_{i=1}^{n2^n} 2^{-n}\nu(f^{-1}\left(i.2^{-n},\infty\right]) \tag{3.8}$$

We are now able to prove a directed version of the bounded monotone convergence theorem.

Theorem 3.13 (Directed Monotone Convergence) *If X is a directed set of u.c. functions with pointwise sup a bounded function f, then*

$$\sup_{g\in X} \int g\,d\nu = \int f\,d\nu.$$

Proof From Equation 3.8 above

$$\sup_{g\in X} \int g\,d\nu = \sup_{g\in X} \sup_{n} \sum_{i=1}^{n2^n} 2^{-n}\nu(g^{-1}\left(i.2^{-n},\infty\right]).$$

We can interchange the sups on the right hand side of the equation above, then we can interchange the sup and the finite sum since g is directed, but then clearly $\bigcup_{g\in X} g^{-1}(a,\infty] = f^{-1}(a,\infty]$ since f is the pointwise sup of $g \in X$. Hence by the continuity of ν we know $\sup_{g\in X}(\nu(g^{-1}(a,\infty])) = \nu(f^{-1}(a,\infty])$. So we see that

$$\sup_{g\in X} \int g\,d\nu = \lim_{n\to\infty} \sum_{i=1}^{n2^n} \nu(f^{-1}(a,\infty]) = \int f\,d\nu$$

by Equation 3.8 again, as required. ∎

Finally we consider linearity. For any positive real number r, it is clear that $\int rf\,d\nu = r\int f\,d\nu$ since if $s \leq f$, then $rs \leq rf$ and similarly $t \leq rf$ implies $t/r \leq f$. So we just need to show that $\int (f+g)\,d\nu = \int f\,d\nu + \int g\,d\nu$.

Theorem 3.14 *For any bounded upper continuous functions f and g*

$$\int (f+g)\,d\nu = \int f\,d\nu + \int g\,d\nu.$$

Proof Express both f and g as the sup of increasing sequences of simple functions s_n and t_n, say. Then clearly $(t_n + s_n)$ is an increasing sequence of simple functions whose sup is $f + g$. Then since integration of simple functions is linear and by the directed monotone convergence theorem, we have

$$\int (f+g)\, d\nu = \sup_n \int s_n + t_n \, d\nu = \sup_n \int s_n \, d\nu + \sup_n \int t_n \, d\nu = \int f \, d\nu + \int g \, d\nu$$

as required. ∎

We note that if μ is a Borel measure which also satisfies the continuity condition for open sets, then integration with respect to the corresponding continuous evaluation obtained by restricting μ to open sets agrees with the usual form of integration with respect to μ (since they agree on χ_O for any open set O and both are linear and satisfy monotone convergence).

3.10 Products of Evaluations

In this section we will consider how to define products of evaluations. As with measures, we shall define two products, one on the product topology of two topological spaces (the equivalent of the usual product on measures) and its extension to the Scott topology on the product of two ipos. We will get a version of Fubini's theorem for the first product and we will show that the products coincide when one of the ipos is continuous.

Suppose that η and ν are continuous evaluations on topological spaces Y and Z. We extend η and ν to give finitely additive set functions on the fields $\mathcal{F}(Y)$ and $\mathcal{F}(Z)$ and then by Theorem 6.2 in [6] we get the unique finitely additive set function on the field generated by sets of the form $A \times B$ where $A \in \mathcal{F}(Y)$ and $B \in \mathcal{F}(Z)$ which takes the value $\eta(A) \times \nu(B)$ on sets of form $A \times B$. This field consists of all finite disjoint unions of products of crescents, and the σ-field it generates is *not* the Borel sets of $X \times Y$, since in forming the product topology we can take arbitrary unions of open rectangles whereas in forming a σ-field we can only take countable unions. However we can use the following general theorem about extending evaluations from the basis of a topology.

Theorem 3.15 *Suppose Σ is some collection of open subsets of Y, such that it is it is closed under finite intersections and unions and every open set is an arbitrary union of sets in Σ, and ν is a modular, continuous set function on Σ (that is $\nu(\emptyset) = 0$ and $\nu(A) + \nu(B) = \nu(A \cap B) + \nu(A \cup B)$ and whenever A_i is a directed collection of sets in Σ whose union is also in Σ then $\nu(\bigcup_{i \in I} A_i) = \sup_{i \in I} \nu(A_i)$) then ν extends uniquely to a continuous evaluation on Y.*

Proof We will define $\bar{\nu}$ on any open set by

$$\bar{\nu}(O) = \sup_{A \subseteq O, A \in \Sigma} \nu(A).$$

Since ν is monotonic (by the continuity condition) it is clear that $\bar{\nu}$ extends ν. This extension is unique since the equation above must hold for any continuous evaluation which extends ν. Now suppose we have some directed collection of sets A_i such that A_i is in Σ for each $i \in I$. The union $\bigcup_{i \in I} A_i$ is an open set, say O. We shall show that

$$\sup_{i \in I} \nu(A_i) = \bar{\nu}(O).$$

It is clear that $\sup_{i \in I} \nu(A_i) \leq \bar{\nu}(O)$ from the definition of $\bar{\nu}$. To show the converse we pick $\epsilon > 0$ and find some $A \in \Sigma$ such that $A \subseteq O$ and $\nu(A) + \epsilon/2 > \bar{\nu}(O)$. Then it is clear that $\bigcup_{i \in I}(A \cap A_i) = A$ and this union is directed, hence by the continuity of ν, $\nu(A) = \sup_{i \in I} \nu(A \cap A_i)$, so there exists $i \in I$ such that $\nu(A \cap A_i) + \epsilon/2 > \nu(A)$, and then since $A_i \cap A \subseteq A$ it is clear that $\nu(A_i) + \epsilon > \bar{\nu}(O)$ as required.

Given this fact we can easily show that $\bar{\nu}$ is a continuous evaluation. For modularity, we note that $\bar{\nu}(\emptyset) = \nu(\emptyset) = 0$ and given open sets U and V, then say $U = \bigcup_{i \in I} A_i$ and $V = \bigcup_{j \in J} B_j$ for A_i, B_j in Σ and both unions directed, then

$\bigcup_{(i,j)\in I\times J}(A_i \cap B_j) = U \cap V$ and $\bigcup_{(i,j)\in I\times J}(A_i \cup B_j) = U \cup V$ and both these unions are directed. Hence

$$\bar{\nu}(U)+\bar{\nu}(V) = \sup_{i,j}(\nu(A_i)+\nu(B_j)) = \sup_{i,j}(\nu(A_i\cap B_j)+\nu(A_i\cup B_j)) = \bar{\nu}(U\cap V)+\bar{\nu}(U\cup V)$$

Suppose that $\bigcup_{i\in I} O_i = O$ and this union is directed. Let $K = \{A \mid A \in \Sigma, A \subseteq O_i \text{ for some } i \in I\}$. Then K is directed (since the set O_i for $i \in I$ is directed) so by the above,

$$\bar{\nu}(O) = \sup_{A\in K} \nu(A)$$

since $\bigcup_{A\in K} A = O$. But by definition

$$\bar{\nu}(O_i) = \sup_{A\in\Sigma, A\subseteq O_i} \nu(A),$$

so clearly

$$\sup_{i\in I} \bar{\nu}(O_i) = \sup_{A\in K} \nu(A) = \bar{\nu}(O)$$

So we see that $\bar{\nu}$ is continuous. ∎

This theorem is similar to Theorem 3.3 (extensions of measures). We will restrict the finitely additive function on the field generated by products of crescents to a function on finite unions of open rectangles, and then apply the theorem above to get a continuous evaluation. We can apply the theorem since finite unions of open rectangles are closed under finite unions and intersections and since every open set in the product topology is a union of finite rectangles. It remains to show that the finitely additive set function restricts to a modular continuous function on the finite unions of open rectangles. Modularity comes from finite additivity, we need to do a little work to show that the set function is continuous on finite unions of open rectangles.

Lemma 3.16 *If η and ν are continuous evaluations and s is the unique finitely additive set function on the field generated by products of crescents, then if $U \times V$ is an open rectangle and $U \times V = \bigcup_{i\in I}(U_i \times V_i)$ where $U_i \times V_i$ are open rectangles for each i, then $s(U \times V) = \sup_{i_1,\ldots,i_n} s((U_{i_1} \times V_{i_1}) \cup \ldots \cup (U_{i_n} \times V_{i_n}))$.*

Pick $\epsilon > 0$. First consider any $x \in U$. For all $y \in V$, there exists i such that $(x, y) \in U_i \times V_i$ hence $\bigcup_{i \text{ s.t. } x \in U_i} V_i = V$. So by the continuity of ν we can find finitely many i, say $i_1^x, \ldots i_{n_x}^x$ such that $x \in U_{i_j^x}$ for each $1 \leq j \leq n_x$ and $\nu(\bigcup_{j=1}^{n_x} V_{i_j^x}) + \epsilon > \nu(V)$. Set $U_x = \bigcap_{j=1}^{n_x} U_{i_j^x}$, so $x \in U_x$. Note also that the finite union $\bigcup_{j=1}^{n_x}(U_{i_j^x} \times V_{i_j^x})$ contains $U_x \times (\bigcup_{j=1}^{n_x} V_{i_j^x})$, hence has s value at least $\eta(U_x)(\nu(V) - \epsilon)$. But also $\bigcup_{x \in U} U_x = U$ so by continuity of η, there exist $x_1, \ldots x_m$ such that $\eta(\bigcup_{k=1}^m U_{x_k}) + \epsilon > \eta(U)$. So now if we consider the finite union $\bigcup_{k=1}^m \bigcup_{j=1}^{n_{x_k}} i_j^{x_k}$ then the corresponding union of open rectangles has s value at least $(\eta(U) - \epsilon)(\nu(V) - \epsilon)$. So as ϵ was arbitrary, we see that $s(U \times V) = \sup_{i_1, \ldots, i_n} s((U_{i_1} \times V_{i_1}) \cup \ldots \cup (U_{i_n} \times V_{i_n}))$ as required. ∎

We can use this lemma to show that the set function s as defined above is continuous in the sense required by Theorem 3.15. We want to apply the theorem to the case where Σ is the collection of sets which are finite, disjoint unions of open rectangles, so we need to show that $s(O) = \sup_{i \in I} s(O_i)$ where $O = \bigcup_{i \in I} O_i$ where this is a directed union and $O = \bigcup_{j=1}^n (U_j \times V_j)$ and $O_i = \bigcup_{j=1}^{n_i} (U_j^i \times V_j^i)$. By finite additivity we can assume that $n = 1$, so O is just an open rectangle $O = U \times V$. Then applying the theorem above to the sets $U_j^i \times V_j^i$ (for all $i \in I$ and $j \in \{1, \ldots, n_i\}$), we see that $s(O)$ is equal to the sup of s over all finite unions of the sets $U_j^i \times V_j^i$. But any finite union of the sets $U_j^i \times V_j^i$ must be contained in some O_i since O_i is directed, so $s(O) = \sup_{i \in I} s(O_i)$ as required.

So given any continuous evaluations η and ν we can form the continuous evaluation $\eta \times \nu$ on the product topology $X \times Y$. We now prove a counterpart to Fubini's Theorem for products of evaluations.

Theorem 3.17 (Fubini for evaluations) *If X and Y are topological spaces and f is upper-continuous in the product topology of X and Y, then for any η and ν continuous evaluations on X and Y respectively, the iterated integrals are defined* *and satisfy*

$$\int_{x\in X}\int_{y\in Y} f(x,y)\,d\nu\,d\eta = \int f\,d(\nu\times\eta) = \int_{y\in Y}\int_{x\in X} f(x,y)\,d\eta\,d\nu.$$

Proof We first assume that f is the characteristic function of an open set in $X\times Y$, say $f=\chi_O$. Suppose that O is an arbitrary union of open rectangles, say $O=\bigcup_{i\in I}(V_i\times U_i)$. Then the functions $\lambda x.\int_{y\in Y} f(x,y)\,d\nu$ and $\lambda y.\int_{x\in X} f(x,y)\,d\eta$ are upper continuous since the first can be seen to be equal to

$$\sup\{\nu(\bigcup_{i\in K} U_i)\chi_{\bigcap_{i\in K} V_i} \mid \text{finite subsets } K \text{ of } I\}$$

and similarly for the second. Then further since any function f is the sup of an increasing sequence of sums of step functions, the iterated integrals exists since sums and sups of upper continuous functions are upper continuous. So the iterated integrals exist.

It is clear that if f is the characteristic function of an open rectangle $U\times V$, then all three expressions give the value $\eta(U)\times\nu(V)$ since then the interated integrals become $\int \eta(U)\chi_V\,d\nu$ and $\int \eta(V)\chi_U\,d\eta$ respectively. But each interated integral defines a modular, continuous evaluation on open sets, by $f=\chi_O$. To see this we use the notation $O_x=\{y \mid (x,y)\in O\}$ and $f_O=\lambda x.\nu(O_x)$ as in Lemma 3.6, then

$$\int_{x\in X}\int_{y\in Y}\chi_\emptyset\,d\nu\,d\eta = \int_{x\in X}\nu(\emptyset)\,d\eta = 0$$

and

$$\int_{x\in X}\int_{y\in Y}\chi_U\,d\nu\,d\eta + \int_{x\in X}\int_{y\in Y}\chi_V\,d\nu\,d\eta = \int_{x\in X} f_U + f_V\,d\eta$$

but $f_U+f_V = f_{U\cap V}+f_{U\cup V}$ since for any x, $\nu((U\cap V)_x)+\nu((U\cup V)_x)=\nu(U_x\cap V_x)+\nu(U_x\cup V_x)=\nu(U_x)+\nu(V_x)$ by the modularity of ν, hence by linearity of integration,

$$\int_{x\in X} f_{U\cap V}+f_{U\cup V}\,d\eta = \int_{x\in X} f_{U\cap V}\,d\eta + \int_{x\in X} f_{U\cup V}\,d\eta$$
$$= \int_{x\in X}\int_{y\in Y}\chi_{U\cap V}\,d\nu\,d\eta + \int_{x\in X}\int_{y\in Y}\chi_{U\cup V}\,d\nu\,d\eta$$

For continuity we suppose that $\bigcup_{i \in I} O^i = O$, this being a directed union. Then clearly for any x, $\bigcup_{i \in I} O^i_x = O_x$ and this union is directed, hence $\sup_{i \in I} f_{O^i} = f_O$ by the continuity of ν and this sup is also directed. Then

$$\sup_{i \in I} \int_{x \in X} \int_{y \in Y} \chi_{O^i} \, d\nu \, d\eta = \sup_{i \in I} \int_{x \in X} f_{O^i} \, d\eta$$

and by directed monotone convergence,

$$\sup_{i \in I} \int_{x \in X} f_{O^i} \, d\eta = \int_{x \in X} f_O \, d\eta = \int_{x \in X} \int_{y \in Y} \chi_O \, d\nu \, d\eta$$

as required. We can similarly show that the integral

$$\int_{y \in Y} \int_{x \in X} \chi_O \, d\eta \, d\nu$$

gives a modular and continuous evaluation.

But these three continuous evaluations must all be equal by the uniqueness of the definition of product evaluations. So the theorem holds for f equal to χ_O for any open set O.

Then since any continuous function f on $X \times Y$ is the limit of linear combinations of step functions, by linearity and continuity of integration, it is clear that

$$\int_{x \in X} \int_{y \in Y} f(x, y) \, d\eta \, d\mu = \int_{y \in Y} \int_{x \in X} f(x, y) \, d\mu \, d\eta$$

as required. ∎

Now we consider the case of evaluations η and ν defined on ipos P and Q. We let O be open in the Scott topology of $P \times Q$ and consider the equation

$$\eta \otimes \nu(O) = \int_{x \in P} \int_{y \in Q} \chi_O(x, y) \, d\nu \, d\eta$$

Since O is open in the Scott topology of $P \times Q$, the inner integral is defined since the x cross-section of O is open in Q. Similarly the function which gives the measure according to ν of the x cross-section of O is upper continuous since it is continuous as a function on ipos. (This was shown in Lemma 3.6.)

Further this definition gives an evaluation since (exactly as before)

$$(\eta \otimes \nu)(\emptyset) = \int_{x \in P} \nu(\emptyset)\, d\eta = 0$$

and

$$(\eta \otimes \nu)(U) + (\eta \otimes \nu)(V) = \int_{x \in P} \nu(U_x) + \nu(V_x)\, d\eta$$

$$= \int_{x \in P} \nu((U \cap V)_x) + \nu((U \cup V)_x)\, d\eta$$

$$= (\eta \otimes \nu)(U \cap V) + (\eta \otimes \nu)(U \cup V)$$

gives modularity.

Now suppose $\bigcup_{i \in I} O^i = O$ and this union is directed, all sets being open in $P \times Q$. Then O^i_x over $i \in I$ is a directed set of open sets in Q with sup O_x so

$$\sup_{i \in I} (\eta \otimes \nu)(O^i) = \sup_{i \in I} \int_{x \in P} \nu(O^i_x)\, d\eta$$

but similarly $f_i = \lambda x.\nu(O^i_x) = f_{O^i}$ is a directed set of functions since if $O^i \cup O^j \subseteq O^k$, then for any x, $f_i(x) \leq f_k(x)$ and $f_j(x) \leq f_k(x)$ so f_k is an upper bound for f_i and f_j, and f_i over $i \in I$ has pointwise lub given by $\lambda x.\nu(O_x)$ hence

$$\sup_{i \in I} \int_{x \in P} f_i(x)\, d\eta = \int_{x \in P} \nu(O_x) = (\eta \otimes \nu)(O)$$

so we have continuity.

Note also that by the continuity and linearity of integration, the operation $\otimes$ is continuous in each variable and bilinear.

By Fubini's theorem $\eta \otimes \nu$ extends the product we defined above, and it is symmetric (again by Fubini's theorem) on sets which are open in the product topology. If one of P and Q is continuous, then the Scott topology on the product of the ipos is the same as the product topology and the two products are identical.

3.11 Concluding Remarks

In this section we have defined measures and integration of measurable functions with respect to them, and products of measures and Fubini's theorem on interchanging the of order of integration. Similarly we have defined evaluations and integration of upper continuous functions with respect to them and also products of evaluations and a version of Fubini's theorem for them. Since only bounded upper continuous functions are used in the rest of this thesis, the definition of integration for evaluations is restricted to them.

The first product of evaluations—defined on the product of the topological spaces, is analogous to the usual product of measures. It uses the continuity of evaluations to give a measure to an arbitrary union of open rectangles. If we took two continuous measures on topological spaces, restricted them to evaluations, and took their product, the result would not generally extend to a measure. However if we took two continuous measures η and ν defined on ipos, the product $\eta \otimes \nu$ (on the Borel sets of the Scott topology of the ipos) would agree on open set with the $\otimes$ product of the restrictions of η and μ to open sets (by the fact that integration of evaluations is agrees with the usual integration). It is not clear whether the $\otimes$ products are symmetric in general.

Chapter 4

The Probabilistic Powerdomain

Given a probabilistic computation over a set of results the naive approach to expressing the result might be to assign a probability to each of the possible outcomes. Probability theorists have long known that this approach is not sufficient if the set of outcomes is uncountable, e.g. in problems like picking a real number at random from $[0, 1]$, since in this case any particular number has probability zero of occurring and so we cannot answer questions like "what is the probability that the random number is greater than half ?". Instead they consider probability distributions which assign a probability to certain sets of results—in the $[0, 1]$ example, probabilities are assigned by to all Lebesgue measurable subsets, with $[a, b]$ having probability $b - a$. As was observed by Saheb-Djahromi [35], this problem occurs with probabilistic computation as well; an example is given of the program which prints infinite strings of zeros and ones, at each stage printing a zero with probability $1/2$ and a one with probability $1/2$. Then the probability of any particular string is zero, but we still want to be able to calculate things like the probability of getting a string beginning with 11 for instance. So we must represent the outcome of a probabilistic computation by a function assigning probabilities to certain subsets of the set of results.

In this thesis we decide to represent probabilistic computations by continuous evaluations. Continuous evaluations give, for every open set, the probability that the result of a probabilistic computation is in this set. Following Smyth [38] we can regard open sets as being the possible tests or properties which the result of a computation may satisfy. So we can interpret the axioms for continuous evaluations as necessary properties of tests and justify them as follows. We take the axioms from Lawson [23] (or Birkhoff [5, Chapters 10,11]) as given in Section 3.9.

With open sets representing properties or tests, clearly the empty set corresponds to a test that always fails, or the property that is never satisfied. Then a probabilistic computation will have probability zero of satisfying this test, hence a function on open sets representing the computation should satisfy Equation 3.5 (that $\nu(\emptyset) = 0$). Similarly if one open set U is contained in another V, then this implies the property that U represents implies the property that V represents. Hence a function representing a probabilistic computation should satisfy Equation 3.6 (that if $U \subseteq V$ then $\nu(U) \leq \nu(V)$). Similarly the function should be modular (Equation 3.4, that $\nu(U) + \nu(V) = \nu(U \cup V) + \nu(U \cap V)$) since the intersection and union of two open sets represents combining the properties with "and" and "or" respectively and so this reduces to the axiom of probability theory that the probability of A or B is equal to the probability of A plus the probability of B minus the probability of A & B. The continuity condition, Equation 3.7, is a natural one to impose when working with ipos and the Scott topology. So we see that the natural way to represent the result of a probabilistic computation is a continuous evaluation over the space of possible results.

Other authors, Saheb-Djahromi, Grajam [35,14] have considered the space of probability distributions over the Borel sets of the result space; we shall see how this compares with evaluations later.

4.1 Evaluations as an Ipo

We shall initially consider the set of continuous evaluations ν on any topological space X with the additional property that $\nu(X) \leq 1$. We give this set the pointwise partial order as follows :-

$$\mu \sqsubseteq \nu \text{ iff for all open sets } O,\ \mu(O) \leq \nu(O) \tag{4.1}$$

So an evaluation is greater than another if the probability of it passing any test (or open set) is greater than that of the other.

It is clear that Equation 4.1 above gives a transitive, reflexive relation and that $\mu \sqsubseteq \eta$ and $\eta \sqsubseteq \mu$ implies $\eta = \mu$. So the definition in Equation 4.1 defines a partial order on the set of evaluations on a topological space X which we will denote by $\mathcal{V}(X)$.

Theorem 4.1 *For any topological space X, $\mathcal{V}(X)$ is a directed complete partial order, with lubs give pointwise and with a least element $O \mapsto 0$.*

Proof First we will see that given a directed set of evaluations they have an upper bound. Let μ_i be a directed set of evaluations when i ranges over some index set I. Then define the value of $\bigsqcup_{i \in I} \mu_i = \eta$ on O by

$$\eta(O) = \sup_{i \in I} \mu_i(O)$$

which is defined since $\mu(O)$ is bounded (by 1) for all μ.

Strictness and modularity follow immediately from the definition.

To show monotonicity and continuity we need to show that finite sums and directed sups commute with sups over directed sets.

To see that

$$\sup_{i \in I}(\mu_i(U)) + \sup_{i \in I}(\mu_i(V)) = \sup_{i \in I}(\mu_i(U) + \mu_i(V))$$

we note that clearly

$$\sup_{i\in I}(\mu_i(U)) + \sup_{i\in I}(\mu_i(V)) \geq \sup_{i\in I}(\mu_i(U) + \mu_i(V))$$

to show this is equality, we pick any $\epsilon > 0$. Then there exists i_1 and i_2 both in I such that $\mu_{i_1}(U) > \sup_{i\in I}(\mu_i(U)) - \epsilon/2$ and $\mu_{i_2}(V) > \sup_{i\in I}(\mu_i(V)) - \epsilon/2$. But since μ_i is directed, there is some i_3 in I such that $\mu_{i_1} \sqsubseteq \mu_{i_3}$ and $\mu_{i_2} \sqsubseteq \mu_{i_3}$; hence

$$\sup_{i\in I}(\mu_i(U) + \mu_i(V)) \geq \mu_{i_1}(U) + \mu_{i_2}(V) > \sup_{i\in I}(\mu_i(U)) + \sup_{i\in I}(\mu_i(V)) - \epsilon$$

but ϵ was arbitrary so we have equality.

To show continuity, we need to see that for an directed set of open sets O_j over $j \in J$,

$$\sup_{i\in I}\sup_{j\in J}\mu_i(O_j) = \sup_{j\in J}\sup_{i\in I}\mu_i(O_j)$$

but this is easy as any sups commute.

Finally we note that the zero evaluation $U \mapsto 0$ which has value zero on any open set is the least element in $\mathcal{V}(X)$. ∎

Another possible candidate for a partial order on $\mathcal{V}(X)$ is

$$\mu \sqsubseteq \eta \iff \forall f\colon X \to [0,1] \text{ upper continuous } \int f\, d\mu \leq \int f\, d\eta$$

we can easily show this is equivalent to the definition we have already discussed. Note the integration sign here refers to integration with respect to evaluations as defined in Section 3.9.

Theorem 4.2

$$\forall O \subseteq \Omega(X) \quad \mu(O) \leq \eta(O) \tag{4.2}$$

if and only if

$$\forall f\colon X \to [0,1] \text{ } u.c., \int f\, d\mu \leq \int f\, d\eta \tag{4.3}$$

Proof Suppose Equation 4.3 holds. Then since for any open set O, the function χ_O is upper continuous and

$$\int \chi_O \, d\mu = \mu(O)$$

then by Equation 4.3 we know

$$\int \chi_O \, d\mu \leq \int \chi_O \, d\eta$$

i.e. $\mu(O) \leq \eta(O)$. So Equation 4.3 implies Equation 4.2.

To show 4.2 implies 4.3 we need to recall from Section 3.8 that every upper continuous function f can be written as an increasing pointwise limit of step functions. For instance, if f takes values only in $[0,1]$, as is the case here then

$$f = \lim_{n \to \infty} \sum_{i=1}^{2^n} 2^{-n} \chi_{f^{-1}(i.2^{-n},1]}.$$

If s is a step function, say $s = \sum_{i=1}^{n} r_i \chi_{O_i}$ then

$$\int s \, d\mu = \sum_{i=1}^{n} r_i \mu(O_i)$$

hence, if Equation 4.2 holds then

$$\int s \, d\mu \leq \int s \, d\eta$$

but for any f, we can write $f = \sup f_n$ where each f_n is a step function and f is the pointwise limit of the sequence of increasing functions, and we know that

$$\int f \, d\mu = \sup \int f_n \, d\mu$$

hence if Equation 4.2 holds,

$$\int f_n \, d\mu \leq \int f_n \, d\eta$$

so

$$\sup_{n \to \infty} \int f_n \, d\mu \leq \sup_{n \to \infty} \int f_n \, d\eta$$

hence

$$\int f \, d\mu \leq \int f \, d\eta$$

as required. ∎

A similar argument can be used to prove that if μ_i is a directed set of evaluations for $i \in I$, then for any upper continuous function $f: X \to [0,1]$

$$\sup_{i\in I} \int f\, d\mu_i = \int f\, d \bigsqcup_{i\in I} \mu_i \tag{4.4}$$

Since for any open set O the equation holds for the function χ_O by the definition of $\bigsqcup \mu_i$, then by linearity and continuity it holds for any upper continuous f.

We shall now assume that the space of results is itself an inductive partial order, P, with the Scott topology. We can relate the partial order on P with the partial order on the set of evaluations on P, $\mathcal{V}(P)$. Recall the definition of a point measure (denoted η_x) from Equation 3.2, this clearly restricts to a point evaluation which we will denote by the same symbol. We shall prove that $\eta_a \sqsubseteq \eta_b$ iff $a \sqsubseteq b$ and $\bigsqcup_{a\in X} \eta_a = \eta_{\bigsqcup_{a\in X} a}$ for a directed set X. To see the first note that if $a \sqsubseteq b$, then for any open set O, if $a \in O$ then $b \in O$, hence if $\eta_a(O) = 1$ then $\eta_b(O) = 1$. Thus $\eta_a(O) \leq \eta_b(O)$ for all open sets O. To show the converse, consider the open set O_b (Section 2.1.2). $b \notin O_b$ so $\eta_b(O_b) = 0$ hence $\eta_a(O_b) = 0$ i.e. $a \notin O_b$ so $a \sqsubseteq b$. For the second part first note that by the first part $\bigsqcup_{a\in X} \eta_a \sqsubseteq \bigsqcup \eta_{\bigsqcup_{a\in X} a}$ To show the converse let O be any open set. If $\eta_{\bigsqcup_{a\in X} a}(O) = 1$, then $\bigsqcup_{a\in X} a \in O$, therefore there is some a_1 in X for which $a_1 \in O$, therefore $\eta_{a_1}(O) = 1$, so clearly $\sup_{a\in X} \eta_a(O) = 1$.

4.2 The Powerdomain Functor

In this section we will define a functor $\mathcal{V}$ on the category **Ipo**.

The category **Ipo** is defined to be the category with ipos as objects and continuous (total) maps between them as morphisms. On objects, the functor $\mathcal{V}$ will map an ipo P to $\mathcal{V}(P)$. We know from Theorem 4.1 above that $\mathcal{V}(P)$ is an ipo with the partial order given by Equation 4.1. Given a continuous ipo morphism $f: P \to Q$ we define

$$\mathcal{V}(f)(\mu)(O) = \mu(f^{-1}(O)) \tag{4.5}$$

for any μ in $\mathcal{V}(P)$ and open set O in Q.

We will now show that the definitions above define a functor **Ipo** $\to$ **Ipo**.

Theorem 4.3 *The operation $\mathcal{V}$ defined above on **Ipo** is a functor.*

Proof We already know that $\mathcal{V}$ on an object in **Ipo** gives another object in **Ipo**. We need to check that if $f: P \to Q$ is a **Ipo** morphism, that $\mathcal{V}(f)$ is an ipo morphism from $\mathcal{V}(P)$ to $\mathcal{V}(Q)$ and that $\mathcal{V}$ preserves identity morphisms and composition.

Given a continuous morphism f, it is clear that Equation 4.5 for fixed μ defines a function from open sets to $[0, 1]$ since if f is continuous we know $f^{-1}(O)$ is open for any open set O. We can show that this function is an evaluation; clearly $f^{-1}(\emptyset) = \emptyset$ so it is strict, if $U \subseteq V$, then $f^{-1}(U) \subseteq f^{-1}(V)$ so it is monotone, it is modular since $f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V)$ and similarly for $\cup$ and finally it is continuous since $f^{-1}(\bigcup U_i) = \bigcup f^{-1}(U_i)$.

Now if $\mu \sqsubseteq \nu$ in $\mathcal{V}(\mathrm{X})$, then $\mathcal{V}(f)(\mu) \sqsubseteq \mathcal{V}(f)(\nu)$ since for any O, $\mu(f^{-1}(O)) \leq \nu(f^{-1}(O))$. So $\mathcal{V}(f)$ preserves $\sqsubseteq$. To see that $\mathcal{V}(f)(\bigsqcup_{i\in I} \mu_i) = \bigsqcup_{i \in I} \mathcal{V}(f)(\mu_i)$ for a directed set μ_i, take any open set O, then

$$\begin{aligned} \mathcal{V}(f)(\bigsqcup_{i\in I} \mu_i)(O) &= (\bigsqcup_{i\in I} \mu_i) f^{-1}(O) = \sup_{i\in I} \mu_i(f^{-1}(O)) \\ &= \sup_{i \in I} \mathcal{V}(f)(\mu_i)(O) = \bigsqcup_{i\in I} \mathcal{V}(f)(\mu_i)(O). \end{aligned}$$

So we see that $\mathcal{V}(f)$ is continuous, i.e. it is a morphism in **Ipo**.

If f is the identity morphism on P, then clearly $\mathcal{V}(f)$ is the identity on $\mathcal{V}(P)$ since $\mathcal{V}(f)(\mu)(O) = \mu(O)$. So $\mathcal{V}$ preserves identity morphisms.

Finally if $g \circ f$ is the composition of f and g then

$$\mathcal{V}(g) \circ \mathcal{V}(f)(\mu)(O) = \mathcal{V}(f)(\mu(g^{-1}(O))) = \mu(f^{-1}(g^{-1}(O)))$$

$$= \mu((g \circ f)^{-1}(O)) = \mathcal{V}(g \circ f)(\mu)(O)$$

so $\mathcal{V}$ is a functor from **Ipo** to **Ipo**. ▮

Given the functor $\mathcal{V}$ it is natural to ask whether any of the interesting subcategories of **Ipo** are preserved. Clearly the subcategory of ω-algebraic ipos is not preserved (for instance $\mathcal{V}(\mathbb{1}) \cong [0,1]$ see Section 4.5). Consistent completeness is not preserved, nor is being a lattice, again see Section 4.5 for this.We will see below (in corollary 5.4) that the functor $\mathcal{V}$ preserves the sub-category of continuous domains.

We will now show that the functor $\mathcal{V}$ is a monad (see Section 2.2). Recall Theorem 2.2 that we can make an function on the objects of a category $\mathcal{V}$ into a monad iff we can find an operation $\dagger$ which takes functions $f: X \to \mathcal{V}(Y)$ to $f^\dagger: \mathcal{V}(X) \to \mathcal{V}(Y)$ and a map $i_X: X \to \mathcal{V}(X)$ so that

$$\mathcal{V}(f) = (i \circ f)^\dagger$$

and which satisfy the equations

$$\begin{aligned} i_X^\dagger &= \mathrm{id}_{\mathcal{V}(X)} \\ f^\dagger \circ i_X &= f \\ g^\dagger \circ f^\dagger &= (g^\dagger \circ f)^\dagger \end{aligned}$$

Then the action of the monad on morphisms is given by $\mathcal{V}(f) = (i_Y \circ f)^\dagger$. So if we can define such an operation and the action on morphisms given by $\dagger$ is the same as that defined above (Equation 4.5) then the functor $\mathcal{V}$ is a monad. We will obviously want to make i the injection map $x \mapsto \eta_x$ and we define $\dagger$, given $f: X \to \mathcal{V}(Y)$, by the equation

$$f^\dagger(\mu)(O) = \int_{x \in X} f(x)(O)\, d\mu.$$

We will later see that $\mathcal{V}(\mathbb{1}) = [0,1]$, which will justify defining $h^\dagger$ for $h\colon X \to [0,1]$ by

$$h^\dagger(\mu) = \int_{x\in X} h(x)\, d\mu.$$

We first prove a useful lemma.

Lemma 4.4 *If h is any upper continuous function $h\colon Y \to [0,1]$ and f is a function $f\colon X \to \mathcal{V}(Y)$, then for any $\mu \in \mathcal{V}(X)$,*

$$\int_{y\in Y} h\, df^\dagger(\mu) = \int_{x\in X} h^\dagger(f(x))\, d\mu.$$

Proof If we substitute for $\dagger$ we see that we need to prove that

$$\int_{y\in Y} h(y)\, d(\lambda O. \int_{x\in X} f(x)(O)\, d\mu) = \int_{x\in X} \int_{y\in Y} h(y)\, df(x)\, d\mu.$$

We first assume that h is a step function, say χ_U for U open in Y. Then both sides of the equation above reduce to the expression

$$\int_{x\in X} f(x)(U)\, d\mu.$$

Then we observe that both equations are linear and continuous in h (by the linearity and continuity of integration), so since every upper continuous function is the sup of an increasing sequence of linear combinations of step functions, the equation holds for any upper continuous h. ∎

Theorem 4.5 *For i as the injection map $x \mapsto \eta_x$ and the operation $\dagger$ defined on a map $f\colon X \to \mathcal{V}(Y)$ by*

$$f^\dagger(\mu)(O) = \int_{x\in X} f(x)(O)\, d\mu$$

as above, the equations above are satisfied, and $\mathcal{V}(f) = (i_Y \circ f)^\dagger$.

Proof To show $\mathcal{V}f = (i_Y \circ f)^\dagger$, simply note that

$$(i_Y \circ f)^\dagger(\mu)(O) = \int_{x\in X} \eta_{f(x)}(O)\, d\mu = \mu(f^{-1}(O)).$$

Substituting the definitions into $i_Y^\dagger = \mathrm{id}_{\mathcal{V}(X)}$ we obtain

$$\int_{x \in X} \eta_x(O)\, d\mu = \mu(O)$$

which is trivial since $\eta_x(O) = \chi_O(x)$.

For the second equation need to show

$$f^\dagger(\eta_x) = f(x)$$

but

$$f^\dagger(\eta_x)(O) = \int_{y \in Y} f(y)(O)\, d\eta_x = f(x)(O).$$

Finally, substituting in the third equation we get

$$\int_{y \in Y} g(y)(O)\, d(f^\dagger(\mu)) = \int_{x \in X} g^\dagger(f(x))(O)\, d\mu.$$

This follows from Lemma 4.4 and by noting that if $h\colon Y \to [0,1]$ is given by $\lambda y.g(y)(O)$ for some $g\colon Y \to \mathcal{V}(Z)$, then

$$h^\dagger(\mu) = \int_{y \in Y} h(y)\, d\mu = \int_{y \in Y} g(y)(O)\, d\mu = g^\dagger(\mu)(O)$$

∎

So we see that $\mathcal{V}$ is a monad. In fact we could have used this theorem to prove that $\mathcal{V}$ is a functor rather than giving the direct proof above. From the theorem we know that $\mathcal{V}(g) = (i_Y \circ g)^\dagger$. Substituting $i_Y \circ g$ for f in Lemma 4.4, we obtain the equation

$$\int h\, d\mathcal{V}(g)(\mu) = \int h^\dagger \circ (i_Y \circ g)\, d\mu$$

but by the second of the equations in the theorem, $h^\dagger \circ i_Y = h$, so we see that

$$\int h\, d\mathcal{V}(g)(\mu) = \int h \circ g\, d\mu \tag{4.6}$$

This equation will be useful later.

The fact that $\mathcal{V}$ is a monad also shows that we can form a category $\mathbf{Ipo}_{\mathcal{V}}$ where objects are ipos and a morphism between X and Y is a continuous function $X \to \mathcal{V}(Y)$, the Kleisli category of the monad $\mathcal{V}$ (see Section 2.2). The identity maps are given by i and composition is given by $f \circ g = f^{\dagger} \circ g$.

Similarly we can also define the category of $\mathcal{V}$-algebras $\mathbf{Ipo}^{\mathcal{V}}$ (see Section 2.2) consisting of ipos with a continuous total map $\alpha\colon \mathcal{V}(X) \to X$ satisfying the appropriate commutative diagrams. Such a map α gives a categorical notion of an ipo with a generalised "probabilistic sum" operation, for instance we can form linear combinations of elements of X by the equation

$$ra + (1-r)b = \alpha(r\eta_a + (1-r)\eta_b) \tag{4.7}$$

Also $\alpha(U \mapsto 0)$ is a least element of X, since for any x in X, $(U \mapsto 0) \sqsubseteq \eta_x$, and applying the continuous map α shows that $\alpha(U \mapsto 0) \sqsubseteq x$. As we know from category theory, for any X in $\mathbf{Ipo}$, $\mathcal{V}(X)$ is a $\mathcal{V}$-algebra with morphism α given by the action of the natural transformation $\mu\colon \mathcal{V}^2 \to \mathcal{V}$ on X, $\mu_X\colon \mathcal{V}^2(X) \to \mathcal{V}(X)$.

Graham [14] gives the definition of an abstract probabilistic domain, an ipo (with bottom) and a continuous function $+\colon [0,1] \times X^2 \to X$ satisfying associativity, commutativity and absorption laws as follows

$$\begin{aligned} a +_1 b &= a \\ (a +_r b) +_s c &= a +_{rs} (b +_{\frac{s(1-r)}{1-sr}} c) \quad (rs \neq 1) \\ a +_r b &= b +_{1-r} a \\ a +_r a &= a \end{aligned}$$

Here $[0,1]$ has the usual (Hausdorff) topology and $+$ is continuous in the product topology of $[0,1] \times X^2$ where X^2 has the Scott topology derived from its partial order as a product ipo. We have slightly generalised Graham in that his definition was for RSFP cpos.

We now define the category with objects as abstract probabilistic domains and morphisms as continuous functions which are linear, that is they satisfy

$$\forall a, b, r \quad f(a +_r b) = f(a) +_r f(b)$$

and which also preserve $\perp$. We denote this category by **Apd**. Equation 4.7 above suggests that an $\mathcal{V}$-algebra is also an abstract probabilistic domain, in fact there is a forgetful functor $\mathcal{U}$ from the category of $\mathcal{V}$-algebras $\mathbf{Ipo}^{\mathcal{V}}$ to the category **Apd**. If (A, α) is an $\mathcal{V}$-algebra, then we define $U(A, \alpha)$ to be A and define the operation $+_r$ on A by

$$a +_r b = \alpha(r\eta_a + (1-r)\eta_b)$$

(as in Equation 4.7), and note that $\perp$ is $\alpha(U \mapsto 0)$. To see that $+$ is continuous, since a composition of continuous functions is continuous we just need to show that the function $f\colon [0,1] \times X^2 \to \mathcal{V}(X)$ given by $f(r, x, y) = r\eta_x + (1-r)\eta_y$ is continuous. Let O be any open set in $\mathcal{V}(X)$ and suppose $f(r, x, y) \in O$, then we can find some $\epsilon > 0$ such that $(r-\epsilon)\eta_x + ((1-r)-\epsilon)\eta_y \in O$ (since O is open and $(r-\epsilon)\eta_x + ((1-r)-\epsilon)\eta_y$ over all $\epsilon > 0$ is a directed set with lub $r\eta_x + (1-r)\eta_y$). With this ϵ define $A = \{(a, b) \mid (r-\epsilon).\eta_a + ((1-r)-\epsilon).\eta_b \in O\}$, then A is open in X^2 (since it is clearly upper-closed and inaccessible by lubs of directed sets) and $(x, y) \in A$ and it is clear that $(r-\epsilon, r+\epsilon) \cap [0,1] \times A \subseteq f^{-1}(O)$. So $f^{-1}(O)$ is open. It satisfies the absorption law since $\alpha(\eta_a) = a$ (from one of the diagrams that α satisfies) and the other laws since (pointwise) addition of evaluations is associative and commutative.

Now if $f\colon (A, \alpha) \to (B, \beta)$ is a morphism in $\mathbf{Ipo}^{\mathcal{V}}$, then f is simply a continuous total function $A \to B$ satisfying the diagram on page 26. So provided it is linear i.e. satisfies the equation

$$f(a +_r b) = f(a) +_r f(b)$$

and preserves $\perp$, then it is a morphism in **Apd**. But the diagram reduces to the equation

$$f(\alpha(\sigma)) = \beta(\lambda O.\sigma(f^{-1}(O)))$$

for all σ in $\mathcal{V}(A)$, and if we set $\sigma = r\eta_a + (1-r)\eta_b$ then $\sigma(f^{-1}(O)) = 1$ if both $a \in f^{-1}(O)$ and $b \in f^{-1}(O)$, and similarly for the other cases; but $a \in f^{-1}(O)$ is equivalent to $f(a) \in O$ so $\lambda O.\sigma(f^{-1}(O)) = r\eta_{f(a)} + (1-r)\eta_{f(b)}$ thus we see that

$$f(a +_r b) = \beta(r.\eta_{f(a)} + (1-r).\eta_{f(b)}) = f(a) +_r f(b)$$

so f is linear. Further f preserves $\perp$ since $\perp = \alpha(U \mapsto 0)$, so by the diagram $f(\perp) = f(\alpha(U \mapsto 0)) = \beta(\mathcal{V}(f)(U \mapsto 0)) = \beta(U \mapsto 0)$ which is $\perp$ in B. So there is indeed a forgetful functor $\mathcal{U}: \mathbf{Ipo}^{\mathcal{V}} \to \mathbf{Apd}$. Later we shall see that on the subcategory of continuous ipos with bottom, this forgetful functor restricts to the identity functor.

We will now consider what it means for a continuous function $f: \mathcal{V}(P) \to \mathcal{V}(Q)$ to be a $\mathcal{V}$-algebra morphism. It has to satisfy the diagram.

$$\begin{tikzcd}
\mathcal{V}^2 P \arrow[r, "\mathcal{V}f"] \arrow[d, "(\mathrm{id}_{\mathcal{V}P})^{\dagger}"] & \mathcal{V}^2 Q \arrow[d, "(\mathrm{id}_{\mathcal{V}Q})^{\dagger}"] \\
\mathcal{V}P \arrow[r, "f"] & \mathcal{V}Q
\end{tikzcd}$$

Written as an equation and substituting the definitions into the diagram we obtain the following

$$\forall \sigma \in \mathcal{V}^2(P) \quad f(\lambda O. \int_{\mu \in \mathcal{V}(P)} \mu(O)\, d\sigma) = \lambda O. \int_{\nu \in \mathcal{V}(Q)} \nu(O)\, d\mathcal{V}(f)(\sigma).$$

We first apply the Equation 4.6 to the R.H.S. of this equation with h as the function $\lambda\nu.\nu(O)$, so simplifying it to

$$f(\lambda O. \int_{\mu \in \mathcal{V}(P)} \mu(O)\, d\sigma) = \lambda O. \int_{\mu \in \mathcal{V}(P)} f(\mu)(O)\, d\sigma \tag{4.8}$$

We call a function which satisfies this equation *super-linear*. From this equation we can also see that the value of f is determined by its value on the point evaluations.

To show this we take any μ in $\mathcal{V}(P)$ and define $\sigma = \lambda O.\mu(\eta_P^{-1}(O))$ where η_P is the natural transformation $P \to \mathcal{V}P$, $\eta_P(x) = \eta_x$. Then σ is in $\mathcal{V}^2(P)$, in fact, $\sigma = \mathcal{V}(\eta_P)(\mu)$. Then by Equation 4.6, the R.H.S. of Equation 4.8 simplifies to $\lambda O. \int_{x \in P} f(\eta_x)(O) d\mu$, and the L.H.S is $f \circ \mu_P \circ \mathcal{V}(\eta_P)$ so from the definition of a monad we know that $\mu_P \circ \mathcal{V}(\eta_P) = \mathrm{id}_{\mathcal{V}(P)}$ i.e. L.H.S. reduces to $f(\mu)$ so super-linearity for this σ gives

$$f(\mu) = \lambda O. \int_{x \in P} f(\eta_x)(O) d\mu.$$

We shall later see that if P is continuous, then any linear function is super-linear.

From Section 2.2 recall that since $\mathcal{V}$ is a monad, $(\mathcal{V}(P), \mu_P)$ is in the category $\mathbf{Ipo}^{\mathcal{V}}$ and we have a theorem which states that given any morphism in $\mathbf{Ipo}$ $f: P \to A$ where (A, α) is a $\mathcal{V}$-algebra there is a unique $\mathcal{V}$-algebra morphism $\bar{f}: \mathcal{V}(P) \to A$ satisfying $\bar{f} \circ i = f$. This morphism is given concretely by $\bar{f} = \mathcal{V}(f) \circ \alpha$. If we assume (A, α) is $(\mathcal{V}(Q), \mu_Q)$ for some ipo Q, this means that the set of continuous functions $f: P \to \mathcal{V}(Q)$ is the same as the set of continuous, super-linear functions $f: \mathcal{V}(P) \to \mathcal{V}(Q)$.

4.3 Evaluations as a Computational Model

In this section we will show that the monadic functor $\mathcal{V}$ that we have defined gives a λ_c-model over $\mathbf{Ipo}$ as defined by Moggi in [26].

We already know that $\mathcal{V}$ is a monad, to further show that it is a λ_c-model we must prove it satisfies three conditions, firstly that for any A, η_A is mono, secondly that it has a *tensorial strength* which respects the monad structure (this is used to define products) and finally a $\mathcal{V}$-exponential (to interpret functions).

That η_A is a mono is trivial, since in Section 4.1 we showed $\eta_A(x) \sqsubseteq \eta_A(y)$ iff $x \sqsubseteq y$, hence $\eta_A(x) = \eta_A(y)$ iff $x = y$, so η_A is a monomorphism, then it is easy to see that it is a mono in the categorical sense.

A *tensorial strength* of a monad T on a category $\mathbf{C}$ with finite products is a natural transformation $t_{A,B}: (A \times TB) \to T(A \times B)$ satisfying certain diagrams. The tensorial strength is given by the formula, for all points $a: 1 \to A$ and $b: 1 \to TB$

$$\langle a, b \rangle; t_{A,B} = b; T(\langle !_B; a, \mathrm{id}_B \rangle)$$

but this may not be a morphism. For $\mathcal{V}$, this says that

$$t_{P,Q}(x, \mu) = \mathcal{V}(\lambda y. \langle x, y \rangle)(\mu)$$

and since $(\lambda y. \langle x, y \rangle)^{-1}(W) = \lambda y. \chi_W(x, y)$, this gives us

$$t_{P,Q}(x, \mu) = W \mapsto \int_{y \in Q} \chi_W(x, y)\, d\mu.$$

We only need to check that this is indeed a morphism, which is trivial by noting that this is equivalent to $W \mapsto \int\int \chi_W(x, y)\, d\eta_P(x)\, d\mu$ which is clearly continuous by the continuity of η, and Equation 4.4. In [26], Moggi derives a morphism $\psi_{A,B}: (TA \times TB) \to T(A \times B)$ from a tensorial strength by the equation

$$\psi_{A,B} = c_{TA,TB}; t_{TB,A}; (c_{TB,A}; t_{A,B})^{\dagger}$$

where c is the natural isomorphism $c_{A,B}: A \times B \to B \times A$. We will show that for $\mathcal{V}$ this natural transformation is actually the product of evaluations $\otimes$ as given in Chapter 3.

We suppose that $\eta \in \mathcal{V}(P)$ and $\nu \in \mathcal{V}(Q)$. Then $c_{\mathcal{V}(P),\mathcal{V}(Q)}(\eta, \nu) = (\nu, \eta)$ and

$$t_{\mathcal{V}(Q),P}(\nu, \eta) = \lambda W \subseteq \mathcal{V}Q \times P. \int_{x \in P} \chi_W(\nu, x) d\eta$$

by substituting the definition of $t_{P,Q}$ given above. But clearly this is just

$$\mathcal{V}(\lambda x. \langle x, \nu \rangle)\eta$$

since

$$(\lambda x. \langle \nu, x \rangle)^{-1}(W) = \{x \mid \langle \nu, x \rangle \in W\}$$

and has characteristic function $\lambda x.\chi(\nu, x)$. So

$$\psi_{P,Q}(\eta, \nu) = (c_{\mathcal{V}(Q),P}; t_{P,Q})^{\dagger}(\mathcal{V}(\lambda x.\langle x, \nu\rangle)\eta) =$$

$$\lambda O. \int_{(\mu,x)\in \mathcal{V}Q\times P} (c_{\mathcal{V}(Q),P}; t_{P,Q})(\mu, x)(O)\, d(\mathcal{V}(\lambda x.\langle \nu, x\rangle)\eta)$$

but applying Equation 4.6, this simplifies to

$$\lambda O. \int_{x\in P} c_{\mathcal{V}(Q),P}; t_{P,Q})^{\dagger}(\nu, x)(O)\, d\eta = \lambda O. \int_{x\in P} t_{P,Q}(x, \nu)(O)\, d\eta$$

and by substituting for $t_{P,Q}$ we get

$$\psi_{P,Q}(\eta, \nu) = \lambda O. \int_{x\in P} \int_{y\in Q} \chi_O(x, y)\, d\nu\, d\eta.$$

Moggi also defines $\tilde{\psi}_{A,B} = c_{A,B}; \psi_{A,B}; Tc_{B,A}$, which can similarly be seen to be

$$\tilde{\psi}_{P,Q}(\eta, \nu) = \lambda O. \int_{y\in Q} \int_{x\in P} \chi_O(x, y)\, d\eta\, d\nu$$

in our case. The natural transformation ψ represents pairing of arguments and Moggi notes that $\tilde{\psi}$ represents evaluating the second argument first. In our case we would intuitively expect that the order of evaluation would make no difference, however we do not know that ψ and $\tilde{\psi}$ are equal in general, although we do know this if one of the ipos is continuous.

Finally to interpret functional types, we need to define a special type of exponential constructor, $(\mathcal{V}B)^A$, denoted $B^A_{\mathcal{V}}$, and the evaluation morphism $\mathrm{eval}^T_{A,B}\colon (B^A_{\mathcal{V}} \times A) \to \mathcal{V}B$ with its universal property that for any $f\colon (C \times A) \to \mathcal{V}B$ there exists a unique $h\colon C \to B^A_{\mathcal{V}}$ such that $f = \mathrm{eval}^{\mathcal{V}}_{A,B} \circ (h \times \mathrm{id}_A)$. In our case this is trivial since **Ipo** is cartesian closed, we just define $B^A_{\mathcal{V}}$ to be the exponential of A and $\mathcal{V}(B)$ and eval to be the usual evaluation morphism.

4.4 Technical Lemmas

This section contains some lemmas which give the conditions under which one evaluation is below another. The general idea is that for finite linear combinations of point evaluations, the definition of $\sqsubseteq$ which involves checking the value of the evaluations on every open set, should be simplifiable to something involving only a finite number of checks.

The first lemmas are simply derived from considering minimal and maximal open sets containing or not containing a particular set of points. For any ipo the maximal open set not containing x is $O_x = \{y \mid y \not\sqsubseteq x\}$. There is no minimal open set containing a point x, as the set $A_x = \{y \mid x \sqsubseteq y\}$, which is the intersection of all open sets containing x is not generally open. However we can get around this by considering $\inf\{\mu(O) \mid O \text{ open}, A_x \subseteq O\}$.

We have seen (Section 4.1) that $\eta_a \sqsubseteq \eta_b$ iff $a \sqsubseteq b$ and clearly $r\mu \sqsubseteq s\mu$ iff $r \leq s$. We give a lemma which shows that in the case of linear combinations of point evaluations, these are essentially the only ways in which we can have $\mu \sqsubseteq \nu$. More precisely we show that given two finite linear combinations of point evaluations, if one is below the other then we can express them as more complicated combinations (i.e. by splitting up say $r\eta_a$ into $(r-q)\eta_a + q\eta_a$) such that every term in the smaller evaluation has a corresponding term in the larger and these pairs are of the form $r\eta_a$ and $s\eta_b$ where $r \geq s$ and $a \sqsubseteq b$.

Finally we give some conditions for $\mu \ll \nu$. For these lemmas we need to assume the ipo is continuous.

Lemma 4.6 *For any evaluation of the form $\sum_{b \in B} r_b \eta_b$ where B is finite,*

$$\mu \sqsubseteq \sum_{b \in B} r_b \eta_b$$

iff $\forall K \subseteq B$ *such that* K *is upward closed in* B *(i.e. if* $c \in K$ *and* $c \sqsubseteq b$ *and* $b \in B$ *then* $b \in K$*)*

$$\mu(\bigcap_{b\in(B\setminus K)} O_b) \leq \sum_{b\in K} r_b.$$

Proof Note that $\bigcap_{b\in(B\setminus K)} O_b$ is the maximal open set containing all $x \in K$, but no $x \in (B \setminus K)$. Then it is clear that

$$\sum_{b\in B} r_b\eta_b(\bigcap_{b\in(B\setminus K)} O_b) = \sum_{b\in K} r_b$$

hence the first equation implies the second. To see the converse, let O be any open set and define $K = \{b \mid b \in O \cap B\}$, hence K is clearly upward closed in B. Now trivially $O \subseteq \bigcap_{b\in(B\setminus K)} O_b$, hence

$$\mu(O) \leq \mu(\bigcap_{b\in(B\setminus K)} O_b) \leq \sum_{b\in K} r_b$$

and the lemma is proven. ∎

Corollary 4.7 *For two linear combinations of point evaluations*

$$\sum_{b\in B} r_b\eta_b \sqsubseteq \sum_{c\in C} s_c\eta_c$$

iff $\forall K \subseteq C$ *with* K *upward closed,*

$$\sum_{\substack{\forall c\in(B\setminus K)\\ b\not\sqsubseteq c}} r_b \leq \sum_{c\in K} s_c.$$

Proof This is proved from Lemma 4.6 by showing that

$$(\sum_{b\in B} r_b\eta_b)(\bigcap_{b\in(B\setminus K)} O_c) = \sum_{\substack{\forall c\in(B\setminus K)\\ b\not\sqsubseteq c}} r_b$$

but this is easy since $b \in \bigcap_{b\in(B\setminus K)} O_c$ iff $\forall c \in (B \setminus K)$, $b \in O_c$, i.e. $b \not\sqsubseteq c$. ∎

For the next lemmas we shall be using sets of the form $\bigcup_{i=1}^{n} A_{x_i}$, where $A_x = \{y \mid x \sqsubseteq y\}$ as before. Clearly A_x is upper closed for any x and so is a union of such sets.

Note that we can extend any continuous evaluation μ to upper closed sets by defining for A upper closed

$$\bar{\mu}(A) = \inf\{\mu(O) \mid A \subseteq O, O \text{ open}\}$$

this is an extension since if A is open then clearly $\mu(A)$ is the inf of the set above. Also the extension is monotonic since if $A \subseteq B$ then $\bar{\mu}(A)$ is the inf over a set of values containing those for $\bar{\mu}(B)$.

We also note that for any upper closed set A,

$$A = \bigcap\{O \mid A \subseteq O, O \text{ open}\}$$

that $A \subseteq \bigcap_{A \subseteq O} O$ is obvious. For the converse, suppose that x is not in A, then since A is upper closed, $x\!\downarrow \cap A = \emptyset$ so $A \subseteq O_x$, hence since $x \notin O_x$, $x \notin \bigcap_{A \subseteq O} O$, so $\bigcap_{A \subseteq O} O = A$.

Finally we note that if μ happens to be a linear combination of point evaluations, say $\mu = \sum_{b \in B} r_b \eta_b$ (B a finite set as usual), then $\bar{\mu}(A) = \sum_{b \in A \cap B} r_b$, since firstly, if $A \subseteq O$, then $\mu(O) \geq \sum_{b \in A \cap B} r_b$, and secondly the set $\bigcap_{b \in B \setminus A} O_b$ is open and $A \sqsubseteq \bigcap_{b \in B \setminus A} O_b$, and so $\mu(O) \leq \mu(\bigcap_{b \in B \setminus A} O_b) = \sum_{b \in A \cap B} r_b$.

In the next lemma (Lemma 4.8) we are really checking the value of μ on infinitely many open sets by conditions on $\mu(A)$. If the underlying space is ω-continuous, then A is a countable intersection of open sets (see Sections 2.1.3 and 3.2) so we are only checking the value on countably many sets and if μ happens to be a measure, the value of $\bar{\mu}(A)$ is also the measure of A.

Lemma 4.8

$$\sum_{b \in B} r_b \eta_b \sqsubseteq \mu$$

iff $\forall K \subseteq B$ *with* K *upward closed in* B*,*

$$\sum_{b \in K} r_b \leq \mu(\bigcup_{b \in K} A_b).$$

Proof Note that $\bigcup_{b\in K} A_b$ is contained in any open set that contains K, and since K is upward closed in B, it does not contain any $x \in (B \setminus K)$.

To prove the if case, recall from above that if $\mu \sqsubseteq \nu$ then $\mu(\bigcap_{i=1}^n A_{x_i}) \leq \nu(\bigcap_{i=1}^n A_{x_i})$. So $\sum_{b\in B} r_b\eta_b \sqsubseteq \mu$ implies that $\sum_{b\in K} r_b \leq \mu(\bigcup_{b\in K} A_b)$.

Now we assume the second equation holds for any K satisfying the conditions above. Let O be any open set and set $K = \{b \mid b \in B \cap O\}$, clearly K is upward closed in B. Then $O \supseteq \bigcup_{b\in K} A_b$ so $\mu(O) \geq \mu(\bigcup_{b\in K} A_b)$. But clearly $(\sum_{b\in B} r_b\eta_b)(O) = \sum_{b\in K} r_b$ hence from the second equation $(\sum_{b\in B} r_b\eta_b)(O) \leq \mu(O)$. ∎

Corollary 4.9

$$\sum_{b\in B} r_b\eta_b \sqsubseteq \sum_{c\in C} s_c\eta_c$$

iff $\forall K \subseteq B$ *with* K *upward closed in* B,

$$\sum_{b\in K} r_b \leq \sum_{\substack{c\in C\, s.t.\\ \exists b\in K, b\sqsubseteq c}} s_c.$$

Proof This corollary is trivial from lemma 4.8 and the fact that $c \in \bigcup_{b\in K} A_b$ implies $\exists b$ such that $c \in A_b$ i.e. $b \sqsubseteq c$. ∎

We will next prove a theorem which says that when one linear combination of point evaluations is below another, say

$$\sum_{b\in B} r_b\eta_b \sqsubseteq \sum_{c\in C} s_c\eta_c,$$

then we can split up each $r_b\eta_b$ into several parts, each of the form

$$\sum_{c, b\sqsubseteq c} t_{b,c}\eta_b,$$

(i.e. ranging over all c s.t. $b \sqsubseteq c$) and then re-assemble the whole sum into parts of form

$$\sum_{b, b\sqsubseteq c} t_{b,c}\eta_b$$

so each part is less than $s_c\eta_c$. This lemma could use either Corollary 4.7 or Corollary 4.9, but the latter is a more natural choice.

Theorem 4.10 (Splitting Lemma) *For two linear combinations of point evaluations if*

$$\sum_{b\in B} r_b\eta_b \sqsubseteq \sum_{c\in C} s_c\eta_c$$

then $\exists t_{b,c}$ *such that*

$$\sum_{c\in C} t_{b,c} = r_b$$

$$\sum_{b\in B} t_{b,c} \leq s_c$$

and $t_{b,c} \neq 0$ *implies* $b \sqsubseteq c$.

Proof The proof uses a directed version of the Max-flow, Min-cut theorem, as in [7] (also [12]). This theorem states that the maximal flow through a directed graph where edges each have a capacity (a real number) and there are two distinguished nodes (the source and the sink) is equal to the value of the minimal cut. Here a cut is simply a set of nodes including the source but not the sink. The value of a cut is the sum of the capacities of all the edges going from a node in the cut to one outside it. A flow is defined in the obvious way, being a set of directed flows along each edge such that the flow into any node is equal to the flow out except at the source or sink and its value is the flow out from the source (or into the sink).

We construct a graph with a line of nodes for each $b \in B$ connected by an edge of capacity r_b to the source, a similar set for $c \in C$ connected to the sink and the nodes b and c connected by a line with "large" capacity (say 1) if $b \sqsubseteq c$. It is clear that a flow through the network represents a way of splitting the r_bs, the flow along a node connecting b to c gives the value of $t_{b,c}$. If there exists a flow which has value r_b along each edge connecting source to b, for all b, then this flow will give values of $t_{b,c}$ which satisfy the conditions of the theorem and this flow will clearly have maximal value.

So it remains to show that the value of any cut is greater than the value of such a flow (i.e. $\sum r_b$). Let T be any cut. If T is such that for some b and c, $b \sqsubseteq c$ and b is in the cut and c is not in the cut, then the value of T will include the capacity of the edge joining b and c and will therefore be large. So suppose T is a cut which does not have any such pairs b and c. Let K be the set of $b \in T$. Now the value of the cut T is $\sum_{c \in T} s_c + \sum_{b \notin T} r_b$. Consider the set $K' = \{c \mid c \in C, \exists b \in K\uparrow \text{ s.t. } b \sqsubseteq c\}$. If c is in K', then there exists some b in $K\uparrow$ with $b \sqsubseteq c$, where $K\uparrow$ is the upward closure on K in B, hence if c is in K' there is some b in K with $b \sqsubseteq c$, Then by the condition that T does not cut any lines of capacity 1, c is in T, i.e. $K' \subseteq T$. But then by Corollary 4.9 above $\sum_{c \in K'} s_c \geq \sum_{b \in K\uparrow} r_b$, hence we have $\sum_{c \in T} s_c \geq \sum_{c \in K'} s_c \geq \sum_{b \in K\uparrow} r_b \geq \sum_{b \in T} r_b$. So the value of the cut T is greater than $\sum_{b \in B} r_b$, i.e. there is a flow with value $\sum_{b \in B} r_b$. ∎

This application of the Max-flow, Min-cut theorem was inspired by a similar application of the theorem in [7] to prove Hall's Theorem.

We now look at the well-below operation. We shall first note that the basic result that if $a \ll b$ then $(1-\epsilon)\eta_a \ll \eta_b$ is the best that can be obtained. If we omit the $1-\epsilon$ multiplier we get a counterexample by considering the directed set $(1-\epsilon).\eta_b$ over $1 > \epsilon > 0$ which has lub η_b, then clearly η_a cannot be below any of these terms since $\eta_a(P) = 1$.

Lemma 4.11 *In a continuous space and for any* $1 > \epsilon > 0$

$$(1-\epsilon).\eta_a \ll \eta_b$$

iff $a \ll b$

Proof We first note that $(1-\epsilon)\eta_a \sqsubseteq \eta_b$ implies $a \sqsubseteq b$ for any $1 > \epsilon > 0$, since we must have $\eta_a(O_b) = 0$, i.e. $a \sqsubseteq b$. We assume $(1-\epsilon)\eta_a \ll \eta_b$ and we want to show $a \ll b$. Now consider any directed set X with $b \sqsubseteq \bigsqcup X$. We know that

$\eta_b \sqsubseteq \eta_{\bigsqcup X} = \bigsqcup_{d \in X} \eta_d$. Hence since $(1-\epsilon)\eta_a \ll \eta_b$ then there is some d in X such that $(1-\epsilon)\eta_a \sqsubseteq \eta_d$ hence $a \sqsubseteq d$, so we have shown that $a \ll b$.

To prove the converse, let X be any directed set of evaluations with $\eta_b \sqsubseteq \bigsqcup X$. Now assuming $a \ll b$, we use the density result (Section 2.1.3) to find x satisfying $a \ll x \ll b$. Now consider V_x, $\eta_b(V_x) = 1$ thus there exists some ν in X such that $\nu(V_x) > (1-\epsilon)$ (since $\eta_b(O) \leq \sup_{\nu \in X} \nu(O)$ for any open set O). But by Corollary 4.9, $(1-\epsilon)\eta_a \sqsubseteq \nu$ if $\nu(A_a) \geq (1-\epsilon)$ where $\nu(A_a)$ is given by $\nu(A_a) = \inf\{\nu(O) \mid A_x \subseteq O, O \text{ open}\}$ but clearly $V_x \subseteq A_a$ since $a \ll x$. Hence if O contains A_x, it contains V_x, so clearly $\nu(O) \geq \mu(V_x)$, hence $\nu(A_a) \geq (1-\epsilon)$, so $\eta_a \sqsubseteq \nu$, hence finally $\eta_a \ll \eta_b$. ∎

There is an obvious generalisation to this, $\mu \ll \eta_b$, iff $\exists \epsilon > 0$ and $a \ll b$ with $\mu \sqsubseteq (1-\epsilon)\eta_a$. One implication is obvious from the theorem, for the other we note that if X is the directed set of points way-below b, then the set given by $\{(1-\epsilon).\eta_d \mid \epsilon > 0, d \in X\}$ is also directed and has least upper bound η_b, then the fact that $\mu \ll \eta_b$ gives a d and ϵ with the desired properties.

The next result gives an obvious condition for when one linear combination of point evaluations is below another.

Lemma 4.12 *For a continuous P and $\epsilon > 0$, if $b' \ll b$ for all $b \in B$ then*

$$\sum_{b \in B} (1-\epsilon) r_b \eta_{b'} \ll \sum_{b \in B} r_b \eta_b.$$

Proof We consider any directed set of evaluations X with

$$\sum_{b \in B} r_b \eta_b \sqsubseteq \bigsqcup X$$

and we want to prove that for some ν in X,

$$\sum_{b \in B} (1-\epsilon) r_b \eta_{b'} \sqsubseteq \nu$$

and to do this, by Lemma 4.8, we simply need to prove that for any $K \subseteq B$ upward closed in B

$$(1-\epsilon)\sum_{b\in K} r_b \leq \nu(\bigcup_{b\in K} A_{b'})$$

but we can show this exactly as before by finding b'' such that $b' \ll b'' \ll b$ and finding ν_K such that

$$\nu_K(\bigcup_{b\in K} V_{b''}) > (1-\epsilon).\sum_{b\in K} r_b$$

then a joint upper bound ν of ν_K over the finitely many K's that are upward closed subsets of B will satisfy $\sum_{b\in B}(1-\epsilon)r_b\eta_b \sqsubseteq \nu$ as required. ∎

With Lemma 4.12 and the Splitting Lemma, we can prove a directed version of the Splitting Lemma.

Lemma 4.13 *For two linear combinations of point evaluations if*

$$\sum_{b\in B} r_b\eta_b \ll \sum_{c\in C} s_c\eta_c$$

then $\exists t_{b,c}$ *such that*

$$\sum_{c\in C} t_{b,c} = r_b$$

$$\sum_{b\in B} t_{b,c} < s_c{}^{1}$$

and $t_{b,c} \neq 0$ *implies* $b \ll c$.

Proof We consider the directed set $\sum_{c\in C}(s_c-\epsilon)\eta_{c'}$ for any $c' \ll c$ and ϵ s.t. $\min_{c\in C}(s_c) > \epsilon > 0$. It clearly has lub $\sum_{c\in C} s_c\eta_c$, so we can find c' and $\epsilon > 0$ such that

$$\sum_{b\in B} r_b\eta_b \sqsubseteq \sum_{c\in C}(s_c-\epsilon)\eta_{c'}$$

---

[1]We assume $s_c > 0$ for all $c \in C$

we than apply the Splitting Lemma (Lemma 4.10) to obtain $t_{b,c}$ such that

$$\sum_{c \in C} t_{b,c} = r_b$$

and

$$\sum_{b \in C} t_{b,c} \le (s_c - \epsilon)$$

and such that $t_{b,c} \ne 0$ implies that $b \sqsubseteq c'$. Then this $t_{b,c}$ will do since $c' \ll c$ so $b \sqsubseteq c'$ implies $b \ll c$ and

$$\sum_{b \in C} t_{b,c} < s_c$$

as required. ∎

We will later use these results to help show that the set of evaluations on an continuous ipo is a continuous ipo itself.

4.5 The Action of the Powerdomain Functor

Let $X = \mathbb{1}$ (the one point ipo). Note that the topology is simply $\mathcal{O}(X) = \{X, \emptyset\}$. It is clear that any evaluation is defined by its value on X, i.e. any evaluation is $a.\eta_1$ for some $a \in [0,1]$. It is also easy to see that $a.\eta_1 \sqsubseteq b.\eta_1$ iff $a \le b$. So we see that $\mathcal{V}(\mathbb{1}) = [0,1]$ with the upper-set topology, i.e. open sets are of the form $(a, 1]$.

We can use the Splitting Lemma to characterise the powerdomain on a finite ipo P. Suppose that P has n elements which we order say as $p_1, \ldots, p_n$, then $\mathcal{V}(P)$ is the set of n-tuples of positive real numbers with sum less than 1, i.e. $(r_1, \ldots, r_n)$ such that $\sum_{i=1}^{n} r_n \le 1$ and $r_i \ge 0$ for all $1 \le i \le n$, with the partial order $(r_1, \ldots, r_n) \sqsubseteq (s_1, \ldots, s_n)$ iff there exists $t_{i,j}$ such that $\sum_{j=1}^{n} t_{i,j} = r_i$ and $\sum_{i=1}^{n} t_{i,j} \le s_j$ and $t_{i,j} \ne 0$ implies $p_i \sqsubseteq p_j$. To see that this is indeed $\mathcal{V}(P)$ we note that any upper closed set in P is open, so every point x is the crescent

$A_x \setminus O_x$, so every evaluation, which gives a finitely additive function on crescents, is determined by its value on the points of P, also every collection of values r_i with sum less that one is an evaluation since it is $\sum_{i=1}^{n} r_i \eta_{p_i}$ and the partial order is given by the Splitting Lemma.

We can show that $\mathcal{V}(P)$ is not consistently complete, even when P is. For this let P have three elements, $P = \{a, b, a \vee b\}$ where neither $a \sqsubseteq b$ or $b \sqsubseteq a$ holds. Clearly P is consistently complete. Consider the two measures $1/2\, \eta_P(a)$ and $1/2\, \eta_P(b)$. They have joint upper bounds $1/2\, \eta_P(a \vee b)$ and $1/2\, \eta_P(a) + 1/2\, \eta_P(b)$ but no least upper bound, since a least upper bound must be below each of these upper bounds, but $\mu \sqsubseteq 1/2\, \eta_P(a) + 1/2\, \eta_P(b)$ implies $\mu(a \vee b) = 0$ and for such a μ to be an upper bound of $1/2\, \eta_P(a)$ and $1/2\, \eta_P(b)$, it must be equal to $1/2\, \eta_P(a) + 1/2\, \eta_P(b)$. But clearly $1/2\, \eta_P(a) + 1/2\, \eta_P(b) \not\sqsubseteq 1/2\, \eta_P(a \vee b)$, since $1/2\, \eta_P(a) + 1/2\, \eta_P(b)$ has value 1 on the whole set and $1/2\, \eta_P(a \vee b)$ has value $1/2$. Similarly $\mathcal{V}(P)$ is not generally a lattice if P is, to see this just consider the example above with an extra element $\perp$ added.

Consider $\mathcal{V}(P+Q)$. An open set in $P+Q$ is the union of two unique open sets in P and Q respectively, so any evaluation is defined by the values it takes on the open sets contained in P or Q (since from modularity $\mu(U) = \mu(U \cap P) + \mu(U \cap Q)$). Also it is clear that the restriction of an evaluation on $P+Q$ to P or Q is also an evaluation. So

$$\mathcal{V}(P+Q) = \{(\mu, \eta) \mid \mu \in \mathcal{V}(P), \eta \in \mathcal{V}(Q), \mu(P) + \eta(Q) \leq 1\}.$$

In fact this is the categorical sum in the category $\mathbf{Ipo}^{\mathcal{V}}$ of $\mathcal{V}(P)$ and $\mathcal{V}(Q)$, and for any monad this is known to be the case.

The definition in Section 3.10 of the product of two evaluations defines a morphism $\mathcal{V}(P) \times \mathcal{V}(Q) \to \mathcal{V}(P \times Q)$ either by the fact that $\psi_{P,Q}$ is a morphism (which is because $t_{P,Q}$ is a morphism) or directly since if $(\eta_1, \nu_1) \sqsubseteq (\eta_2, \nu_2)$ then

$$\eta_1 \otimes \nu_1(O) \leq \eta_2 \otimes \nu_2(O)$$

(by monotonicity of integration) and given $\bigsqcup_{i \in I}(\eta_i, \nu_i) = (\eta, \nu)$, η_i, ν_i directed sets, then

$$\bigsqcup_{i \in I}(\eta_i \otimes \nu_i)(O) = \sup_{i \in I}(\int_{x \in P} \nu_i(\{y \mid (x,y) \in O\})\, d\eta_i)$$

which by directedness of η_i and ν_i equals

$$\sup_{i \in I} \sup_{j \in I} \int_{x \in P} \nu_j(\{y \mid (x,y) \in O\})\, d\eta_i.$$

But then as the functions of x inside form a directed set by directed monotone convergence,

$$R.H.S. = \sup_i \int_{x \in P} \nu(\{y \mid (x,y) \in O\})\, d\eta_i$$

and then again by Equation 4.4

$$= \int_{x \in P} \sup_i(\nu_i(\{y \mid (x,y) \in O\}))\, d\eta = \eta \otimes \nu(O).$$

4.6 Concluding Remarks

So far as I know , the idea in this section of using evaluations to represent the outcome of a probabilistic computation is original. The partial order was inspired by Saheb-Djahromi's work (although obvious by itself). The categorical aspects of this chapter are entirely new.

The discussion in the previous section on the action of the powerdomain is continued in Chapter 7, we define probabilistic versions of useful functions such as projections and injections.

Chapter 5

Evaluations on Continuous Ipos

Computationally we expect to use only the minimal set of evaluations containing all point evaluations, and which is closed under finite linear combinations and increasing limits. For a continuous ipos this is just the ipo generated by limits of linear combinations of point evaluations. In this chapter we will show that if P is a continuous ipo, all continuous evaluations are a directed limit of linear combinations of point evaluations. This main result will enable us to show that $\mathcal{V}$-algebras and abstract probabilistic domains are isomorphic on continuous ipos. We will also show that this makes all continuous evaluations on continuous ipos extend to measures, so showing that the powerdomain of evaluations is identical to the ones of measures constructed by Saheb-Djahromi and Graham.

For measures, the question of whether all measures can be generated from point measures has been examined before, Djahromi (in [35]) shows this can be done for ω-algebraic ipos and Graham [14] shows it for RSFP domains.

5.1 Characterising Continuous Evaluations

Here we prove the main result of this chapter, that every continuous evaluation is the sup of a directed set of linear combinations of point evaluations.

We prove this theorem by constructing a set of linear combinations of point evaluations below some continuous evaluation μ using dissections which are a collection of disjoint sets of the form $V_b \setminus O$. We form linear combinations of point evaluations from these dissections by summing terms $\mu(V_b \setminus O)\eta_P(b)$. The hardest part is showing that the set of linear combinations of point evaluations is directed.

We will need the theorem of Pettis [28], that any evaluation extends to a unique additive function on the field generated by the Borel sets, and one other lemma about the continuity of this extension which we will state and prove now.

Lemma 5.1 *Let η be the extension of a continuous evaluation to the field generated by the open sets. If C is closed and the sets O_i over some index set I are open, then*

$$\eta((\bigcup_{i\in I} O_i) \cap C) = \sup\{\eta(\bigcup_{j=1}^{n} O_{i(j)} \cap C) \mid \textit{any finite } i(1), \ldots, i(n) \textit{ from } I\}.$$

Proof Since η is the extension of an evaluation and is finitely additive, we know that for any open set O and closed set C,

$$\eta(O \cap C) = \eta(O) - \eta(O \setminus C)$$

so we need to show that

$$\eta(\bigcup_{i\in I} O_i) - \eta((\bigcup_{i\in I} O_i) \setminus C) = \sup\{\eta(O) - \eta(O \setminus C) \mid \text{for } O \text{ any finite union of } O_i\text{'s}\}.$$

We first see that we can distribute the sup on the right of this equation since if we increase O (by adding more O_i's), then not only do $\eta(O)$ and $\eta(O \setminus C)$

increase, but so does $\eta(O) - \eta(O \setminus C)$ (since we know that the extension of ν is monotone), hence by a simple directedness argument, $\sup(\eta(O) - \eta(O \setminus C)) = \sup(\eta(O)) - \sup(\eta(O \setminus C))$.

Now if we consider any union of open sets $\bigcup_{i \in I} U_i$, then the set consisting of all finite unions of U_is is directed, since the union of two finite unions of U_i's is also a finite union of U_i's, and its lub is clearly the whole union, so we know from the continuity of η as an evaluation that

$$\sup\{\eta(U) \mid U \text{ is a finite union of } U_i\text{'s}\} = \eta(\bigcup_{i \in I} U_i),$$

and this formula holds for the O_i's. Similarly we can form a directed set of finite unions of $O_i \setminus C$'s which has lub $(\bigcup_i O_i) \setminus C$ then we also have

$$\sup\{\eta(O \setminus C) \mid O \text{ is a finite union of } O_i\text{'s}\} = \eta((\bigcup_i O_i) \setminus C).$$

Then if we subtract the second equality from the first we obtain the desired result.
∎

We now prove the main theorem.

Theorem 5.2 *If P is a continuous ipo, then every continuous evaluation in $\mathcal{V}(P)$ is the lub of a directed set of linear combinations of point evaluations.*

Proof Recall from Section 2.1.3, the notation

$$V_b = \{x \in P \mid b \ll x\}$$

and the fact that for all $b \in P, V_b$ is open in the Scott Topology on P and that $\{V_b \mid b \in B\}$ generate the Scott topology.

We define a *dissection* D, to be a finite set of pairwise disjoint sets $C_1, \ldots, C_n$ where each C_i is of the form $V_{x_i} \setminus U_i$, for x_i in P and U_i open. Given a continuous evaluation μ and any $0 < r < 1$ we can then form a linear combination of point evaluations $\mu(D, r)$ as follows.

$$\mu(D, r) = \sum_{i=1}^{n} r\mu(C_i)\eta_P(x_i)$$

where by $\mu(C_i)$ we mean the measure of C_i in the extension of μ to the field, and so $\mu(C_i) = \mu(V_{x_i}) - \mu(V_{x_i} \setminus U_i)$.

We observe that for any r and D, $\mu(r, D) \sqsubseteq \mu$. This is clear since for any open set O, if $x_i \in O$, then clearly $C_i \subseteq O$ and since the C_i are disjoint, then $\mu(O) \geq \sum_{x_i \in O} \mu(C_i)$ and $\mu(D, r)(O) = r \sum_{x_i \in O} \mu(C_i)$.

We will now state and prove a lemma which we will use in the main proof.

Lemma 5.3 *Let P be a continuous ipo and μ a continuous evaluation on P. Then for any open sets $O_1, \ldots, O_n$ and $0 < r < 1$, there exists a dissection D and $0 < s < 1$ such that for any $B \subseteq \{1, \ldots, n\}$,*

$$r\mu(\bigcup_{i \in B} O_i) < \mu(D, s)(\bigcup_{i \in B} O_i).$$

Proof Suppose T is the difference between two open sets (a *crescent*), $T = O \setminus V$ and $x_1, \ldots, x_n$ are points in T. Then we can construct a dissection of sets contained in T by setting

$$T_i = V_{x_i} \setminus (\bigcup_{j<i} V_{x_j} \cup V).$$

These T_i's are clearly pairwise disjoint and contained in T.

Further if we set $U = \bigcup_{x \in T} V_x$, then U is open and $U \setminus V = O \setminus V$, since if $x \in O$ and $x \notin U$ this implies $x \in V$. We then apply Lemma 5.1 with the closed set C as $P \setminus V$ and the family of sets O_i as V_x over $x \in T$. So we know that

$$\sup_{x_1, \ldots, x_n \in C} \mu(\bigcup_{i=1}^{n} V_{x_i} \cap C) = \mu(C).$$

So we take $x_1, \ldots, x_n$ such that

$$\mu(\bigcup_{i=1}^{n} V_{x_i} \cap C) > \sqrt{r}\mu(C)$$

and so if we form a dissection, D_C of C with these points as described above, then

$$\mu(D_C, s)(C) = s\mu(\bigcup_{i=1}^{n} V_{x_i} \cap C) > s\sqrt{r}\mu(C) \tag{5.1}$$

(Again we have extended $\mu(D_C, s)$ to the field generated by the open sets.)

We carry out this process this with the (disjoint) crescents given by

$$C_J = \bigcap_{i \in J} O_i \setminus \bigcap_{i \notin J} O_i$$

for any $J \subseteq \{1, \ldots, n\}$ and put all the dissections D_{C_J} together, to give a dissection D.

Now consider $\mu(D, \sqrt{r})(\bigcup_{i \in B} O_i)$ for any $B \subseteq \{1, \ldots, n\}$. Clearly $\bigcup_{i \in B} O_i = \bigcup_{J \supseteq B} C_J$ and the C_J's are pairwise disjoint. Clearly

$$\mu(D, \sqrt{r})(C_J) = \mu(D_{C_J}, \sqrt{r})(C_J)$$

so by Equation 5.1 above,

$$\mu(D, \sqrt{r})(\bigcup_{i \in B} O_i) = \sum_{J \supseteq B} \mu(D_{C_J}, \sqrt{r})(C_J) > \sum_{J \supseteq B} r\mu(C_J) = r\mu(\bigcup_{i \in B} O_i)$$

as required. ∎

To prove the main theorem we consider the collection of linear combinations of point evaluations A, given by

$$A = \{\mu(D, r) \mid 0 < r < 1, D \text{ a dissection of P}\}.$$

We will first show that A is directed. Let $\mu(D_1, r_1)$ and $\mu(D_2, r_2)$ be any two members of A. Suppose that D_1 has components $C_1 = V_{x_1} \setminus U_1, \ldots, C_n = V_{x_n} \setminus U_n$. Recall Lemma 4.8 which stated that

$$\sum_{b \in B} r_b \eta_b \sqsubseteq \mu$$

iff $\forall K \subseteq B$ with K upward closed in B,

$$\sum_{b \in K} r_b \leq \mu(\bigcup_{b \in K} A_b)$$

where by $\mu(A)$ for A is upper closed, we mean $\inf\{\mu(O) \mid A \subseteq O, O \text{ open}\}$. Note that $A_b \supseteq V_b$, so a sufficient condition for $\mu(D_1, r_1) \sqsubseteq \nu$ is that for any $K \subseteq \{1, \ldots, n\}$, such that the set of points x_i over $\in K$ is upper closed in $\{x_1, \ldots, x_n\}$

$$r_1 \sum_{i \in K} \mu(C_i) \leq \nu(\bigcup_{i \in K} V_{x_i}).$$

Also $C_i \subseteq V_{x_i}$ so $\bigcup_{i \in K} C_i \subseteq \bigcup_{i \in K} V_{x_i}$ hence it is also sufficient to merely show that

$$r_1 \mu(\bigcup_{i \in K} V_{x_i}) \leq \nu(\bigcup_{i \in K} V_{x_i}).$$

Also if D_2 is given by sets $V_{y_1} \setminus W_1, \ldots V_{y_n} \setminus W_n$, then a similar condition is sufficient to show $\mu(D_2, r_2) \sqsubseteq \nu$. But we can obtain a D and an s so that both these conditions are satisfied by $\mu(D, s)$ by applying the lemma above to $\max(r_1, r_2)$ and the sets $V_{x_1}, \ldots, V_{x_n}, V_{y_1}, \ldots, V_{y_m}$. So A is directed.

Finally, to show that the least upper bound of A is μ first note that μ is clearly an upper bound of A. To show it is the least upper bound consider any open set O, then apply the lemma above to see that for any $0 < r < 1$, there is some $\nu \in A$ such that $\nu(O) > r\mu(O)$ so $\sup_{\nu \in A} \nu(O) = \mu(O)$. ∎

Corollary 5.4 *If P is continuous, then $\mathcal{V}(P)$ is continuous.*

Proof Recall that any linear combination of point evaluations can easily be written as the limit of a well-below directed set of other linear combinations of point evaluations from Section 4.4.

So given any evaluation μ, we can take the directed set constructed from Theorem 5.2 above, then for each linear combination of point evaluations in this set we can find a well-below directed set of point evaluations for it and take the union as in Section 2.1.3 to get a well-below directed set of linear combinations of point evaluations for μ. Then the set of evaluations well-below μ is directed with lub μ since between any $\eta \ll \mu$ and μ is a linear combination of point evaluations also well-below μ. ∎

Corollary 5.5 *If P is ω-continuous, then so is $\mathcal{V}(P)$.*

Proof We need to find a countable basis of $\mathcal{V}(P)$. But if B is a countable basis of P then

$$\{\sum_{i=1}^{n} r_i \eta_{b_i} \mid b_i \in B, r_i \text{ rational and } \sum r_i \leq 1\}$$

is easily seen to be a basis for $\mathcal{V}(P)$ since it is clear that we can approximate any linear combination of point evaluations by one with rational coefficients, and to get a well-below set for any linear combination of point evaluations, we only need to use basis elements. Further this the set is clearly countable. Hence $\mathcal{V}(P)$ is ω-continuous. ∎

5.2 Continuous Ipos and Abstract Probabilistic Domains

In this section we will return to the definition of an abstract probabilistic domain and its comparison via a forgetful functor to the $\mathcal{V}$-algebras as described in Section 4.2. We will give a concrete proof of the free algebra property for abstract probabilistic domains and from this show that the forgetful functor is the identity when restricted to continuous ipos. We will need the theorems developed in Section 4.4, particularly the Splitting Lemma.

Recall the definition of an *abstract probabilistic domain* as as ipo A with $\perp$ and a continuous function $+_r : A \times A \times [0,1] \to A$ that satisfies associativity, commutativity,one and absorption laws as follows

$$\begin{aligned}
(a +_r b) +_s c &= a +_{rs} (b +_{\frac{s(1-r)}{1-rs}} c) \quad (rs \neq 1) \\
a +_r b &= b +_{1-r} a \\
a +_1 b &= a \\
a +_r a &= a
\end{aligned}$$

We considered the category of such ipos with morphisms as continuous maps which respected $+_r$ i.e. satisfied

$$\forall a, b, r \quad f(a +_r b) = f(a) +_r f(b)$$

and preserved $\perp$, and showed that there was a forgetful functor U to **Apd** from $\mathbf{Ipo}^{\mathcal{V}}$.

We will now construct n-ary operators $\sum_{i=1}^{n} r_i a_i$ on an abstract probabilistic domain A, for elements a_i of A and r_i real numbers such that $\sum_{i=1}^{n} r_i \leq 1$, and prove some facts about these operators; we do not claim that this is an alternative axiomatisation although it could be made to be with some more work.

We first assume that $\sum_{i=1}^{n} r_i = 1$ and define $\sum_{i=1}^{n}$ by induction on n. If $n = 1$ then since $\sum_{i=1}^{1} r_i = 1$ i.e. $r_1 = 1$ then we define $1a_1$ to be a_1. Given the definition for $\sum_{i=1}^{n-1} r_i a_i$, and assuming $r_1 < 1$ we define

$$\sum_{i=1}^{n} r_i a_i = a_1 +_{r_1} \left(\sum_{i=2}^{n} \frac{r_i}{(1 - r_1)} a_i \right)$$

and if $r_1 = 1$ then all the other coefficients must be zero and we define $\sum_{i=1}^{n} r_i a_i = a_1$. In the first equation note that $1 - r_1 = \sum_{i=2}^{n} r_2$ so dividing each coefficient by $1 - r_1$ ensures the inside term is a sum whose coefficients add up to 1. Also note that if $n = 2$ this gives $a_1 +_{r_1} a_2$.

For $\sum_{i=1}^{n} r_i a_i$ where $\sum_{i=1}^{n} r_i < 1$, define

$$\sum_{i=1}^{n} r_i a_i = \sum_{i=1}^{n} \frac{r_i}{\sum_{j=1}^{n} r_j} a_i +_{\sum_{j=1}^{n} r_j} \perp,$$

unless $\sum_{i=1}^{n} r_i = 0$ in which case $\sum_{i=1}^{n} r_i a_i = \perp$.

Now note that we can consider $\sum_{i=1}^{n}$ as an operation from $G \times A^n$ where G is the subset of $[0, 1]^n$ such that $\sum_{i=1}^{n} r_i \leq 1$ and we will show that it is continuous in this sense, with $G \times A^n$ given the Scott topology of the product and with G given the Scott topology as a subset of the ipo $[0, 1]^n$. Here we don't use the Hausdorff topology on $[0, 1]$.

Lemma 5.6 *The definition above of $\sum_{i=1}^n r_i a_i$ is continuous as a function $\sum_{i=1}^n : G \times A^n \to A$ and satisfies the following:*

$$\sum_{i=1}^n r_i \left(\sum_{j=1}^{n_i} r_j^i a_j^i \right) = \sum_{i,j} r_i r_j^i a_j^i \tag{5.2}$$

$$\sum_{i=1}^n r_i a_i = \sum_{i=1}^n r_{\pi(i)} a_{\pi(i)} \quad \textit{for any permutation } \pi \textit{ of } \{1, \ldots, n\} \tag{5.3}$$

$$\sum_{i=1}^n r_i a_i = (r_1 + r_2) a_1 + \sum_{i=3}^n r_i a_i \quad \textit{when } a_1 = a_2 \tag{5.4}$$

$$\sum_{i=1}^n r_i a_i = \sum_{i=2}^n r_i a_i \quad \textit{when } r_1 = 0 \tag{5.5}$$

Proof We first prove the equations. For each equation we can assume that $\sum_{i=1}^n r_i = 1$ (and for the first equation that $\sum_{j=1}^{n_i} = r_j^i = 1$) since otherwise we just apply the proof to $(1 - \sum_{j=1}^n r_j)\perp + \sum_{i=1}^n \frac{r_i}{\sum_{j=1}^n r_j} a_i$, since by the definition and commutativity, $(1 - \sum_{j=1}^n r_j)\perp + \sum_{i=1}^n \frac{r_i}{\sum_{j=1}^n r_j} a_i = \sum_{i=1}^n r_i a_i$

We will first show that for any $\sum_{i=1}^n r_i = \sum_{j=1}^m s_j = 1$,

$$\left(\sum_{i=1}^n r_i a_i \right) +_r \left(\sum_{j=1}^m s_j b_j \right) = \sum_{i=1}^n r r_i a_i + \sum_{j=1}^m (1-r) s_j b_j \tag{5.6}$$

(where the left hand side is the obvious $(m+n)$-ary operator). We use induction on n; the base case, $n = 1$, is trivial. For the induction step consider

$$\sum_{i=1}^n r r_i a_i + \sum_{j=1}^m (1-r) s_j b_j = a_1 +_{r r_1} \left(\sum_{i=2}^n \frac{r r_i}{1 - r r_1} a_i + \sum_{j=1}^m \frac{(1-r) s_j}{1 - r r_1} b_j \right)$$

(clearly if $r r_1 = 1$ the equation is trivial). Then we can rewrite the co-efficients on the left hand side to

$$a_1 +_{r r_1} \left(\sum_{i=2}^n \frac{r(1-r_1)}{1 - r r_1} \frac{r_i}{1 - r_1} a_i + \sum_{j=1}^m (1 - \frac{r(1-r_1)}{1 - r r_1}) s_j b_j \right)$$

and then by the inductive hypothesis this is equal to

$$a_1 +_{r r_1} \left(\sum_{i=2}^n \frac{r_i}{1 - r_1} a_i +_{\frac{r(1-r_1)}{1 - r r_1}} \sum_{j=1}^m s_j b_j \right)$$

which by the commutativity law is equal to

$$(a_1 +_{r_1} \sum_{i=2}^{n} \frac{r_i}{1-r_1} a_i) +_r \left(\sum_{j=1}^{m} s_j b_j\right) = \left(\sum_{i=1}^{n} r_i a_i\right) +_r \left(\sum_{j=1}^{m} s_j b_j\right)$$

as required.

Now we will prove the first equation by induction on n. The case $n = 1$ is trivial. Then

$$\sum_{i=1}^{n} r_i \sum_{j=1}^{n_i} r_j^i a_j^i = \left(\sum_{j=1}^{n_1} r_j^1 a_j^1\right) +_{r_1} \sum_{i=2}^{n} \frac{r_i}{1-r_1} \left(\sum_{j=1}^{n_i} r_j^i a_j^i\right)$$

(and if $r_1 = 1$ the equation is trivial) but by the inductive hypothesis, the left-hand side is equal to

$$\left(\sum_{j=1}^{n_1} r_j^1 a_j^1\right) +_{r_1} \sum_{i>1,j} \frac{r_i r_j^i}{1-r_1} a_j^i.$$

But by Equation 5.6 this is equal to

$$\sum_{j=1}^{n_1} r_1 r_j^1 a_j^1 + \sum_{i>1,j} r_i r_j^i a_j^i$$

which is equal to $\sum_{i,j} r_i r_j^i a_j^i$ as required.

We next show that for any permutation π, that n

$$\sum_{i=1}^{n} r_i a_i = \sum_{i=1}^{n} r_{\pi(i)} a_{\pi(i)}.$$

Since any permutation is generated by "swaps" of two adjacent elements, we can just prove the above for π given by $\pi(k) = k+1$, $\pi(k+1) = k$ and $\pi(i) = i$ for all other i. Calculating from the associativity and commutativity laws we get that $x +_\alpha (y +_\beta z) = y +_{\beta-\alpha\beta} (x +_{\frac{\alpha}{1-\beta+\alpha\beta}} z)$. Since the definition of $\sum_{i=1}^{n} r_i a_i$ is inductive, it is clear that to prove the equation above we need to show that

$$a_k +_{r_k} (a_{k+1} +_{\frac{r_{k+1}}{1-r_k}} \sum_{i=k+2}^{n} \frac{r_i}{K} a_i) = a_{k+1} +_{r_{k+1}} (a_k +_{\frac{r_k}{1-r_{k+1}}} \sum_{i=k+2}^{n} \frac{r_i}{K} a_i)$$

(where K is $\sum_{i=k+2}^{n} r_i$), since the cases where we are dividing by zero are trivial. But just substituting for α and β in the expression above gives the desired result, since $r_{k+1}/(1-r_k) - r_{k+1} r_k/(1-r_k) = r_{k+1}$ and the other term similarly simplifies.

Now we prove the next equation holds. Suppose $a_1 = a_2$, then since by the commutativity law given above $x +_\alpha (x +_\beta z) = (x +_{\frac{\alpha}{\alpha+\beta-\alpha\beta}} y) +_{\alpha+\beta-\alpha\beta} z$ and then by absorption $x +_\alpha (x +_\beta z) = x +_{\alpha+\beta-\alpha\beta} z$ so we see that

$$a_1 +_{r_1} (a_1 +_{\frac{r_2}{1-r_1}} \sum_{i=3}^{n} \frac{r_i}{K} a_i) = a_1 +_{r_1+r_2} \sum_{i=3}^{n} \frac{r_i}{K} a_i,$$

where $K = \sum_{i=3}^n r_i$, since if $\beta = r_2/(1-r_1)$ and $\alpha = r_1$ then $\alpha + \beta - \alpha\beta = r_1 + r_2$.

For the final equation, we use the one law, that $a +_1 b = a$, or, with commutativity, that $a +_0 b = b$. Then if $r_1 = 0$,

$$\sum_{i=1}^{n} r_i a_i = (a_1 +_0 \sum_{i=2}^{n} \frac{r_i}{K} a_i) +_{0+K} \perp = \sum_{i=2}^{n} r_i a_i$$

where $K = \sum_{i=2}^n r_i$.

To show continuity of $\sum_{i=1}^n$ as a function from $G \times A^n$ to A, since all the terms in the product are ipos, we only need to show continuity in each term separately. By Equation 5.3 we only need to show continuity in a_1 and since G is itself a subset of the product $[0,1]^n$ we only need to show continuity in r_1.

To show continuity in a_1, we can assume $\sum_{i=1}^n r_i = 1$ otherwise the proof is by continuity of $x +_{\sum_{i=1}^n r_i} \perp$ in x. Then continuity in a_1 is trivial from continuity in x of $x +_{r_1} \sum_{i=2}^n \frac{r_i}{1-r_1} a_i$.

To show continuity in r_1 is slightly more tricky. We first note that by definition $\sum_{i=1}^n r_i a_i = \sum_{i=1}^n r_i a_i + (1 - \sum_{i=1}^n r_i)\perp$, then we can show monotonicity since supposing that $r \leq s$, then by absorption, $sa_1 + \sum_{i=2}^n r_i a_i + (1 - (s + \sum_{i=1}^n r_i))\perp = (s-r)a_1 + ra_1 + \sum_{i=2}^n r_i a_i + (1 - (s + \sum_{i=1}^n r_i))\perp$ and $ra_1 + \sum_{i=2}^n r_i a_i + (1 - (r + \sum_{i=1}^n r_i))\perp = (s-r)\perp + a_1 + \sum_{i=2}^n r_i a_i + (1 - (s + \sum_{i=1}^n r_i))\perp$ so since $\perp \sqsubseteq a_1$ we see that $ra_1 + \sum_{i=2}^n r_i a_i \sqsubseteq sa_1 + \sum_{i=2}^n r_i a_i$. Then if $\sup_j r_j = r_1$, we know that $\lim_j r_j/(r_j + K) = r_1$ so by the continuity of $+_r$ in r, we know that $\bigsqcup_j (a_1 +_{r_j/(r_j+K)} \sum_{i=2}^n r_i/K a_i) +_{r_j} \perp = \bigsqcup_j (a_1 +_{r_1/(r_1+K)} \sum_{i=2}^n r_i/K a_i) +_{r_1} \perp$. ∎

Now we will prove a lemma relating $\sum_{i=1}^n r_i a_i$ and $\sum_{i=1}^n r_i \eta(a_i)$ using the Splitting Lemma.

Lemma 5.7 *In any abstract probabilistic domain A, $\sum_{i=1}^{n} r_i \eta_A(a_i) \sqsubseteq \sum_{j=1}^{m} s_j \eta_A(b_j)$ implies $\sum_{i=1}^{n} r_i a_i \sqsubseteq \sum_{j=1}^{m} s_j b_j$.*

Proof From the Splitting Lemma we know that $\sum_{i=1}^{n} r_i \eta_A(a_i) \sqsubseteq \sum_{j=1}^{m} s_j \eta_A(b_j)$ implies there exists $t_{i,j}$ such that $t_{i,j} = 0$ unless $a_i \sqsubseteq b_j$, $\sum_{i=1}^{n} t_{i,j} \leq s_j$ and $\sum_{j=1}^{m} t_{i,j} = r_i$.

We use repeated applications of Equation 5.4 to show that $\sum_{i=1}^{n} r_i a_i = \sum_{i,j} t_{i,j} a_i$ and by Equation 5.5 above the zero coefficients vanish, so by monotonicity of $\sum_{i,j}$ $\sum_{i,j} t_{i,j} a_i \sqsubseteq \sum_{i,j} t_{i,j} b_j$. Then by Equation 5.4 again, $\sum_{i,j} t_{i,j} b_j \sqsubseteq \sum_{j=1}^{m} (\sum_{i=1}^{n} t_{i,j}) b_j$ and then by monotonicity in coefficients,

$$\sum_{j=1}^{m} (\sum_{i=1}^{n} t_{i,j}) b_j \sqsubseteq \sum_{j=1}^{m} s_j b_j$$

as required. ∎

Now we give another lemma to show that if f respects $+_r$ and $\perp$, then it respects the sum operations we have just defined. The proof is just a simple induction on the definition of $\sum_{i=1}^{n} r_i a_i$.

Lemma 5.8 *If $f: A \to B$ where A and B are abstract probabilistic domains and $f(a +_r b) = f(a) +_r f(b)$ for any a and b in A and $f(\perp) = \perp$ (i.e. f is an **Apd** morphism), then*

$$f(\sum_{i=1}^{n} r_i a_i) = \sum_{i=1}^{n} r_i f(a_i).$$

Proof We first assume $\sum_{i=1}^{n} r_i = 1$ and work by induction on n. The base case where $n = 1$ is trivial. For the induction step we see

$$f(\sum_{i=1}^{n} r_i a_i) = f(a_1 +_{r_1} \sum_{i=2}^{n} \frac{r_i}{1 - r_1} a_i) = f(a_1) +_{r_1} f(\sum_{i=2}^{n} \frac{r_i}{1 - r_1} a_i)$$

since f preserves $+_r$. Then by the inductive hypothesis,

$$f(\sum_{i=2}^{n} \frac{r_i}{1 - r_1} a_i) = \sum_{i=2}^{n} \frac{r_i}{1 - r_1} f(a_i)$$

so we see that

$$f(\sum_{i=1}^{n} r_i a_i) = f(a_1) +_{r_1} \sum_{i=2}^{n} \frac{r_i}{1-r_1} f(a_i) = \sum_{i=1}^{n} r_i f(a_i)$$

as required.

Finally if $\sum_{i=1}^{n} r_i < 1$, then

$$f(\sum_{i=1}^{n} r_i a_i) = f(\sum_{i=1}^{n} \frac{r_i}{\sum_{j=1}^{n} r_j} a_i +_{\sum_{j=1}^{n} r_j} \perp)$$

and by the assumptions on f and the first part,

$$= \sum_{i=1}^{n} \frac{r_i}{\sum_{j=1}^{n} r_j} f(a_i) +_{\sum_{j=1}^{n} r_j} \perp = \sum_{i=1}^{n} r_i f(a_i)$$

as required. ∎

We will now give a free algebra theorem for abstract probabilistic domains; we will show that any continuous map $f: P \to A$ where P is a continuous ipo and A an abstract probabilistic domain extends to an abstract probabilistic domain morphism $\bar{f}: \mathcal{V}(P) \to A$.

Theorem 5.9 *If P is a continuous ipo and A an abstract probabilistic domain with $f: P \to A$ a continuous map, there exists a unique **Apd** morphism $\bar{f}: \mathcal{V}(P) \to A$ such that $f = \bar{f} \circ i$.*

Proof We define

$$\bar{f}(\mu) = \bigsqcup\{\sum_{i=1}^{n} r_i f(x_i) \mid \sum_{i=1}^{n} r_i \eta_P(x_i) \ll \mu\}.$$

We know that the set of $\sum_{i=1}^{n} r_i \eta_P(x_i) \ll \mu$ is directed by Theorem 5.4, so the set $\sum_{i=1}^{n} r_i \eta_A(f(x_i))$ is directed, (as we have just applied $\mathcal{V}(f)$) and then by Lemma 5.7, $\sum_{i=1}^{n} r_i f(x_i)$ is also directed, so it has a lub. So $\bar{f}$ is well defined.

Pick any μ. Let A be any directed set of linear combinations of point evaluations way below μ with lub μ. Then

$$\bigsqcup\{\sum_{i=1}^{n} r_i f(x_i) \mid \sum_{i=1}^{n} r_i \eta_P(x_i) \in A\} = \bar{f}(\mu) \tag{5.7}$$

since for any other $\sum_{i=1}^{n} r_i \eta_P(x_i) \ll \mu$ which is not in A there exists $\sum_{i=1}^{n} r_i \eta_P(x_i) \sqsubseteq \sum_{j=1}^{m} s_j \eta_P(y_i)$ in A, hence $\sum_{i=1}^{n} r_i f(x_i) \sqsubseteq \sum_{j=1}^{m} s_j f(y_i)$.

It is clear that $\bar{f}$ is continuous since if $\bigsqcup \mu_i = \mu$ and μ_i is directed, then the set of linear combinations of point evaluations way below μ is just the union over all i of the set of linear combinations of point evaluations way below μ_i and, where all sups are over directed sets, if $\bigsqcup_j a_{i,j} = a_i$ and $\bigsqcup a_i = a$, then $\bigsqcup_{i,j} a_{i,j} = a$.

We need to show that $\bar{f}$ respects $+_r$. Now suppose that $\mu = \mu_1 +_r \mu_2$, that is $\mu = r\mu_1 + (1-r)\mu_2$ then

$$\mu = \bigsqcup \{ r\eta_1 + (1-r)\eta_2 \mid \eta_i \ll \mu_i, \eta_i \text{ a linear combination of point evaluations} \}$$

since $\mu_i = \bigsqcup \eta_i$ and $+_r$ is continuous. Then from Equation 5.7 above, we know that

$$\bar{f}(\mu) = \bigsqcup \{ \sum_{i=1}^{n} r r_i f(x_i) + \sum_{j=1}^{n} (1-r) s_j f(y_j) \mid \sum_{i=1}^{n} r_i \eta_P(x_i) \ll \mu_1, \sum_{j=1}^{n} s_j \eta_P(y_j) \ll \mu_2 \}.$$

But from Equation 5.6 we know that

$$\sum_{i=1}^{n} r_i f(x_i) +_r \sum_{j=1}^{n} s_j f(y_j) = \sum_{i=1}^{n} r r_i f(x_i) + \sum_{j=1}^{n} (1-r) s_j f(y_j).$$

so

$$\begin{aligned} \bar{f}(\mu) &= \bigsqcup \{ \sum_{i=1}^{n} r_i f(x_i) \mid \sum_{i=1}^{n} r_i \eta_P(x_i) \ll \mu_1 \} +_r \bigsqcup \{ \sum_{j=1}^{n} s_j f(y_j) \mid \sum_{j=1}^{n} s_j \eta_P(y_j) \ll \mu_2 \} \\ &= \bar{f}(\mu_1) +_r \bar{f}(\mu_2) \end{aligned}$$

as required.

Finally we need to show that $\bar{f}(\eta_P(x)) = f(x)$. We note that $(1-2^{-n})\eta_P(y) \ll \eta_P(x)$ for any $n \geq 1$ and $y \ll x$ and further that $\bigsqcup_{n,y}(1-2^{-n})\eta_P(y) = \eta_P(x)$. So by Equation 5.7, $\bar{f}(\eta_P(x)) = \bigsqcup_{n,y}(1-2^{-n})f(y)$ and by the continuity of rx, this is equal to $1f(x) = f(x)$ as required. ∎

With this theorem we can show that a continuous ipo which is an abstract probabilistic domain can be given a map $\alpha: \mathcal{V}(A) \to A$ such that (A, α) is an $\mathcal{V}$-algebra.

Theorem 5.10 *Every continuous abstract probabilistic domain is an $\mathcal{V}$-algebra.*

Proof Let A be continuous abstract probabilistic domain. We consider the identity function $\mathrm{id}_A: A \to A$ and apply Theorem 5.9 above to obtain a continuous function $\alpha: \mathcal{V}(A) \to A$. From Lemma 5.8 we know that α is linear with respect to finite sums.

Now we have to show that α satisfies the diagrams in Section 2.2.

To see the first diagram is satisfied we need to show $\alpha(\eta_A(x)) = x$. But this we already know from the fact that $\alpha \circ \eta_A = \mathrm{id}_A$.

For the second diagram we need to prove that for all σ in $\mathcal{V}^2(A)$,

$$\alpha(\mu_A(\sigma)) = \alpha(\mathcal{V}(\alpha)\sigma).$$

Since A is continuous we know $\mathcal{V}^2(A)$ is continuous, so since each side of this equation is a continuous function of σ, we only need to prove this for σ a linear combination of point evaluations. So suppose that $\sigma = \sum_{i=1}^n r_i\eta(\nu_i)$. By definition, $\mu_A(\sigma) = \sum_{i=1}^n r_i\nu_i$ and similarly $\mathcal{V}(\alpha)(\sigma) = \sum_{i=1}^n r_i\eta(\alpha(\nu_i))$. So we need to prove that

$$\alpha(\sum_{i=1}^n r_i\nu_i) = \alpha(\sum_{i=1}^n r_i\eta(\alpha(\nu_i)))$$

but since α is linear with respect to finite sums, so

$$\alpha(\sum_{i=1}^n r_i\eta(\alpha(\nu_i))) = \sum_{i=1}^n r_i\alpha(\eta(\alpha(\nu_i)))$$

and by the fact that $\alpha \circ \eta_A = \mathrm{id}_A$, this is equal to $\sum_{i=1}^n r_i\alpha(\nu_i)$ which equals $\alpha(\sum_{i=1}^n r_i\nu_i)$ as required, again by the linearity of α. ∎

This theorem defines a map from the objects in the category **Apd** to $\mathbf{Ipo}^{\mathcal{V}}$. Note that this map is the inverse to the one given by the forgetful functor U. To see this suppose we have a continuous ipo A and a sum operation $+_r$. If we define the function α as in the theorem above to make A an $\mathcal{V}$-algebra and then use this to define another operation $\oplus_r$ we get back to the original since $\oplus_r$ is defined by

$$a \oplus_r b = \alpha(r.\eta_a + (1-r).\eta_b)$$

and

$$\alpha(r.\eta_a + (1-r).\eta_b) = r.a + (1-r).b = a +_r b$$

from the definition of α above. Similarly if A is an $\mathcal{V}$-algebra with a map α and we define a sum operation as above and use this to derive a function α' say, then

$$\alpha'(\sum_{i=1}^{n} a_i \eta_{x_i}) = \sum_{i=1}^{n} a_i x_i = \alpha(\sum_{i=1}^{n} a_i \eta_{x_i}).$$

We now compare the morphisms in the two categories. We have already seen that a morphism of $\mathcal{V}$-algebras is a morphism in the category of abstract probabilistic domains; we now show that if P is continuous, then a morphism in **Apd** is a $\mathcal{V}$-algebra morphism.

Theorem 5.11 *For abstract probabilistic domains P and Q where P is continuous, a continuous function $f: P \to Q$ which preserves $\perp$ and is respects $+_r$ is a morphism in **Ipo**$^{\mathcal{V}}$.*

Proof Recall the diagram on page 26, which defines morphisms in **Ipo**$^{\mathcal{V}}$.

To show this diagram commutes we need to show for all σ in $\mathcal{V}(P)$, the equation

$$f(\alpha(\sigma)) = \beta(\mathcal{V}(f)\sigma).$$

Since each side is a continuous function of σ and P is a continuous ipo, it is sufficient to prove this for σ as a sum of point evaluations.

Suppose $\sigma = \sum_{i=1}^{n} a_i \eta_P(x_i)$ then $\alpha(\sigma) = \sum_{i=1}^{n} a_i x_i$ by definition. Further $\mathcal{V}(f)(\sigma) = \sum_{i=1}^{n} a_i \eta_Q(f(x_i))$, and similarly $\beta(\mathcal{V}(f)(\sigma)) = \sum_{i=1}^{n} a_i f(x_i)$, so the above equation becomes

$$f(\sum_{i=1}^{n} a_i x_i) = \sum_{i=1}^{n} a_i x_i$$

but by Lemma 5.8, since f is preserves $+_r$ and $\perp$, then this equation holds. ∎

5.3 Extending Evaluations to Measures

In this section we will prove a slightly stronger version of Saheb-Djahromi's theorem that the lub of a directed set of linear combinations of point evaluations extends to a measure [35]. His theorem was proved for increasing sequences of evaluations although the extension is easy. With this theorem, we will be able to prove another corollary to the main result of the chapter, Theorem 5.2, stating that every continuous evaluation on a continuous ipo extends uniquely to a measure.

First we introduce some notation, borrowed from Saheb-Djahromi [35]. Given two *finite* subsets of an ipo P, say K and L, then we write $K \sqsubseteq L$ to mean that for every b in L, there exists an a in K such that $a \sqsubseteq b$. Also given a finite linear combination of point evaluations $\eta = \sum_{a \in A} r_a \eta(a)$, then for any for any subset K of P we write $\eta(K)$ to mean $\sum_{a \in K} r_a$.

Lemma 5.12 *Let P be an ipo and $U \setminus V$ be a crescent in P. Pick some $\epsilon > 0$. Suppose that η and ν are linear combination of point evaluations, say $\eta = \sum_{a \in A} r_a \eta(a)$ and $\nu = \sum_{b \in B} s_b \eta(b)$ such that $\eta \sqsubseteq \nu$, $\nu(U) - \eta(U) < \epsilon/4$ and $\nu(P) - \eta(P) < \epsilon/4$. We set $\alpha = \eta(U \setminus V)$, then if T is any subset of $U \setminus V$ such that $\eta(T) = 0$ and $\nu(T) \geq \alpha$ then there exists $K \subseteq B \cap T$ such that $\nu(K) \geq \alpha - \epsilon/2$ and $(A \cap (U \setminus V)) \sqsubseteq K$ Further if $L \subseteq A \cap (U \setminus V)$ with $\eta(L) = \beta$, then there exists some $M \subseteq B \cap T$ such that $L \sqsubseteq M$ and $\nu(M) \geq \beta - \epsilon$.*

Proof Suppose $\eta = \sum_{a \in A} r_a \eta(a)$ and $\nu = \sum_{b \in B} s_b \eta(b)$ satisfy the assumptions above for some crescent $U \setminus V$ and T. Then since $\eta \sqsubseteq \nu$, by the Splitting Lemma 4.10 we can find a set of values $t_{a,b}$ such that $t_{a,b} = 0$ unless $a \sqsubseteq b$ and such that $\sum_{a \in A} t_{a,b} \leq s_b$ and $\sum_{b \in B} t_{a,b} = r_a$.

We will first show that for any subset $H \subseteq B$,

$$\sum_{b \in H} s_b \leq \sum_{b \in H} \sum_{a \in A} t_{a,b} + \epsilon/4 \tag{5.8}$$

By substituting for r_a, we see that $\nu(P)-\eta(P) = \sum_{b\in B}(s_b - \sum_{a\in A} t_{a,b})$ so $\sum_{b\in B}(s_b - \sum_{a\in A} t_{a,b}) < \epsilon/4$. But since $H \subseteq B$, $\sum_{b\in H}(s_b - \sum_{a\in A} t_{a,b}) < \epsilon/4$, so $\sum_{b\in H} s_b < \sum_{b\in H}\sum_{a\in A} t_{a,b} + \epsilon/4$.

Now we will prove a useful fact about subsets $H \subseteq B \cap U$. We know that $\nu(U) - \eta(U) < \epsilon/4$ hence $\sum_{b\in B\cap U} s_b - \sum_{a\in B\cap U}\sum_{b\in B} t_{a,b} < \epsilon/4$. Now $\sum_{a\in A\cap U}\sum_{b\in B} t_{a,b} = \sum_{a\in A\cap U}\sum_{b\in B\cap U} t_{a,b}$ since $t_{a,b} \neq 0$ implies $a \sqsubseteq b$ and so $a \in U$ implies $b \in U$. Also $\sum_{b\in B\cap U} s_b \geq \sum_{b\in B\cap U}\sum_{a\in A} t_{a,b}$ so $\sum_{b\in B\cap U}\sum_{a\in A} t_{a,b} - \sum_{b\in B\cap U}\sum_{a\in A\cap U} t_{a,b} < \epsilon/4$. So $\sum_{b\in B\cap U}\sum_{a\in A\setminus U} t_{a,b} < \epsilon/4$. Therefore for any subset $H \subseteq B \cap U$,

$$\sum_{b \in H} \sum_{a \in A \setminus U} t_{a,b} < \epsilon/4 \tag{5.9}$$

Now let $S = \{b \in B \cap (U \setminus V) \mid t_{a,b} \neq 0 \Rightarrow a \in A \setminus U\}$. We set $K = B \cap T \setminus S$, and to prove the lemma we need to show that $\nu(K) > \alpha - \epsilon$, and that $(A \cap (U \setminus V)) \sqsubseteq K$. For the first condition we use the Equation 5.8 see that $\sum_{b\in S} s_b < \sum_{b\in S}\sum_{a\in A} t_{a,b} + \epsilon/4$. But if $b \in S$ and $t_{a,b} \neq 0$ this implies $a \in A \setminus U$, hence $\sum_{b\in S}\sum_{a\in A} t_{a,b} = \sum_{b\in S}\sum_{a\in A\setminus U} t_{a,b}$. Then applying Equation 5.9, we see that $\sum_{b\in S}\sum_{a\in A\setminus U} t_{a,b} < \epsilon/4$ so $\sum_{b\in S} s_b < \epsilon/2$. Now $\sum_{b\in K} s_b = \sum_{b\in T} s_b - \sum_{b\in(T\cap S)} s_b \leq \alpha - \sum_{b\in S} s_b < \alpha - \epsilon/2$ as required. For the second part, note that $b \in K$ implies $b \notin S$, hence there exists a such that $t_{a,b} \neq 0$ and $a \in A \cap U$, hence $a \sqsubseteq b$ and $a \in U \setminus V$.

For the next part, we assume L is a subset of $A \cap (U \setminus V)$, such that $\eta(L) = \beta$, clearly $\beta \leq \alpha$. We set $S' = \{b \in B \cap (U \setminus V) \mid \forall a \in L, t_{a,b} = 0\}$ and define $M = K \setminus S'$. We now need to show that $\nu(M) \geq \beta - \epsilon$ and also that $M \sqsubseteq L$. For the first condition we note that $\sum_{b\in M} s_b = \sum_{b\in K} s_b - \sum_{b\in(S'\cap K)} s_b$. By the first part of the lemma, $\sum_{b\in K} s_b > \alpha - \epsilon/2$ so it remains to show that $\sum_{b\in(S'\cap K)} s_b < (\alpha - \beta) + \epsilon/2$. By Equation 5.8, $\sum_{b\in S'} s_b < \sum_{b\in S'}\sum_{a\in A} t_{a,b} + \epsilon/4$. But $b \in S'$ implies $t_{a,b} = 0$ if

$a \in L$ so for $b \in S'$, $\sum_{a \in A} t_{a,b} = \sum_{a \in A \setminus U} t_{a,b} + \sum_{a \in (U \setminus V) \setminus L} t_{a,b}$. By Equation 5.9, this first sum $\sum_{b \in S'} \sum_{a \in A \setminus U} t_{a,b} < \epsilon/4$ and the second sum $\sum_{b \in S'} \sum_{a \in (U \setminus V) \setminus L} t_{a,b} \leq \sum_{a \in (U \setminus V) \setminus L} r_a = \alpha - \beta$. For the second condition, note that $b \in M$ implies $b =\notin S'$ and $b \in B \cap (U \setminus V)$ hence there exists $a \in L$ such that $t_{a,b} \neq 0$ i.e. $a \sqsubseteq b$ as required. ∎

We now prove the theorem that the lub of a directed set of linear combinations of point evaluations extends to a measure. This proof is based on Saheb-Djahromi, [35, lemma 2], extending it to a directed set instead of a chain, and using the lemma above.

Theorem 5.13 *If X is a directed set of linear combinations of point evaluations, then $\bigsqcup X$ extends uniquely to a measure on the Borel sets of X.*

Proof We set $\mu = \bigsqcup X(O) = \sup_{\nu \in X} \nu(O)$ by definition. By Pettis, we can extend μ to a finitely additive set function on the field of finite unions of crescents of X. By Theorem 3.3, if this extension is countably additive then it extends uniquely to a measure. So we need to prove that whenever $\bigcup_{i=1}^{n} A_i = A$ for A and A_i in the ring of finite, disjoint unions of crescents and A_i pairwise disjoint, that $\sum_{i=1}^{\infty} \mu(A_i) = \mu(A)$. We can easily reduce this to the case where A is a crescent, say $A = U \setminus V$, by intersecting each of the A_i with the components of A and then by considering each component of A_i separately we can assume each A_i is also a crescent, say $U_i \setminus V_i$. By the monotonicity of μ, it is clear that $\sum_{i=1}^{\infty} \mu(A_i) \leq \mu(A)$. We now prove the converse by contradiction.

Suppose that $\mu(A) - \sum_{i=1}^{\infty} \mu(A_i) = \alpha > 0$, so for any n, $\mu(A) - \sum_{i=1}^{n} \mu(A_i) \geq \alpha$, i.e. for all n, there exists $\eta \in X$ such that for all ν in X, such that $\eta \sqsubseteq \nu$ $\nu(\bigcup_{i>n} A_i) \geq \alpha$. We will construct a sequence of sets of points $B_1, B_2, \ldots$ and integers $P_1 < P_2 \ldots$ such that for all $x \in B_{n+1}$ there exists $y \in B_n$ such that $y \sqsubseteq x$, that is $B_n \sqsubseteq B_{n+1}$ and such that $B_n \subseteq \bigcup_{i=P_{n-1}+1}^{P_n} A_i$. Then applying König's infinity lemma (e.g. [42]) to these points we get an increasing sequence

x_n with $x_n \in B_n$. Then consider the limit of this sequence $\bigsqcup_n x_n$. It must be in A, since for all n $x_n \in U$ and $x_n \notin V$, so it must be in some A_i, say A_k. But then for sufficiently large n, all x_n must be in $U_k \setminus V_k$ since U_k and V_k are open, but this is impossible, since $x_n \in A_k$ implies $P_{n+1} > n$, hence $x_{n+1} \notin A_k$. So if we can construct such a set of points and integers, we're done.

We are going to want to use the lemma above, so we first note that if O is any open subset of P, then for any $\epsilon > 0$ there exists $\eta \in X$ such that for all $\nu \in X$ such that $\eta \sqsubseteq \nu$, then $\nu(O) - \eta(O) < \epsilon$. This is proved by recalling that $\mu(O) = \sup_{\nu \in X} \nu(O)$, and η in X such that $\mu(O) - \eta(O) < \epsilon$ will satisfy the above.

We start by setting $\epsilon = \alpha/2$ and picking η_1 in X so that for all $\eta_1 \sqsubseteq \nu$, $\nu(U) - \eta_1(U) < \epsilon/4$ and $\nu(P) - \eta_1(P) < \epsilon/4$ and such that $\eta(A) > \alpha$. Then if we let B_1 be the set of points of η and since there can only be finitely many of then, there must be some P_1 such that $B_1 \subseteq \bigcup_{i=1}^{P_1} A_i$. Then we find an η_2 above η_1 such that for all ν above η_2 we have $\nu(\bigcup_{i \geq P_1} U_i) - \eta_2(\bigcup_{i \geq P_1} U_i) < \alpha/4$ and $\nu(P) - \eta_2(P) < \alpha/4$ and applying the lemma with T as $\bigcup_{i > P_1} A_i$ we find a subset B_2 with $\eta_2(B_2) \geq \alpha/2$ and $B_1 \sqsubseteq B_2$. Then P_2 is such that $B_2 \subseteq \bigcup_{i \leq P_2} A_i$. Similarly, suppose we have η_n, B_n and P_n, such that for all ν in X above η_n, $\nu(P) - \eta_n(P) < \alpha 2^{-(n+2)}$ and $\nu(\bigcup_{i > P_{n-1}} A_i) - \eta_n(\bigcup_{i > P_{n-1}} A_i) < \alpha 2^{-(n+2)}$ and $B_n \subseteq \bigcup_{i=P_{n-1}+1}^{P_n} A_i$, in fact P_n is such that η_n has no points in A_i for $i \geq P_n$. Then we find $\eta_{n+1} \in X$ so that first it has value at least α on $\bigcup_{i > P_n} A_i$ and secondly that for all ν above η_{n+1}, both $\nu(P) - \eta_{n+1}(P) < \alpha 2^{-(n+3)}$ and $\nu(\bigcup_{i > P_n} A_i) - \eta_{n+1}(\bigcup_{i > P_n} A_i) < \alpha 2^{-(n+3)}$. We then apply the second part of Lemma 5.12 above to η_{n+1} with $T = \bigcup_{i > P_n} A_i$, $U = \bigcup_{i > P_{n-1}} A_i$, V as in the theorem, and L as B_n to get a subset of points of T with value at least $\alpha 2^{-(n+2)}$ which is the new B_{n+1}. Then since there are only finitely many points in A from η_{n+1}, they are covered by finitely many A_i, so let P_{n+1} be an upper bound of i. ∎

Corollary 5.14 *If X is a continuous ipo, and ν is an continuous evaluation in $\mathcal{V}(X)$, then ν extends uniquely to a measure on the Borel sets of X.*

Proof Use Theorem 5.2 to express ν as the lub of a directed set of point evaluations. Then apply Theorem 5.13 above to see that ν extends uniquely to a measure. ∎

5.4 Concluding Remarks

In this chapter we have proved the result that on continuous ipos, all evaluations are given by directed limits of linear combinations of point evaluations. We have also shown how Graham's finitary algebraic characterisation [14] of the probabilistic powerdomain can be extended to continuous ipos.

The continuity condition is essential for proving Theorem 5.2 (that all evaluations are limits of linear combinations of point evaluations); I conjecture that without this condition, the result fails.

I have been unable to find counter-examples to show that these results cannot be extended to all ipos. I feel that it is possible (although not very likely) that any evaluation on an ipo can be extended to a measure, but highly unlikely that either on any ipo, all measures are generated by linear combinations of point measures or that the finitary algebraic characterisation extends to all ipos. We do know that Lebesgue measure on $[0, 1]$ gives a counter-example to the limits theorem for $[0, 1]$ as a topological space.

The result that all measures are given by limits of linear combinations of point measures was essentially proved for ω-algebraic ipos by Saheb-Djahromi [35]. This can be easily extended to ω-continuous ipos using the fact that they are retractions of ω-algebraic ipos and the fact that the powerdomain is a functor. The result that these theorems lead to, that all evaluations extend, is derivable for ω-continuous ipos from the theorem by Lawson [23], that evaluations extend when the topology is second countable, although this result uses the continuity condition on measures.

Chapter 6

Duality and Probabilistic Ipos

This chapter shows we can form an analogy to the Stone dualities, where sets of functions to the interval $[0,1]$ replace sets of functions to the two point space $\{0,1\}$. We give a duality theorem for the category $\mathbf{Ipo}_{\mathcal{V}}$ of ipos and continuous maps $P \to \mathcal{V}(Q)$.

6.1 Preliminary Remarks

The Stone Duality Theorem is formed by taking sets of appropriate functions to the two point space $\{0,1\}$. Here we investigate the dualities that arise from taking the sets of functions to the continuous analogue of the two-point space, namely $[0,1]$. This is motivated in part from the natural isomorphism in finite dimensional vector spaces between a space and its double dual, where the dual of a vector space is the set of functionals (linear continuous maps $X \to \Re$). The isomorphism uses the observation that "evaluate at x" for a fixed x is a functional from the double dual.

We will use the upper-set topology on $[0,1]$ and the set of upper continuous functions from an ipo to $[0,1]$. This topology is the Scott topology which arises from the natural ordering on $[0, 1]$. Also $[0, 1]$ with $\leq$ is isomorphic to the ipo of evaluations on the singleton set, thus it is the simplest probabilistic powerdomain.

We will get a theorem similar to the Riesz Representation theorem (which says any functional on the space of functions $X \to \Re$ on a compact, Hausdorff X is integration by some measure); we will show that every functional on the set of upper-continuous maps is integration w.r.t. some evaluation.

6.2 The Duality

For any category **C** and any object A in **C** the contravariant hom-set functor on A takes an object to the set of **C** morphisms from B to A (an object in **Set**) and takes a morphism $f: B \to C$ to the function from the set of morphisms from C to A to the set of morphisms from B to A by composition with f (that is $h: C \to A \mapsto h \circ f: B \to A$). The image of **C** in **Set** under this functor is clearly dual to **C** since the objects and morphisms are mapped to distinct sets and functions. One way to find a category dual to **C** is to find a concrete representation of this image.

We will use this idea to find a dual category to $\mathbf{Ipo}_{\mathcal{V}}$. Recall that this category has the same objects as **Ipo** and that a morphism in $\mathbf{Ipo}_{\mathcal{V}}$ from P to Q is a morphism in **Ipo** from P to $\mathcal{V}Q$. We consider the image of $\mathbf{Ipo}_{\mathcal{V}}$ under the contravariant hom-set functor on $\mathbb{1}$ (the one point ipo) to this category.

The $\mathbf{Ipo}_{\mathcal{V}}$ morphisms $P \to \mathbb{1}$ are the **Ipo** morphisms $P \to \mathcal{V}\mathbb{1}$, but we know from Section 4.5 that $\mathcal{V}\mathbb{1}$ is isomorphic to the ipo $[0, 1]$ with the partial order $\leq$ hence each $\mathbf{Ipo}_{\mathcal{V}}$ morphism $P \to \mathbb{1}$ can be uniquely associated with an upper continuous function $P \to [0, 1]$. Furthermore each $\mathbf{Ipo}_{\mathcal{V}}$, $f: P \to Q$ induces a transformation from the upper continuous functions $Q \to [0, 1]$ to the upper continuous functions $P \to [0,1]$ given by

$$f^\circ(h: Q \to [0,1]) = \lambda x: P. \int h \, df(x)$$

since in the Kleisli category,

$$h \circ f\,(x) = O \mapsto \int_{y \in Q} h(y)(O)\, df(x)$$

and we transform a measure μ in $\mathcal{V}\mathbb{1}$ to an element of $[0,1]$ by evaluating $\mu(\mathbb{1})$, so $h \circ f\,(x)(\mathbb{1}) = \int h(y)(\mathbb{1})\, df(x)$. We will show that distinct (non-equal) $\mathbf{Ipo}_{\mathcal{V}}$ morphisms $P \to Q$ under the hom-set functor on $\mathbb{1}$ give distinct functions in **Set**. This is equivalent to showing that in $\mathbf{Ipo}_{\mathcal{V}}$ for any $f, g: P \to Q$, if for all $h: Q \to \mathbb{1}$, $h \circ f = h \circ g$ then $f = g$.

Lemma 6.1 *In $\boldsymbol{Ipo}_{\mathcal{V}}$ if $f, g: P \to Q$ are such that for all $h: Q \to \mathbb{1}$, $h \circ f = h \circ g$ then $f = g$.*

Suppose that $f, g: P \to Q$ are $\mathbf{Ipo}_{\mathcal{V}}$ morphisms which satisfy the above.

As previously noted, we know

$$h \circ f(x) = O \mapsto \int_{y \in Q} h(y)(O)\, df(x)$$

and

$$h \circ g(x) = O \mapsto \int_{y \in Q} h(y)(O)\, dg(x).$$

We choose any x in P and any open subset O of Q, we want to show that $f(x)(O) = g(x)(O)$. But the function χ_O is an upper continuous function $Q \to [0,1]$ and is associated with the $\mathbf{Ipo}_{\mathcal{V}}$ morphism $h: Q \to [0,1]$ given by $h(x) = O \mapsto 0$ for x not in O and $h(x) = \eta_{\mathbb{1}}(1)$ for x in O. Furthermore we can see that $h \circ f(x)(\mathbb{1}) = f(x)(O)$, so since $h \circ f = h \circ g$ then $f(x)(O) = g(x)(O)$ as required. ∎

We can now get an "external" description of the image as a category we will denote by **F**. The objects of **F** are the sets of upper continuous functions $h: P \to$

$[0,1]$ for every ipo P. Each object is a cHa (from Section 3.8) and also an ipo with the pointwise ordering. Also the objects are closed under taking (pointwise) finite linear combinations $\sum_{i=1}^{n} r_i h_i$ where $\sum_{i=1}^{n} r_i \leq 1$. We will denote the set of upper continuous functions $P \to [0,1]$ by $F(P)$. We have seen above how $\mathbf{Ipo}_{\mathcal{V}}$ morphisms induce functionals on these cHas, we shall now show that the functionals which preserve the pointwise order and finite sums are exactly those which are images of $\mathbf{Ipo}_{\mathcal{V}}$ morphisms. Conversely, we will also show that the functionals induced by a $\mathbf{Ipo}_{\mathcal{V}}$ morphism also have these properties.

Theorem 6.2 *Suppose g is a functional from the set of upper continuous functions $P \to [0,1]$ to $Q \to [0,1]$ which preserves the pointwise order and finite sums. Then g is the image of a **Ipo**$_{\mathcal{V}}$ morphism under the hom-set functor as described above. Conversely the functional induced by any morphism preserves limits and finite sums as well.*

Suppose g is such a function. Define $g^\star\colon Q \to \mathcal{V}P$ as follows;

$$g^\star(y) = \lambda O.g(\chi_O)(y)$$

where y is in Q and O ranges over open subsets of P. We first need to show $g^\star(y)$ is a continuous evaluation.

Since g is linear, $g(\chi_\emptyset)(y) = 0$. Then by linearity, $g(1/2\,\chi_U + 1/2\,\chi_V) = 1/2\,g(\chi_U) + 1/2\,g(\chi_V)$ but $1/2\,\chi_U + 1/2\,\chi_V = 1/2\,\chi_{U\cap V} + 1/2\,\chi_{U\cup V}$ so $1/2\,g^\star(y)(U) + 1/2\,g^\star(y)(V) = 1/2\,g^\star(y)(U\cap V) + 1/2\,g^\star(y)(U \cup V)$ i.e. $g^\star(y)$ is modular.

If $U \subseteq V$, then $\chi_U \sqsubseteq \chi_V$ so $g(\chi_U) \sqsubseteq g(\chi_V)$ so for any y, $g(\chi_U)(y) \leq g(\chi_V)(y)$ i.e. $g^\star(y)$ is monotone.

For continuity, we suppose that $O = \bigcup_{i\in I} O_i$ where this union is directed. Then the functions χ_{O_i} where i ranges over I are directed and have sup χ_O, so for any y, $\sup_{i\in I} g(\chi_{O_i})(y) = g(\chi_O)(y)$, so $g^\star(y)$ is a continuous evaluation.

Now we need to show that $g^\star$ is continuous as a function $Q \to \mathcal{V}P$. First we note it is monotone since for any open set O, the function $g(\chi_O)$ is upper continuous, hence if $y \sqsubseteq y'$, then $g^\star(y)(O) \leq g^\star(y')(O)$ so $g^\star(y) \sqsubseteq g^\star(y')$. Similarly, since $g(\chi_O)$ is continuous, if X is a directed subset of Q with lub y, then

$$\sup_{z\in X} g^\star(z)(O) = \sup_{z\in X} g(\chi_O)(z) = g(\chi_O)(y) = g^\star(y)(O)$$

so $g^\star$ is continuous.

We now need to show that g is the image of $g^\star$ under the hom-set functor, that is $(g^\star)^\circ = g$. We need to show that for all upper continuous functions $h\colon P \to [0,1]$,

$$g(h) = \lambda y. \int h \, dg^\star(y)$$

or, substituting the definition of $g^\star$, that

$$g(h) = \lambda y. \int h \, d(O \mapsto g(\chi_O)(y)).$$

Now if h is a step function, say $h = \chi_U$, then the left-hand side of the equation above reduces to $g(\chi_U)$, so the equation holds for step functions. Furthermore, since both sides of the equation are finitely linear and continuous in h, it must hold for any upper continuous function h. So we see that g is the image of $g^\star$ under the hom-set functor.

Finally we now show that the image of any continuous function $f\colon Q \to \mathcal{V}P$ gives a finitely linear and continuous functional on the sets of upper continuous functions. The functional is given by

$$h \mapsto \lambda y. \int h \, df(y)$$

so

$$\sum_{i=1}^n r_i h_i \mapsto \lambda y. \int \sum_{i=1}^n r_i h_i \, df(y) = \lambda y. \sum_{i=1}^n r_i \int h_i \, df(y) = \sum_{i=1}^n r_i \lambda y. \int h_i \, df(y)$$

and we see that the functional is finitely linear. Similarly if h_i over $i \in I$ is a directed collection of upper continuous functions, then by the directed monotone convergence theorem for evaluations,

$$\sup_{i\in I}\int h_i\, df(y) = \int \sup_{i\in I} h_i\, df(y)$$

so

$$\sup_{i\in I}\lambda y.\int h_i\, df(y) = \lambda y.\sup_{i\in I}\int h_i\, df(y) = \lambda y.\int \sup_{i\in I} h_i\, df(y)$$

so the functional is continuous. Furthermore, $(f^\circ)^\star = f$ since

$$(f^\circ)^\star(y) = O \mapsto \int \chi_O\, df(y) = f(y).$$

∎

So we can conclude that the category $\mathbf{Ipo}_{\mathcal{V}}$ is dual to the category $\mathbf{F}$ whose objects are the sets of upper continuous functions and morphisms are the continuous, linear functionals between them. The dualities are given explicitly by $f\colon P \to \mathcal{V}Q \mapsto f^\circ\colon F(Q) \to F(P)$ and $g\colon F(P) \to F(Q) \mapsto g^\star\colon Q \to \mathcal{V}P$.

6.3 The Category F

We look further at the category $\mathbf{F}$ described above. We will use another hom-set functor to get an image in $\mathbf{Set}$ which is isomorphic to the original category $\mathbf{Ipo}_{\mathcal{V}}$. Furthermore the duality derived from this isomorphism is the same as the one given in the previous section.

The simplest object in $\mathbf{F}$ is the one associated with the one point ipo $\mathbb{1}$, and consists of the upper continuous functions $\mathbb{1} \to [0,1]$, which is clearly isomorphic to $[0,1]$. We apply the hom-set functor on $\mathbb{1}$ to $\mathbf{F}$. As previously discussed, the image of this functor in $\mathbf{Set}$ is dual to $\mathbf{F}$, therefore by general principles, it is isomorphic to $\mathbf{Ipo}_{\mathcal{V}}$, since the dual of a category is obtained by just reversing the direction of the arrows.

In fact we will show that this image is isomorphic to $\mathbf{Ipo}_{\mathcal{V}}$ in a more concrete way. We will show that each element of the set of morphisms $F(P) \to F(\mathbb{1})$ corresponds to an evaluation in $\mathcal{V}P$ and that given a morphism $g: F(Q) \to F(P)$ which is the image of $g^{\star}: P \to \mathcal{V}Q$, the corresponding function $\mathcal{V}P \to \mathcal{V}Q$ is equal to $(g^{\star})^{\dagger}$. In other words, the image in **Set** can be given an "external" definition as the category with objects $\mathcal{V}P$ for any ipo P and the super-linear, continuous functions between these objects. This category is clearly isomorphic to $\mathbf{Ipo}_{\mathcal{V}}$.

Theorem 6.3 *The image of **F** under the hom-set functor on $\mathbb{1}$ is isomorphic to the category with objects $\mathcal{V}P$ for any ipo P and the super-linear maps $f: \mathcal{V}P \to \mathcal{V}Q$.*

Objects in the image of **F** under this functor are the sets of **F** morphisms $F(P) \to F(\mathbb{1})$. But from the first duality, every **F** morphism $F(P) \to F(\mathbb{1})$ corresponds to an $\mathbf{Ipo}_{\mathcal{V}}$ morphism $\mathbb{1} \to P$, that is, a continuous function $\mathbb{1} \to \mathcal{V}P$ which naturally gives an element of $\mathcal{V}P$.

Now we consider the image of a morphism $f: F(P) \to F(Q)$ under the hom-set functor as a function on the underlying sets $\mathcal{V}Q \to \mathcal{V}P$. We need to show that this function is continuous and super-linear; this will come automatically if we can show that it is equal to $(f^{\star})^{\dagger}$, since the extension of a continuous function is always super-linear. Recall that a continuous evaluation μ is represented by the dual of the function $\mu: \mathbb{1} \to \mathcal{V}Q$, i.e. μ°. We know f acts by composition with μ°, and gives a function $\mathbb{1} \to \mathcal{V}P$ by applying $\star$, i.e. the evaluation which μ is mapped to is given by $(\mu^{\circ} \circ f)^{\star}$, which is clearly equal to $\mu \circ f^{\star}$, this composition being in the Kleisli category. But then $\mu \circ f^{\star}$ in the Kleisli category is $\mu \circ (f^{\star})^{\dagger}$ i.e. μ is mapped to $(f^{\star})^{\dagger}(\mu)$ as required.

Finally we need to check that every super-linear continuous function $f: \mathcal{V}Q \to \mathcal{V}P$ is the image of a **F** morphism $F(P) \to F(Q)$. But we already know that every such f is the extension of the function $f \circ \eta_Q: Q \to \mathcal{V}P$ and then $(f \circ \eta_Q)^{\circ}: F(P) \to F(Q)$ is an **F** morphism and from the above we see that the image of $(f \circ \eta_Q)^{\circ}$ is f as required. ∎

We note that the duality we derive from the hom-set functor above is the one given by $\star$ and $\circ$, since we show above that the image of a morphism f in $\mathbf{F}$ is the function $(f^\star)^\dagger$ which is isomorphic to $f^\star$ in the Kleisli category.

6.4 Concluding Remarks

It is clear from the above that under relatively weak conditions we can generally find a dual category to the Kleisli category of a monad. The problem is to find a satisfactory "external" characterisation of it.

The original Stone dualities arose as restrictions of adjoint functors; it is not clear how we might enlarge the two categories and get an adjunction. We might try to enlarge $\mathbf{F}$, say to all cHa's with some sort of linearity (perhaps abstract probabilistic domains).

We might be able to enlarge the other category, by extending the powerdomain functor say to topological spaces; the problem in doing this is the topology on $\mathcal{V}(X)$. If X is a topological space, then we have an alternative topology, that generated by sets of the form $U_{O,r} = \{\mu \mid \mu(O) > r\}$ over all open subsets of X and $0 \leq r \leq 1$. It is this topology which has the necessary property that the injection map $x \mapsto \eta(x)$ is continuous. In general this topology is weaker then the Scott topology induced by the partial order. It is easy to see that this topology coincides with the usual one on continuous ipos, using the results in Chapter 5.

Chapter 7

Probabilistic Logic

In Chapter 6 we described a duality between the ipos of evaluations and continuous maps between them and the cHa's of upper continuous functions from an ipo to $[0, 1]$. We hoped that this duality would be useful in relating logic to semantics for a probabilistic language, as was done with the Stone duality by Abramsky [1] for deterministic languages. To exploit this duality we would need to define a logic using the cHa's of functions.

Probably the simplest programming logic is Hoare logic for a simple while language (see e.g. [3]). We add a probabilistic construct, $\mathbf{P}\ \mathtt{or}_p\ \mathbf{Q}$, meaning do $\mathbf{P}$ with probability p or $\mathbf{Q}$ with probability $1-p$ and define a logic of triples $f\,\{\,\mathbf{P}\,\}\,g$ where f and g are functions rather than terms in first order logic. It turns out that for a logic of total program correctness we need the functions to be upper continuous, since they need to be closed under limits of increasing sequences, but for the logic of partial correctness they should be *lower* continuous functions as they need to be closed under limits of decreasing sequences. But since upper and lower continuous functions have a natural 1–1 correspondence, by corresponding to f the function $\lambda x.1 - f(x)$, which we can use.

In this chapter then, we give simple probabilistic while language and a total and partial correctness logic for it. We interpret the logics in terms of a semantics and prove consistency and completeness of them. We give an example of using the logic and then discuss how each logic arises from the duality given in the Chapter 6.

7.1 Preliminaries

Recall from Section 3.8 that a function $f: X \to [0,1]$ is upper continuous if the inverse image of any upper set is open, that is $f^{-1}(a,1]$ is open for any a. In this section we showed that sums and arbitrary sups of upper continuous functions are upper continuous. Furthermore pointwise products of upper continuous functions are upper continuous since

$$(fg)^{-1}(a,1] \bigcup_{s,t\text{s.t. } st=a} f^{-1}(s,1] \cap g^{-1}(t,1]$$

which is clearly open. Similarly we say that a function $f: X \to [0,1]$ is *lower continuous* if for any lower set $[0,a)$, $f^{-1}[0,a)$ is open in X. It is easy to see that if f is upper continuous, then $(1-f)$ is lower continuous and vice versa. So lower continuous functions are closed under arbitrary infs.

We define a simple while language by:

$$\mathbf{P} ::= \texttt{skip} \mid \mathrm{a}{:=}\ \mathbf{E} \mid \mathbf{P}\ \texttt{or}_p\ \mathbf{Q} \mid \mathbf{P};\mathbf{Q} \mid \texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q} \mid \texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P}$$

For assignment, we will assume a probabilistic denotational function $\mathcal{E}: Exp \times S \to \mathcal{V}(Vals)$ where Exp is a set of expressions, and $Vals$ is an ipo of values, $\mathcal{E}(\mathbf{E})$ being continuous for any expression $\mathbf{E}$. Furthermore we assume a function "update", which takes a state s, a variable a and a value x and produces the state $s[x/a]$ which is the state produced from s by replacing the value of a by x. For fixed x, this is a continuous function $S \times Vals \to S$. Similarly, we assume a continuous function $\mathcal{B}[\![\cdot]\!]$ from boolean expressions $\mathbf{B}$, and states s to $\mathcal{V}(\mathbb{T})$ (where $\mathbb{T}$ is the ipo of truth values, containing tt, ff and $\perp$). We will derive from any expression $\mathbf{B}$ the two functions b and $\neg b$ given by $b(s) = \mathcal{B}[\![\mathbf{B}]\!](s)(\{tt\})$ and $\neg b(s) = \mathcal{B}[\![\mathbf{B}]\!](s)(\{\mathit{ff}\})$. Clearly these are upper continuous functions of s for any $\mathbf{B}$.

We use the idea that a state s satisfies a term f with probability $f(s)$ and a generalised state μ (an evaluation over the ipo of states) satisfies a term f with the probability

$$\int f \, d\mu.$$

Given a total logic of triples $f \,[\, \mathbf{P} \,]\, g$ where f and g are upper continuous functions and $\mathbf{P}$ is a program, we want f to be a precondition and g and postcondition such that the triple is valid if whenever f is "true" before executing $\mathbf{P}$ then afterwards g is "true". We generalise the notion of "true" to a probabilistic one, thus given a (generalised) state μ and a probability p that f satisfies μ, for the triple to be valid, the state we reach after executing $\mathbf{P}$ say μ' must satisfy g with probability greater than p, or

$$\int f \, d\mu \leq \int g \, d\mu'.$$

For the partial logic of triples $f \,\{\, \mathbf{P} \,\}\, g$ where f and g are lower continuous functions then f satisfies a generalised state μ with probability

$$\mu(S) - \int (1 - f) \, d\mu$$

(this is just $\int f \, d\mu$ except that we can't directly integrate f since is it lower continuous) and we want the probability of g satisfying the state reached after executing $\mathbf{P}$ say μ' plus the probability of $\mathbf{P}$ failing, that is, $\mu(S) - \mu'(S)$, to be greater than the probability that f satisfies μ. Thus we want

$$\mu'(S) - \int (1 - g) \, d\mu' + (\mu(S) - \mu'(S)) \geq \mu(S) - \int (1 - f) \, d\mu$$

which simplifies to

$$\int (1 - g) \, d\mu' \leq \int (1 - f) \, d\mu.$$

Before we define the logics, we give a useful lemma, showing how upper continuous functions can be derived from continuous functions $S \to \mathcal{V}(S)$.

Lemma 7.1 *For any continuous function $G: S \to \mathcal{V}(S)$ and upper continuous function $f: S \to [0,1]$, the function $g: S \to [0,1]$ given by*

$$g(s) = \int f \, dG(s)$$

is upper continuous.

Proof To see this, note that $g = G; f$ in the Kleisli category with $[0,1]$ identified with $\mathcal{V}(\mathbb{1})$, since $f^{\dagger}(\mu) = \int f \, d\mu$. ∎

Note also a trivial corollary, that for G as above, and f lower continuous, then

$$g(s) = 1 - \int (1 - f) \, dG(s)$$

is lower continuous.

7.2 Partial Correctness Logic

We set up a logic for partial correctness, in which the terms of the logic are given by *lower continuous functions* from a set of states S to $[0,1]$. We have mentioned that pointwise products of upper continuous functions are upper continuous. In the logic, we will frequently use a certain combination of upper and lower continuous functions which we will now show is lower continuous. We take b and $\neg b$ (as defined in the previous section) which are upper continuous, and suppose that f and f' lower continuous, then $b(1-f) + \neg b(1-f')$ is upper continuous, hence $1 - (b(1-f) + \neg b(1-f'))$ or $bf + \neg b f' + (1 - (b + \neg b))$ is lower continuous.

We give the logic in terms of axioms for each atomic program, rules for each program combinator and a consequence rule. These operate on triples of the form $f \,\{\, \mathbf{P} \,\}\, g$ where $\mathbf{P}$ is a program and f and g are lower continuous functions.

$$f\,\{\,\texttt{skip}\,\}\,f$$

$$h\,\{\,\text{a:= }\mathbf{E}\,\}\,g \quad (\text{if } 1-h(s) = \int_{x\in\mathit{Vals}} (1-g(s[a/x]))d\mathcal{E}(\mathbf{E})(s))$$

$$\frac{f_1\,\{\,\mathbf{P}\,\}\,g \quad f_2\,\{\,\mathbf{Q}\,\}\,g}{pf_1+(1-p)f_2\,\{\,\mathbf{P}\ \texttt{or}_p\ \mathbf{Q}\,\}\,g}$$

$$\frac{f\,\{\,\mathbf{P}\,\}\,g \quad g\,\{\,\mathbf{Q}\,\}\,h}{f\,\{\,\mathbf{P};\mathbf{Q}\,\}\,h}$$

$$\frac{f\,\{\,\mathbf{P}\,\}\,g \quad f'\,\{\,\mathbf{Q}\,\}\,g}{bf+\neg bf'+(1-(b+\neg b))\,\{\,\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q}\,\}\,g}$$

$$\frac{f\,\{\,\mathbf{P}\,\}\,bf+\neg bk+(1-(b+\neg b))}{bf+\neg bk+(1-(b+\neg b))\,\{\,\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P}\,\}\,k}$$

$$\frac{f\,\{\,\mathbf{P}\,\}\,g}{f'\,\{\,\mathbf{P}\,\}\,g'} \qquad\qquad (\text{if } f'\leq f \text{ and } g\leq g')$$

Note that in the rule for a:= **E** the expression in the side-condition defines h to be lower continuous by the corollary to Lemma 7.1 since $\mathcal{E}(\mathbf{E})(s)$ is continuous in s.

A proof of a triple $\mathcal{F}_n$ is a sequence $\mathcal{F}_1,\ldots,\mathcal{F}_n$ where each $\mathcal{F}_i$ either follows from the previous ones by the application of a rule, or is an instance of an axiom.

7.3 Total Correctness Logic

For the total logic many of the rules are the same. However the triples consist of a program and two upper continuous functions, rather than lower continuous ones as before.

$$f\,[\,\texttt{skip}\,]\,f$$

$$h\,[\,\mathrm{a}{:=}\ \mathbf{E}\,]\,g \qquad\qquad \left(\text{if } h(s) = \int_{x\in\,Vals} g(s[a/x])d\mathcal{E}(\mathbf{E})(s)\right)$$

$$\frac{f_1\,[\,\mathbf{P}\,]\,g \quad f_2\,[\,\mathbf{Q}\,]\,g}{pf_1 + (1-p)f_2\,[\,\mathbf{P}\ \texttt{or}_p\ \mathbf{Q}\,]\,g}$$

$$\frac{f\,[\,\mathbf{P}\,]\,g \quad g\,[\,\mathbf{Q}\,]\,h}{f\,[\,\mathbf{P};\mathbf{Q}\,]\,h}$$

$$\frac{f\,[\,\mathbf{P}\,]\,g \quad f'\,[\,\mathbf{Q}\,]\,g}{bf + \neg bf'\,[\,\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q}\,]\,g}$$

$$\frac{f_{n+1}\,[\,\mathbf{P}\,]\,bf_n + \neg bk \ \ (n \geq 0)}{b(\bigsqcup_n f_n) + \neg bk\,[\,\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P}\,]\,k} \qquad (\text{where } f_0 = \lambda s.0)$$

$$\frac{f\,[\,\mathbf{P}\,]\,g}{f'\,[\,\mathbf{P}\,]\,g'} \qquad (\text{if } f' \leq f \text{ and } g \leq g')$$

Again a proof of a triple is a sequence each term either being an instance of an axiom or derived from previous terms by one of the rules above.

7.4 Semantics and Interpretation

We will interpret the two logics via a denotational semantics. It can be equivalently expressed as super-linear continuous functions $\mathcal{V}(S) \to \mathcal{V}(S)$ or continuous functions $S \to \mathcal{V}(S)$. We choose the second alternative to avoid checking linearity and because it will simplify the dualities.

The denotation of $\mathbf{P}$ will be given as a function $\mathcal{C}[\![\mathbf{P}]\!]: S \to \mathcal{V}(S)$.

$$
\begin{array}{rcl}
\mathcal{C}[\![\texttt{skip}]\!]s & = & \eta(s) \\
\mathcal{C}[\![\mathrm{a}{:=}\ \mathbf{E}]\!]s & = & \lambda O. \displaystyle\int_{x \in \mathit{Vals}} \chi_O \circ s[a/x]\, d\mathcal{E}(\mathbf{E})(s) \\
\mathcal{C}[\![\mathbf{P}\ \texttt{or}_p\ \mathbf{Q}]\!]s & = & p\mathcal{C}[\![\mathbf{P}]\!]s + (1-p)\mathcal{C}[\![\mathbf{Q}]\!](s) \\
\mathcal{C}[\![\mathbf{P};\mathbf{Q}]\!]s & = & \mathcal{C}[\![\mathbf{Q}]\!]^{\dagger} \circ \mathcal{C}[\![\mathbf{P}]\!]s \\
\mathcal{C}[\![\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q}]\!]s & = & b(s)\mathcal{C}[\![\mathbf{P}]\!]s + \neg b(s)\mathcal{C}[\![\mathbf{Q}]\!]s \\
\mathcal{C}[\![\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P}]\!]s & = & (\displaystyle\bigsqcup_n f_n)(s) \\
 & & \text{where } f_0(s) = W \mapsto 0 \\
 & & f_{n+1}(s) = b(s) f_n^{\dagger} \circ \mathcal{C}[\![\mathbf{P}]\!]s + \neg b(s)\eta(s)
\end{array}
$$

The last clause could have been written

$$\mathcal{C}[\![\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P}]\!]s = \mu f: S \to \mathcal{V}S.(b(s)f^{\dagger}(\mathbf{P}(s)) + \neg b(s)\eta(s))$$

where we abbreviate $\mathcal{C}[\![\mathbf{P}]\!](s)$ by $\mathbf{P}(s)$ and μ is the usual fix point operator.

Note the use of the extension of functions $f^{\dagger}$ as defined in Section 4.2. We will find the equation proved in Lemma 4.4, namely

$$\int_{y \in Y} h\, df^{\dagger}(\mu) = \int_{x \in X} h^{\dagger}(f(x))\, d\mu \qquad (7.1)$$

useful in manipulating $f^{\dagger}$. Also note that the definition of $\mathcal{C}[\![\mathrm{a}{:=}\ \mathbf{E}]\!](s)$ is equal to $\mathcal{V}(k)(\mathcal{E}(\mathbf{E})(s))$ where k is the update function $\mathit{Vals} \to S$ given by $x \mapsto s[a/x]$, so by Equation 4.6,

$$\int h \, d\mathcal{V}(g)(\mu) = \int h \circ g \, d\mu$$

from this we see immediately that for any upper continuous function f,

$$\int_{t \in S} f(t) \, d(\text{a:= } \mathbf{E})(s) = \int_{x \in \mathit{Vals}} g(s[a/x]) \, d\mathcal{E}(\mathbf{E})(s) \tag{7.2}$$

since $g \circ k(x) = g(k(x)) = g(s[a/x])$.

7.5 Consistency and Completeness for the Partial Logic

We will now show that the partial logic presented above is consistent and complete. In the consistency proof we will use induction on length of proof to show $f \{ \mathbf{P} \} g$ and for completeness, induction on the structure of commands.

Theorem 7.2 (Consistency) *For all s in S, if there is a proof of $f \{ \boldsymbol{P} \} g$ then*

$$\int (1 - g) \, d\boldsymbol{P}(s) \leq 1 - f(s).$$

Proof Proof is by induction on the length of proof that $f \{ \mathbf{P} \} g$, we show that if the theorem holds for all but the last step in the proof, then it holds for that step too.

Suppose the last step is by the rule $f \{ \texttt{skip} \} f$ then we need to prove

$$\int (1 - f) \, d(\texttt{skip})(s) \leq 1 - f(s)$$

and since $(\texttt{skip})(s) = \eta(s)$ this is trivial.

If the rule used is $f \{ \text{a:= } \mathbf{E} \} g$, then we need to show

$$\int (1 - g) \, d(\text{a:= } \mathbf{E})(s) \leq 1 - f(s)$$

given

$$1 - f(s) = \int_{x \in Vals} (1 - g(s[a/x])) d\mathcal{E}(\mathbf{E})(s)$$

but by Equation 7.2 above,

$$\int (1-g)\, d(\text{a:= } \mathbf{E})(s) = \int (1 - g(t[a/x]))\, d\mathcal{E}(\mathbf{E})(s)$$

which equals $1 - f(s)$ by the side condition on the rule.

Now suppose the last rule used is

$$\frac{f_1 \{ \mathbf{P} \} g \quad f_2 \{ \mathbf{Q} \} g}{pf_1 + (1-p)f_2 \{ \mathbf{P} \ \mathtt{or}_p \ \mathbf{Q} \} g}$$

so by the inductive hypothesis we assume the inequality holds for the antecedents of the rule and we need to prove it for $\mathbf{P}\ \mathtt{or}_p\ \mathbf{Q}$. So we have

$$\int (1-g)\, d\mathbf{P}(s) \leq 1 - f_1(s)$$

and

$$\int (1-g)\, d\mathbf{Q}(s) \leq 1 - f_2(s)$$

and we need to prove

$$\int (1-g)\, d(\mathbf{P}\ \mathtt{or}_p\ \mathbf{Q})(s) \leq 1 - pf_1(s) + (1-p)f_2(s)$$

but since $(\mathbf{P}\ \mathtt{or}_p\ \mathbf{Q})(s) = p\mathbf{P}(s) + (1-p)\mathbf{Q}(s)$ simply substituting for $(\mathbf{P}\ \mathtt{or}_p\ \mathbf{Q})(s)$ and using linearity gives us

$$p \int (1-g)\, d\mathbf{P}(s) + (1-p) \int (1-g)\, d\mathbf{Q}(s) \leq p(1 - f_1(s)) + (1-p)(1 - f_2(s))$$

which we can see holds as it is the result of multiplying the two inequalities from the hypothesis by p and $1-p$ respectively and adding them together.

Now suppose the last rule applied is

$$\frac{f \{ \mathbf{P} \} g \quad g \{ \mathbf{Q} \} h}{f \{ \mathbf{P};\mathbf{Q} \} h}$$

then we need to show

$$\int (1-h)\, d\mathbf{Q}^{\dagger}(\mathbf{P}(s)) \leq 1 - f(s).$$

But Equation 7.1, proved in Lemma 4.4, we know that

$$\int (1-h)\, d\mathbf{Q}^{\dagger}(\mathbf{P}(s)) = \int (1-h)^{\dagger} \circ \mathbf{Q}\, d\mathbf{P}(s)$$

and by the definition of †,

$$(1-h)^{\dagger} \circ \mathbf{Q}(t) = \int (1-h)\, d\mathbf{Q}(t)$$

so we see that

$$\int (1-h)\, d\mathbf{Q}^{\dagger}(\mathbf{P}(s)) = \int_{t \in S} \left(\int (1-h)\, d\mathbf{Q}(t) \right)\, d\mathbf{P}(s) \tag{7.3}$$

but from the inductive hypothesis on the second antecedent,

$$\int (1-h)\, d\mathbf{Q}(t) \leq 1 - g(t)$$

so

$$\int (1-h)\, d\mathbf{Q}(\mathbf{P}(s)) \leq \int (1-g)\, d\mathbf{P}(s)$$

and by the inductive hypothesis on by the first antecedent

$$\int (1-g)\, d\mathbf{P}(s) \leq 1 - f(s)$$

as required.

Now suppose the last rule applied was

$$\frac{f\,\{\,\mathbf{P}\,\}\,g \quad f'\,\{\,\mathbf{Q}\,\}\,g}{bf + \neg bf' + (1-(b+\neg b))\,\{\,\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q}\,\}\,g}$$

we know from the inductive hypothesis that

$$\int (1-g)\, d\mathbf{P}(s) \leq 1 - f(s)$$

and

$$\int (1-g)\, d\mathbf{Q}(s) \leq 1 - f'(s)$$

then

$$\int (1-g)\, d(\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q})(s) = b(s) \int (1-g)\, d(\mathbf{P}(s)) + \neg b(s) \int (1-g)\, d(\mathbf{Q}(s))$$

and substituting with the above equations we get

$$\begin{aligned}\int 1 - g\, d(\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q})(s) &\leq b(s)(1-f(s))) + \neg b(1-f'(s)) \\ &= 1 - \Big(b(s)f(s) + \neg b(s) f'(s) + (1-(b(s)+\neg b(s)))\Big)\end{aligned}$$

as required.

Now suppose the final rule is

$$\frac{f\,\{\,\mathbf{P}\,\}\,bf + \neg bk + (1-(b+\neg b))}{bf + \neg bk + (1-(b+\neg b))\,\{\,\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P}\,\}\,k}$$

so we need to show that

$$\int (1-k)\,(\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P})(s) \leq (1-f(s)).b(s) + (1-k(s)).\neg b(s)$$

we substitute for $(\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P})(s) = \bigsqcup f_n(s)$ where $f_0(s) = W \mapsto 0$ and $f_{n+1}(s) = b(s) f_n^{\dagger} \circ \mathcal{C}[\![\mathbf{P}]\!] + \neg b(s)\eta(s)$ and proceed by induction. If $n = 0$, then the right-hand side is zero, while the left-hand side is clearly non-negative since it is a sum of products of positive terms. Now assuming the result holds for $n = k$ we consider $k + 1$.

$$\int (1-k)\, df_{k+1}(s) = b(s) \int (1-k)\, d(f_n^{\dagger}(P(s)) + \neg b(s)).\eta(s)$$

which, by linearity and Equation 7.1 and the definition of $\dagger$ simplifies to

$$\neg b(s)(1-k(s)) + b(s) \int_{t\in S} \left(\int (1-k)\, df_n(t) \right)\, d\mathbf{P}(s) \tag{7.4}$$

but then by the result for $n = k$ we see that

$$\int (1-k)\, df_n(t) \leq b(t)(1-f(t)) + \neg b(t)(1-k(t))$$

so

$$\int 1 - k\, df_{k+1}(s) \leq \int_{t\in S} b(t)(1-f(t)) + \neg b(t)(1-k(t))\, d\mathbf{P}(s) b(s) + \neg b(s)(1-k(s))$$

but then the main induction hypothesis tells us that

$$\int_{t \in S} b(t)(1 - f(t)) + \neg b(t)(1 - k(t))\, d\mathbf{P}(s) \leq 1 - f(s)$$

so substituting we see that

$$\int (1 - k)\, df_{k+1}(s) \leq b(s)(1 - f(s)) + \neg b(s)(1 - k(s))$$

as required.

Finally suppose that the last rule applied is

$$\frac{f \{ \mathbf{P} \} g}{f' \{ \mathbf{P} \} g'} \quad (\text{if } f' \leq f \text{ and } g \leq g')$$

then we have

$$\int (1 - g)\, d\mathbf{P}(s) \leq 1 - f(s)$$

so

$$\int (1 - g')\, d\mathbf{P}(s) \leq \int (1 - g)\, d\mathbf{P}(s) \leq 1 - f(s) \leq 1 - f'(s)$$

since $g' \geq g$ and $f' \geq f$. ∎

Now we go on to a completeness proof, by induction on the structure of $\mathbf{P}$.

Theorem 7.3 (Completeness) *If for all s in S we have*

$$\int (1 - g)\, d\boldsymbol{P}(s) \leq 1 - f(s)$$

then there is a proof of $f \{ \boldsymbol{P} \} g$.

Proof Proof is by induction on the structure of $\mathbf{P}$.

Suppose $\mathbf{P} = \texttt{skip}$, and the above equation holds for f and g. Then since $\texttt{skip}(s) = \eta(s)$ we know

$$\forall s. 1 - g(s) \leq 1 - f(s)$$

hence $g \geq f$. So we can show $f\,\{\,\texttt{skip}\,\}\,g$ by using the skip rule to deduce $g\,\{\,\texttt{skip}\,\}\,g$ and then the consequence rule since $g \geq f$.

Suppose $\mathbf{P} = (\mathrm{a}{:=}\ \mathbf{E})$. Then if for all s,

$$\int (1-g)\, d(\mathrm{a}{:=}\ \mathbf{E})(s) \leq 1 - f(s)$$

then if we set

$$1 - k(s) = \int (1-g)\, d(\mathrm{a}{:=}\ \mathbf{E})(s)$$

we can deduce $k\,\{\,\mathrm{a}{:=}\ \mathbf{E}\,\}\,g$, by Equation 7.2 and the rule for $\mathrm{a}{:=}\ \mathbf{E}$, and use the consequence rule to get $f\,\{\,\mathrm{a}{:=}\ \mathbf{E}\,\}\,g$.

Now suppose $\mathbf{P} = (\mathbf{Q}\ \texttt{or}_p\ \mathbf{R})$ and for some f and g,

$$\int (1-g)\, d(\mathbf{Q}\ \texttt{or}_p\ \mathbf{R})(s) \leq 1 - f(s)$$

for any s and by the inductive hypothesis the theorem holds for $\mathbf{Q}$ and $\mathbf{R}$. Set

$$1 - f_1(s) = \int (1-g)\, d\mathbf{Q}(s)$$

and

$$1 - f_2(s) = \int (1-g)\, d\mathbf{R}(s)$$

(these are both clearly upper continuous functions by Lemma 7.1. So by the inductive hypothesis on $\mathbf{Q}$ and $\mathbf{R}$ this implies that we can deduce $f_1\,\{\,\mathbf{Q}\,\}\,g$ and $f_2\,\{\,\mathbf{R}\,\}\,g$, so by the rule for $\mathbf{Q}\ \texttt{or}_p\ \mathbf{R}$ we can deduce $pf_1 + (1-p)f_2\,\{\,\mathbf{Q}\ \texttt{or}_p\ \mathbf{R}\,\}\,g$. But the above condition on f and g implies

$$p\left(\int (1-g)\, d\mathbf{Q}(s)\right) + (1-p)\left(\int (1-g)\, d\mathbf{R}(s)\right) \leq 1 - f(s)$$

hence

$$pf_1(s) + (1-p)f_2(s) \geq f(s)$$

so applying the consequence rule with this inequality we deduce that $f\,\{\,\mathbf{Q}\ \texttt{or}_p\ \mathbf{R}\,\}\,g$ as required.

Now suppose $\mathbf{P} = \mathbf{Q}; \mathbf{R}$ and we have f and g satisfying the condition of the theorem. We set

$$1 - h(s) = \int (1-g)\, d\mathbf{R}(s)$$

then by the inductive hypothesis on $\mathbf{R}$, we can deduce $h \{ \mathbf{R} \} g$. Then since we have

$$\int (1-g)\, d(\mathbf{Q}; \mathbf{R})(s) \leq 1 - f(s)$$

we expand and use Equation 7.3 to see that

$$\int 1 - g\, d(\mathbf{R}^{\dagger})(\mathbf{P}(s)) = \int_{t \in S} \left(\int (1-g)\, d\mathbf{R}(t) \right)\, d\mathbf{Q}(s) \leq 1 - f(s)$$

and substituting for g with h, we see

$$\int_{t \in S} (1 - h(t))\, d\mathbf{Q}(s) \leq 1 - f(s)$$

hence by the inductive hypothesis on $\mathbf{Q}$ we can deduce $f \{ \mathbf{Q} \} h$ so finally by the rule for composition we can deduce $f \{ \mathbf{Q}; \mathbf{R} \} g$.

If $\mathbf{P} = \texttt{if } \mathbf{B} \texttt{ then } \mathbf{Q} \texttt{ else } \mathbf{R}$, and suppose for any s,

$$\int (1-g)\, d(\texttt{if } \mathbf{B} \texttt{ then } \mathbf{Q} \texttt{ else } \mathbf{R})(s) \leq 1 - f(s)$$

then we set

$$1 - h_1(s) = \int (1-g)\, d\mathbf{Q}(s)$$

and

$$1 - h_2(s) = \int (1-g)\, d\mathbf{R}(s)$$

then from the inductive hypothesis we can deduce $h_1 \{ \mathbf{Q} \} g$ and $h_2 \{ \mathbf{R} \} g$, hence by the if rule, $bh_1 + \neg bh_2 \{ \texttt{if } \mathbf{B} \texttt{ then } \mathbf{Q} \texttt{ else } \mathbf{R} \} g$. Then substituting in the definition of the semantics of if, we get

$$b(s) \int (1-g)\, d\mathbf{Q}(s) + \neg b(s) \int (1-g)\, d\mathbf{R}(s) \leq 1 - f(s)$$

hence we easily see that $bh_1 + \neg bh_2 \geq f(s)$, so by applying the consequence rule we deduce that $f \{ \texttt{if } \mathbf{B} \texttt{ then } \mathbf{Q} \texttt{ else } \mathbf{R} \} g$.

Finally we take the case of $\mathbf{P} = \texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{Q}$. Suppose we have some g and h such that

$$\int (1-g)\, d(\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{Q})(s) \leq 1 - h(s)$$

for all s. Define

$$1 - k_n(s) = \int (1-g)\, df_n(s)$$

where f_n is from the definition of the semantics of $\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P}$. Then by definition

$$f_{n+1}(s) = \neg b(s)\eta(s) + b(s)\lambda O. \int_{t \in S} f_n(t)(O)\, d\mathbf{Q}(s)$$

and so

$$1 - k_{n+1}(s) = \neg b(s)(1 - g(s)) + b(s) \int_{t \in S} \left(\int (1-g)\, df_n(t) \right)\, d\mathbf{Q}(s)$$

(by Equation 7.1) so

$$1 - k_{n+1}(s) = \neg b(s)(1 - g(s)) + b(s) \int (1 - k_n)\, d\mathbf{Q}(s).$$

So if $1 - k = \bigsqcup (1 - k_n)$ (since $1 - k_n$ is clearly increasing), then

$$1 - k = \neg b(1 - g) + b \int (1 - k) d\mathbf{Q}$$

so if we set $1 - f = \int (1-k)\, d\mathbf{Q}$, then firstly, by the inductive hypothesis on $\mathbf{Q}$ we can deduce $f\,\{\,\mathbf{Q}\,\}\,k$, or $f\,\{\,\mathbf{Q}\,\}\,bf + \neg bg + (1 - (b + \neg b))$ substituting for the equation above. Then, by the while rule we can deduce $k\,\{\,\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{Q}\,\}\,g$. But by the construction of k_n, we see that $1 - k(s) = \int (1-g)\, d(\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{Q})(s)$, so $1 - k \leq 1 - h$, i.e. $h \leq k$ so by the consequence rule we can deduce $h\,\{\,\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{Q}\,\}\,g$ as required. ∎

In this logic we can define the weakest liberal precondition $wlp(\mathbf{P}, g)$ to be the weakest pre-condition for $f\,\{\,\mathbf{P}\,\}\,g$, i.e. the weakest f such that $f\,\{\,\mathbf{P}\,\}\,g$ holds. Formally this is the sup of all functions h such that $h\,\{\,\mathbf{P}\,\}\,g$. It is clear that the formula

$$1 - f(s) = \int (1-g)\, d\mathbf{P}(s)$$

gives $wlp(\mathbf{P}, g)$ from the completeness and consistency proofs (we will prove this fact in Section 7.8). Then the following equations hold.

$$
\begin{aligned}
wlp(\texttt{skip}, g) &= g \\
wlp(\mathrm{a}{:=}\ \mathbf{E}, g) &= 1 - \lambda s. \int_{v \in Vals} (1 - g(s[a/v]))\, d\mathcal{E}(\mathbf{E})(s) \\
wlp(\mathbf{P}\ \texttt{or}_p\ \mathbf{Q}, g) &= p\, wlp(\mathbf{P}, g) + (1-p) wlp(\mathbf{Q}, g) \\
wlp(\mathbf{P}; \mathbf{Q}, g) &= wlp(\mathbf{Q}, wlp(\mathbf{P}, g)) \\
wlp(\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q}, g) &= b\, wlp(\mathbf{P}, g) + \neg b\, wlp(\mathbf{Q}, g) + (1 - (b + \neg b)) \\
wlp((\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P}), g) &= \inf f_n \\
& \quad f_0 = \lambda s.1 \\
& \quad f_{n+1} = b\, wlp(\mathbf{P}, f_n) + \neg b\, g + (1 - (b + \neg b))
\end{aligned}
$$

So a language which allows us to construct terms from some atomic ones and which contains the constructions $pf + (1-p)g$, $bf + \neg bg + (1 - (b + \neg b))$, the substitution equation and taking decreasing limits, will allow us to express the weakest precondition of any term continuous function g and program $\mathbf{P}$.

In comparing this logic to the usual Hoare logic for a simple while language we find some interesting differences. A proposition which defines a closed set of states in Hoare logic (a "property") is represented by the characteristic function of that set. Analogues of the usual logical connectives and symbols appear and on propositions correspond to the usual connectives, e.g. $\lambda s.1$ represents true and $\lambda s.0$, false, $f \& g$ is given by $\min(f, g)$ (taking the pointwise minimum) and $f \vee g$ given by $\max(f, g)$. Omitting the probabilistic elements gives exactly the usual Hoare logic, with all functions being the characteristic function of some set; b and $\neg b$ also become characteristic functions, the parts of the rules involving $1-(b+\neg b)$ disappear, and the rules for if and while can be transformed into their usual form.

The rule

$$
\frac{f\,\{\,\mathbf{P}\,\}\,g \quad h\,\{\,\mathbf{P}\,\}\,g}{f \vee h\,\{\,\mathbf{P}\,\}\,g}
$$

from Hoare logic is derivable from the logic, since by the completeness and consistency we need only show that

$$(1 - f(s)) \geq \int (1 - g)\, d\mathbf{P}(s)$$

and

$$(1 - h(s)) \geq \int (1 - g)\, d\mathbf{P}(s)$$

implies

$$(1 - \max(f(s), h(s))) \geq \int (1 - g)\, d\mathbf{P}(s)$$

which is easy since $1 - \max(a, b) = \min(1 - a, 1 - b)$.

However the similar rule

$$\frac{f \,\{\, \mathbf{P} \,\}\, g \quad f \,\{\, \mathbf{P} \,\}\, h}{f \,\{\, \mathbf{P} \,\}\, g \& h}$$

which is true for Hoare logic in the deterministic language, does not hold. If we consider the simple example of $\mathbf{P} =$ x:= a $\texttt{or}_{1/2}$ x:= b where $S = \{a, b\}$ (the flat ipo), then $(\lambda s.1/2) \,\{\, \mathrm{P} \,\}\, \chi_a$ and $(\lambda s.1/2) \,\{\, \mathrm{P} \,\}\, \chi_b$ but $\chi_a \& \chi_b = \lambda s.0$ and it is clearly not true that $(\lambda s.1/2) \,\{\, \mathrm{P} \,\}\, (\lambda s.0)$ as $\int (1 - (\lambda s.0))\, d\mathbf{P}(s) = 1$ for any s (here note that $\mathbf{P}(s) = 1/2\, \eta(a) + 1/2\, \eta(b)$) and $1 - (\lambda s.1/2) = 1/2$.

We can also note that it is impossible in this logic to find a strongest postcondition formula, that is some $g(\mathbf{P}, f)$ such that $f \,\{\, \mathbf{P} \,\}\, g$ and g is the least such function. Take $\mathbf{P}$ as above so $\mathbf{P}(s) = 1/2\, \eta(a) + 1/2\, \eta(b)$ for $S = \{a, b\}$ (the flat ipo), then if $f = \lambda s.1/2$, then $g = \chi_a$ and $g = \chi_b$ are both functions which satisfy $f \,\{\, \mathbf{P} \,\}\, g$ whose only joint lower bound is $g = \lambda s.0$ which does not satisfy $f \,\{\, \mathbf{P} \,\}\, g$.

7.6 Consistency and Completeness for the Total Logic

The proofs of consistency and completeness for the total logic are very similar to those of the partial logic.

Recall that the interpretation of the total logic is

$$f\,[\,\mathbf{P}\,]\,g \iff \forall s \int g\, d\mathbf{P}(s) \geq f(s).$$

We will now show that the total logic is consistent.

Theorem 7.4 (Consistency) *If there exists a proof of $f\,[\,\boldsymbol{P}\,]\,g$, then for all s,*

$$\int g\, d\boldsymbol{P}(s) \geq f(s).$$

Proof Proof is by induction on the length of proof that $f\,[\,\mathbf{P}\,]\,g$.

If the last rule used is $f\,[\,\texttt{skip}\,]\,f$ then we need to prove

$$\int f\, d(\texttt{skip})(s) \geq f(s)$$

and since $(\texttt{skip})(s) = \eta(s)$ this is trivial.

If the rule used is $h\,[\,\mathrm{a}{:=}\ \mathbf{E}\,]\,g$, where

$$h(s) = \int_{x \in \mathit{Vals}} g(s[a/x])\, d\mathcal{E}(\mathbf{E})(s)$$

then we need to show

$$\int g\, d(\mathrm{a}{:=}\ \mathbf{E})(s) \geq h(s)$$

but by Equation 7.2,

$$\int g\, d(\mathrm{a}{:=}\ \mathbf{E})(s) = \int_{x \in \mathit{Vals}} g(s[a/x])\, d\mathcal{E}(\mathbf{E})(s) = h(s).$$

Now suppose the last rule applied is

$$\frac{f_1 \,[\, \mathbf{P} \,]\, g \quad f_2 \,[\, \mathbf{Q} \,]\, g}{pf_1 + (1-p)f_2 \,[\, \mathbf{P} \;\texttt{or}_p\; \mathbf{Q} \,]\, g}$$

so we assume the equation holds for the antecedents of the rule and need to prove it for $\mathbf{P} \;\texttt{or}_p\; \mathbf{Q}$. So we have

$$\int g \, d\mathbf{P}(s) \geq f_1(s)$$

and

$$\int g \, d\mathbf{Q}(s) \geq f_2(s)$$

and we need to prove

$$\int g \, d(\mathbf{P} \;\texttt{or}_p\; \mathbf{Q})(s) \geq pf_1(s) + (1-p)f_2(s)$$

but since $(\mathbf{P} \;\texttt{or}_p\; \mathbf{Q})(s) = p\mathbf{P}(s) + (1-p)\mathbf{Q}(s)$ simply substituting for $(\mathbf{P} \;\texttt{or}_p\; \mathbf{Q})(s)$ and using linearity gives us

$$p \int g \, d\mathbf{P}(s) + (1-p) \int g \, d\mathbf{Q}(s) \geq pf_1(s) + (1-p)f_2(s)$$

as required.

Now suppose the last rule applied is

$$\frac{f \,[\, \mathbf{P} \,]\, g \quad g \,[\, \mathbf{Q} \,]\, h}{f \,[\, \mathbf{P};\mathbf{Q} \,]\, h}$$

then we need to show

$$\int h \, d(\mathbf{P};\mathbf{Q})(s) \geq f(s)$$

but by the induction hypothesis on $\mathbf{P}$ we know

$$\int g \, d\mathbf{P}(s) \geq f(s)$$

and by expanding the semantics and by Equation 7.3,

$$\int h \, d(\mathbf{P};\mathbf{Q})(s) = \int_{t \in S} \left(\int h \, d\mathbf{Q}(t) \right) \, d\mathbf{P}(s)$$

and by the hypothesis for $\mathbf{Q}$,

$$\int h \, d\mathbf{Q}(t) \geq g(t)$$

so

$$\int h \, d(\mathbf{P};\mathbf{Q})(s) \geq \int_{t \in S} g(t) \, d\mathbf{P}(s) \geq f(s)$$

as required.

Now suppose the last rule applied was

$$\frac{f\,[\,\mathbf{P}\,]\,g \quad f'\,[\,\mathbf{Q}\,]\,g}{bf + \neg bf'\,[\,\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q}\,]\,g}$$

then by the induction hypothesis

$$\int g \, d\mathbf{P}(s) \geq f(s)$$

and

$$\int g \, d\mathbf{Q}(s) \geq f'(s)$$

then

$$\int g \, d(\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q})(s) = \int g \, d(b(s)\mathbf{P}(s) + \neg b(s)\mathbf{Q}(s))$$

and substituting with the above equations we get

$$\int g \, d(\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q})(s) \geq b(s)f(s) + \neg b(s)f'(s)$$

as required.

Now suppose the final rule is

$$\frac{h_{n+1}\,[\,\mathbf{P}\,]\,\neg bg + bh_n \quad h_0 = \lambda s.0}{\neg bg + b\bigsqcup h_n\,[\,\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P}\,]\,g}$$

so we need to show that

$$\int g \, d(\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P})(s) \geq \neg b(s)g(s) + b(s)\bigsqcup h_n(s)$$

we substitute for $(\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P})(s) = \bigsqcup f_n(s)$ where $f_0(s) = W \mapsto 0$ and $f_{n+1}(s) = b(s)f_n^\dagger \circ \mathcal{C}[\![\mathbf{P}]\!] + \neg b(s)\eta(s)$ and show that for any n,

$$\int g\, df_{n+1}(s) \geq \neg b(s)g(s) + b(s)h_n(s).$$

If $n = 0$, then as $f_1(s) = \neg b(s)\eta(s)$ and $h_0(s) = 0$ so the two sides are equal. Suppose the inequality holds for $n = k$. We consider it for $n = k+1$, i.e.

$$\int g\, df_{k+2}(s) \geq \neg b(s)g(s) + b(s)h_{k+1}(s).$$

Now we know (by Equation 7.4),

$$\int g\, df_{k+2}(s) = \neg b(s)g(s) + b(s)\int_{t\in S}\left(\int g\, df_{k+1}(t)\right)\, d\mathbf{P}(s)$$

and by substituting for $\int g\, df_{k+1}(t)$ using the inequality for $n = k$ we see that

$$\int g\, df_{n+1}(s) \geq \neg b(s)g(s) + b(s)\int \neg bg + bh_n\, d\mathbf{P}(s)$$

but by the main inductive hypothesis,

$$\int \neg bg + bh_k\, d\mathbf{P}(s) \geq h_{k+1}(s)$$

so we see that

$$\int g\, f_{k+2}(s) \geq \neg b(s)g(s) + b(s)h_{k+1}(s)$$

as required.

Finally suppose that the last rule applied is

$$\frac{f\,[\,\mathbf{P}\,]\,g}{f'\,[\,\mathbf{P}\,]\,g'} \quad (\text{if } f' \leq f \text{ and } g \leq g')$$

then we have

$$\int g\, d\mathbf{P}(s) \geq f(s)$$

so

$$\int g'\, d\mathbf{P}(s) \geq \int g\, d\mathbf{P}(s) \geq f(s) \geq f'(s)$$

since $g' \geq g$ and $f' \geq f$. ∎

Now we go on to a completeness proof for this logic.

Theorem 7.5 (Completeness) *If for any s we have*

$$\int g \, d\boldsymbol{P}(s) \geq f(s)$$

then we can prove $f\,[\,\boldsymbol{P}\,]\,g$.

Proof Proof is by induction on the structure of $\mathbf{P}$.

Suppose $\mathbf{P} = \texttt{skip}$. If we have f and g such that

$$\int g \, d(\texttt{skip})(s) \geq f(s)$$

for any s, then since $(\texttt{skip})(s) = \eta(s)$ this means $g(s) \geq f(s)$. So we can show $f\,[\,\texttt{skip}\,]\,g$ by using the skip rule to deduce $g\,[\,\texttt{skip}\,]\,g$ and then the consequence rule since $g \geq f$.

Suppose $\mathbf{P} = (\text{a:=}\ \mathbf{E})$ and we have f and g such that

$$\int g \, d(\text{a:=}\ \mathbf{E})(s) \geq f(s).$$

Then we can deduce $h\,\{\,\text{a:=}\ \mathbf{E}\,\}\,g$ where

$$h(s) = \int_{x \in \mathit{Vals}} g(s[a/x]) d\mathcal{E}(\mathbf{E})(s)$$

and by Equation 7.2

$$h(s) = \int g \, d(\text{a:=}\ \mathbf{E})(s).$$

So $h(s) \geq f(s)$ and we use the consequence rule to get $f\,[\,\text{a:=}\ \mathbf{E}\,]\,g$.

Now suppose $\mathbf{P} = (\mathbf{Q}\ \texttt{or}_p\ \mathbf{R})$ and we have some f, g such that

$$\int g \, d(\mathbf{Q}\ \texttt{or}_p\ \mathbf{R})(s) \geq f(s)$$

for any s. Then by the inductive hypothesis the theorem holds for $\mathbf{Q}$ and $\mathbf{R}$. So we set

$$f_1(s) = \int g \, d\mathbf{Q}(s)$$

and

$$f_2(s) = \int g \, d\mathbf{R}(s)$$

and then we can deduce $f_1 \, [\, \mathbf{Q} \,] \, g$ and $f_2 \, [\, \mathbf{R} \,] \, g$ and hence by the or rule, $pf_1 + (1-p)f_2 \, [\, \mathbf{Q} \; \texttt{or}_p \; \mathbf{R} \,] \, g$. But

$$\int g \, d(\mathbf{Q} \; \texttt{or}_p \; \mathbf{R})(s) = p(\int g \, d\mathbf{Q}(s)) + (1-p)(\int g \, d\mathbf{R}(s))$$

i.e.

$$\int g \, d(\mathbf{Q} \; \texttt{or}_p \; \mathbf{R})(s) = pf_1(s) + (1-p)f_2(s)$$

so $pf_1 + (1-p)f_2 \geq f$ and by the consequence rule we can deduce $f \, [\, \mathbf{Q} \; \texttt{or}_p \; \mathbf{R} \,] \, g$ as required.

Now suppose $\mathbf{P} = \mathbf{Q}; \mathbf{R}$ and we have f, g as above. We set

$$h(s) = \int g \, d\mathbf{R}(s)$$

then by the inductive hypothesis for $\mathbf{R}$, we can deduce $h \, [\, \mathbf{R} \,] \, g$. But from Equation 7.3 we know that

$$\int g \, d(\mathbf{Q}; \mathbf{R})(s) = \int_{t \in S} \left(\int g \, d\mathbf{R}(t) \right) \, d\mathbf{Q}(s)$$

and substituting in h we see that

$$\int g \, d(\mathbf{Q}; \mathbf{R})(s) = \int h \, d\mathbf{Q}(s) \geq f(s)$$

and by the inductive hypothesis on $\mathbf{Q}$ we can deduce $f \, [\, \mathbf{Q} \,] \, h$ so finally by the rule for composition we can deduce $f \, [\, \mathbf{Q};\mathbf{R} \,] \, g$.

If $\mathbf{P} = \texttt{if} \; \mathbf{B} \; \texttt{then} \; \mathbf{Q} \; \texttt{else} \; \mathbf{R}$, suppose we set

$$h_1(s) = \int g \, d\mathbf{Q}(s)$$

and

$$h_2(s) = \int g \, d\mathbf{R}(s)$$

then as before we can deduce $h_1\,[\,\mathbf{Q}\,]\,g$ and $h_2\,[\,\mathbf{R}\,]\,g$, hence we can deduce $bh_1 + \neg bh_2\,[\,\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{Q}\ \texttt{else}\ \mathbf{R}\,]\,g$. But then

$$\int g\, d(\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{Q}\ \texttt{else}\ \mathbf{R})(s) = b(s)\int g\ d\mathbf{Q}(s) + \neg b(s)\int g\ d\mathbf{R}(s)$$

so

$$b(s)f_1(s) + \neg b(s)f_2(s) \geq f(s)$$

and by applying the consequence rule we deduce that $f\,[\,\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{Q}\ \texttt{else}\ \mathbf{R}\,]\,g$ as required.

Finally we consider the case of $\mathbf{P} = (\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{Q})$. Suppose we have f and g such that

$$\int g\, d(\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{Q})(s) \geq f(s).$$

Define

$$k_n(s) = \int g\, df_n(s)$$

where f_n is as defined in the semantics for the while loop. Then we know from Equation 7.4 that

$$k_{n+1}(s) = \neg b(s)g(s) + b(s)\int k_n\ d\mathbf{Q}(s)$$

and setting $h_n = \int k_n\ d\mathbf{Q}(s)$, we know $k_0(s) = 0$ so $h_0(s) = 0$ and so by the inductive hypothesis we can deduce $h_n\,[\,\mathbf{Q}\,]\,k_n$ or $h_{n+1}\,[\,\mathbf{Q}\,]\,bh_n + \neg bg$. So by the while rule, we can deduce $b\bigsqcup h_n + \neg bg\,[\,\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{Q}\,]\,g$. But then

$$b\bigsqcup h_n + \neg bg = \bigsqcup k_n$$

and

$$\bigsqcup k_n(s) = \int g\, d(\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{Q})(s)$$

so $b\bigsqcup h_n + \neg bg \geq f$ so by the consequence rule we can deduce $f\,[\,\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{Q}\,]\,g$ as required. ∎

Similarly we can define the weakest precondition $wp(\mathbf{P}, g)$ to be the weakest f such that $f\,[\,\mathbf{P}\,]\,g$ holds. Note that weakest in the total logic means the largest function. We define $wp(\mathbf{P}, g)$ to be the sup of all functions h such that $h\,\{\,\mathbf{P}\,\}\,g$. As before we can see from the consistency and completeness that

$$wp(\mathbf{P}, g)(s) = \int g \, d\mathbf{P}(s)$$

(and will prove it in Section 7.8). Then $wp(\mathbf{P}, g)$ satisfies the equations,

$$\begin{array}{rcl}
wp(\texttt{skip}, g) & = & g \\
wp(\mathrm{a}{:=}\ \mathbf{E}, g) & = & \lambda s. \displaystyle\int_{v \in Vals} g(s[a/v]) \, d\mathcal{E}(\mathbf{E})(s) \\
wp(\mathbf{P}\ \texttt{or}_p\ \mathbf{Q}, g) & = & p\, wp(\mathbf{P}, g) + (1-p) wp(\mathbf{Q}, g) \\
wp(\mathbf{P}; \mathbf{Q}, g) & = & wp(\mathbf{Q}, wp(\mathbf{P}, g)) \\
wp(\texttt{if}\ \mathbf{B}\ \texttt{then}\ \mathbf{P}\ \texttt{else}\ \mathbf{Q}, g) & = & b\, wp(\mathbf{P}, g) + \neg b\, wp(\mathbf{Q}, g) \\
wp((\texttt{while}\ \mathbf{B}\ \texttt{do}\ \mathbf{P}), g) & = & \bigsqcup f_n,\ f_0 = \lambda s.0 \\
 & & f_{n+1} = b\, wp(\mathbf{P}, f_n) + \neg b\, g
\end{array}$$

So for the total logics the operations needed to construct all the terms needed in the logic are $pf + (1-p)g$, $bf + \neg bg$, the substitution equation and increasing limits.

Exactly as with the partial logic, we can compare this to the usual (Dijkstra) logic of total correctness. Propositions are denoted by characteristic functions of sets; and, or, true and false are as before. The rule

$$\frac{f\,[\,\mathbf{P}\,]\,g \quad h\,[\,\mathbf{P}\,]\,g}{f \vee h\,[\,\mathbf{P}\,]\,g}$$

holds, but not the rule

$$\frac{f\,[\,\mathbf{P}\,]\,g \quad f\,[\,\mathbf{P}\,]\,h}{f\,[\,\mathbf{P}\,]\,g \& h}$$

and there is no strongest postcondition (all for essentially the same reasons as before).

7.7 An Example

In this section we give an example to show how the logics work. We take the program $\mathbf{P}$ given by n := 1 ; `while` $\mathbf{B}$ `do` n := n+1 where the conditional $\mathbf{B}$ has functions $b(s) = 1/2$ and $\neg b(s) = 1/2$, that is, its denotation gives true with probability 1/2 and false with probability 1/2. We expect this program to terminate with probability 1 and have probability 2^{-N} of ending on $n = N$. The set of states is simply the natural numbers. We first make the general analysis of the meaning of the assignment statements.

For n := 1, we note that $\mathcal{E}(\text{n:=1})(n) = \eta(1)$ and $s[n/1] = 1$, so we have $f\,\{\,\text{n := 1}\,\}\,g$ where for all n, $f(n) = g(1)$ and similarly $f\,[\,\text{n := 1}\,]\,g$. For n := n+1, $\mathcal{E}(\text{n:=n+1})(n) = \eta(n+1)$ so we have $f\,\{\,\text{n := n+1}\,\}\,g$ when for all n, $f(n) = g(n+1)$ and similarly for $f\,[\,\text{n := n+1}\,]\,g$.

To show that the program terminates with probability 1, we consider the value of f such that $f\,[\,\mathbf{P}\,]\,\lambda n.1$, that is the weakest precondition of "true". If $f_0(n) = 0$ and $f_{m+1}(n) = 1/2\,f_m(n) + 1/2$ then clearly $f_{m+1}\,[\,\text{n:=n+1}\,]\,f_m.b + (\lambda n.1).\neg b$ so by the while rule, $(\bigsqcup f_m).b + (\lambda n.1).\neg b\,[\,\texttt{while}\ \mathbf{B}\ \texttt{do}\ \text{n := n+1}\,]\,\lambda n.1$ but then clearly $f_m = \lambda n.1 - 2^{-m}$, so $\bigsqcup f_m = \lambda n.1$ hence we actually have

$$\lambda n.1\,[\,\texttt{while}\ \mathbf{B}\ \texttt{do}\ \text{n := n+1}\,]\,\lambda n.1.$$

Finally we get $f\,[\,\mathbf{P}\,]\,\lambda 1.n$ where $f(n) = (\lambda n.1)(1) = 1$. So we see that the program terminates with probability 1.

Now we consider the precondition of the functions χ_N, that is 1 on $n = N$ and 0 elsewhere. We will use the partial logic, the interesting part is seeing what the loop invariant is. Consider the function $f_N(n)$ given by

$$f_N(n) = \begin{cases} 2^{n-N} & \text{if } m < N \\ 0 & \text{otherwise} \end{cases}.$$

To see that f_N is the loop invariant with $k = \chi_N$ we need to check that for any N, $f_N \{ \text{n:=n+1} \} \, bf_N + \neg b\chi_N + (1 - (b + \neg b))$. We can drop the term $(1 - (b + \neg b))$ as in this case it is always zero. We need to see that $f_N(n) = 1/2\, f_N(n + 1) + 1/2\, \chi_N(n + 1)$. But if $n + 1 < N$, then $f_N(n) = 2^{n-N} = 1/2\, 2^{n+1-N} = 1/2\, f_N(n + 1)$ and if $n + 1 = N$ then $f_N(n) = 2^{-1} = 1/2$ and for $n + 1 > N$, both sides are zero. So f_N is the loop invariant and from the while rule, $bf_N + \neg b\chi_N\, [\, \texttt{while}\ \mathbf{B}\ \texttt{do}\ \text{n := n+1} \,]\, \chi_N$. Then $(bf_N + \neg b\chi_N)(1) = 2^{-N}$ so we derive $\lambda x.2^{-N} \{ \mathbf{P} \} \chi_N$ as expected.

7.8 Duality and Logic

We now consider how we can use the duality given in Chapter 6 to relate the logic and semantics we have given above. The semantics of a program $\mathbf{P}$ is defined as a continuous function $S \to \mathcal{V}S$, the weakest precondition semantics is a function $F(S) \to F(S)$ where we recall $F(S)$ is the set of upper continuous functions $f: S \to [0, 1]$ and the weakest liberal precondition semantics is a function $L(S) \to L(S)$ where by $L(S)$ we denote the set of lower continuous functions $f: S \to [0, 1]$. Recall from Chapter 6 the duality between functions $F: S \to \mathcal{V}(S)$ and functions $G: F(S) \to F(S)$ (linear and continuous functionals). It is given by

$$F^{\circ}(k: S \to [0, 1]) = \lambda x. \int k \, dF(x)$$

and

$$G^{\star}(s: S) = O \mapsto G(\chi_O)(s).$$

Furthermore we can form a similar duality between the semantics and functionals on the space of lower continuous functions $S \to [0, 1]$ using the fact that if f is upper continuous, $1 - f$ is lower continuous and vice versa. This duality is given by

$$F^{\bullet}(k) = \lambda x.1 - \int (1 - k) \, dF(x)$$

and

$$G^*(s{:}\,S) = \lambda O.1 - G(1 - \chi_O)(s).$$

Here the functionals $G{:}\,L(S) \to L(S)$ are not the continuous linear ones, but they are continuous with respect to decreasing sequences and satisfy $G(\sum_{i=1}^{n} r_i k_i) = \sum_{i=1}^{n} r_i G(k_i) + (1 - \sum_{i=1}^{n})(1 - G(\lambda x.0))$.

Having duality between the total logic and semantics means that for any program $\mathbf{P}$ the semantics, $\mathcal{C}[\![\mathbf{P}]\!]$, is dual to the functional given by the weakest precondition semantics, namely $\lambda g.wp(\mathbf{P}, g)$. Symbolically this is just

$$\mathcal{C}[\![\mathbf{P}]\!]^\circ = \lambda g.wp(\mathbf{P}, g).$$

Then from the duality we know that $(\lambda g.wp(\mathbf{P}, g))^\star = \mathcal{C}[\![\mathbf{P}]\!]$. We can represent this duality as a diagram

$$\begin{tikzcd}
 & \mathit{Sem} \arrow[dd, shift left, "\circ"] \\
\mathit{Com} \arrow[ur, "\mathcal{C}[\![\cdot]\!]"] \arrow[dr, "wp"'] & \\
 & \mathit{Log} \arrow[uu, shift left, "\star"]
\end{tikzcd}$$

where Com is the set of programs, Sem is the set of continuous functions $S \to \mathcal{V}S$ and Log is the set of continuous linear functionals $F(S) \to F(S)$. Similarly, saying the partial logic is dual to the semantics can be expressed as

$$\mathcal{C}[\![\mathbf{P}]\!]^\bullet = \lambda g.wlp(\mathbf{P}, g)$$

which also shows $(\lambda g.wlp(\mathbf{P}, g))^* = \mathcal{C}[\![\mathbf{P}]\!]$.

We can easily see that both logics are dual to the semantics from the completeness and consistency proofs.

Theorem 7.6 (Total Duality) *For any $\boldsymbol{P}$, the semantics considered as a function $\boldsymbol{P}: S \to \mathcal{V}S$ is the dual of the weakest precondition for $\boldsymbol{P}$ considered as a function $wp(\boldsymbol{P}): F(S) \to F(S)$, where the duality is given by $\star$ and $\circ$ as defined above.*

Proof We need to show that $wp(\mathbf{P}) = \mathcal{C}[\![\mathbf{P}]\!]^\star$. So we need to show that $wp(\mathbf{P}, g) = \lambda s. \int g \, d\mathbf{P}(s)$ for any g. But from completeness we know that if $f = \lambda s. \int g \, d\mathbf{P}(s)$ then $f \, [\, \mathbf{P} \,] \, g$ so $f \leq wp(\mathbf{P}, g)$ and from consistency we know that if $h \, [\, \mathbf{P} \,] \, f$ then $h \leq f$ so f is clearly the sup of all h such that $h \, [\, \mathbf{P} \,] \, g$ i.e. $wp(\mathbf{P}, g)$. ∎

Theorem 7.7 (Partial Duality) *For any $\boldsymbol{P}$, the semantics considered as a function $\boldsymbol{P}: S \to \mathcal{V}$ is the dual of the weakest liberal precondition for $\boldsymbol{P}$ considered as a function $wlp(\boldsymbol{P}): L(S) \to L(S)$, where the duality is given by $*$ as $\bullet$ as defined above.*

Proof We need to show that $wlp(\mathbf{P}) = \mathcal{C}[\![\mathbf{P}]\!]^*$ i.e. that for any g, $wlp(\mathbf{P}, g) = \lambda s. 1 - \int 1 - g \, d\mathbf{P}(s)$, or $1 - wlp(\mathbf{P}, g) = \lambda s. \int 1 - g \, d\mathbf{P}(s)$. We set $f = \lambda s. 1 - \int 1 - g \, d\mathbf{P}(s)$, then by completeness we know that $f \, \{\, \mathbf{P} \,\} \, g$ since $1 - f \geq \int 1 - g \, d\mathbf{P}(s)$ and by consistency, if $h \, \{\, \mathbf{P} \,\} \, g$, then $1 - h \geq \int 1 - g \, d\mathbf{P}(s) = 1 - f$ so $h \leq f$ so f is the required sup and $wlp(\mathbf{P}, g) = f$. ∎

We can also use the duality to derive the logics from the semantics or vice versa. For instance given the semantics for or_p as $\text{or}_p(f, g) = pf + (1-p)g$ as a function $Sem \times Sem \to Sem$ we can derive the weakest pre-condition semantics for or_p as the function $(F, G) \mapsto \text{or}_p(F^\circ, G^\circ)^*$. Expanding this gives us the function $(F, G) \mapsto pF + (1-p)G$ as expected. The duality replaces the interpretation of the logic, since we say a rule is valid exactly when its dual is a correct rule in the semantics.

However it is still not clear exactly how the wp and wlp semantics are related to the logics.

7.9 Concluding Remarks

In this chapter we have given Hoare-style total and partial logics for a simple while language extended with a probabilistic choice operation. Then we have shown how these logics may be related to a semantics via the duality in Chapter 6.

If we consider only that fragment of the logics without or and with only non-deterministic assignment and test, and with only characteristic functions as terms then they are equivalent to the usual program logics, interpreting first order logic terms as the characteristic functions of the set of states satisfying the terms. This is because when all the functions are characteristic functions, pointwise multiplication becomes equivalent to logical and.

In our logics, the relation between the conditions for the while loop in the partial and total case is interesting. In both cases, the weakest liberal pre-condition is given by the limit of a constructed sequence, which naturally satisfies a recursive equation (the expansion of a while loop). In the partial case, the constructed function is the least fixed point of the recursive equation, and because of the consequence rule, any fixed point of the equation would satisfy the condition, thus the rule just requires the precondition to be a solution. In the total case, the solution is the greatest fixed point so it must be constructed explicitly as a limit.

We might possibly get a third duality by considering monotone functions which could give yet another program logic although it is not clear what sort of logic this might be.

Chapter 8

Metalanguage

In this chapter we develop a semantics for a typed, functional metalanguage similar to that in of Plotkin [31], but with an extra operation, probabilistic choice.

Since the language is probabilistic, we will expect the denotation of an expression to be a evaluation, that is it will lie in $\mathcal{V}(X)$ where X is the ipo associated with the type of the expression. Similarly the type of a functional expression must be $X \to \mathcal{V}(Y)$ (or equivalently $\mathcal{V}X \to \mathcal{V}Y$). So we use a different function space type constructor, denoted by $X \to_{\mathcal{V}} Y$ which is the function space of total functions $X \to \mathcal{V}(Y)$.

We will show how to solve recursive domain equations with this new type constructor. We will then give the metalanguage a denotational and operational semantics and prove that they are equivalent. This metalanguage will be similar to the one given by Moggi in [26] although with more features.

8.1 Solving Domain Equations

As discussed above, the natural way to denote the type of the probabilistic function space is $X \to \mathcal{V}(Y)$ rather than $X \rightharpoonup Y$. We need to extend the theory in [39] to cover this new function space operator so we can solve recursive equations thus giving us ipos to represent recursive types.

The approach of [39] uses the category of embeddings (a category of ipos with projection-pairs as morphisms), and expresses solutions to domain equations as colimits of sequences of ipos and embeddings. The sequences are generated by a functor obtained from the type constructors in the recursive type equation; the only condition on the functor for the limit to exist is that it is continuous in a special sense (i.e. it preserves colimits and the way they are colimits) described in [39].

Since we will be working in the category of embeddings, $\mathbf{Ipo}^E$, we will first show that $\mathcal{V}$ is a functor on this category, not just the category $\mathbf{Ipo}$. Then we will show that the new function space functor is continuous.

Recall from Section 4.2 the definition of the functor $\mathcal{V}$. We only defined the functor on continuous maps $f: P \to Q$, but in fact we can define it on partial functions in the same way; we set $\mathcal{V}(f)(\mu) = O \mapsto \mu(f^{-1}(O))$ and this defines an evaluation for any μ and is a morphism exactly as before. This extension makes $\mathcal{V}$ a functor on the category of ipos with partial functions and morphisms, the proof is as in Theorem 4.3. To see that $\mathcal{V}$ is a functor on the category $\mathbf{Ipo}^E$, which has projection-pairs as morphisms we need to show that $\mathcal{V}$ acts appropriately on projection pairs. Given a pair of functions

$$P \overset{f}{\rightharpoonup} Q \overset{g}{\rightharpoonup} P$$

these transform to

$$\mathcal{V}P \overset{\mathcal{V}f}{\to} \mathcal{V}Q \overset{\mathcal{V}g}{\to} \mathcal{V}P.$$

Then if f and g are a projection pair, we know that $g \circ f = \mathrm{id}_P$ and $f \circ g \sqsubseteq \mathrm{id}_Q$ and we need to prove the corresponding results for $\mathcal{V}f$ and $\mathcal{V}g$. But since $\mathcal{V}$ is a functor we know that

$$\mathcal{V}g \circ \mathcal{V}f = \mathcal{V}(g \circ f) = \mathrm{id}_{\mathcal{V}P}.$$

It remains to show that

$$\mathcal{V}f \circ \mathcal{V}g \sqsubseteq \mathrm{id}_{\mathcal{V}Q},$$

but this can be written as

$$\forall \mu (\mathcal{V}f \circ \mathcal{V}g)\mu \sqsubseteq \mu$$

i.e. for all evaluations μ and open sets O

$$\mu(f^{-1}(g^{-1}(O))) \leq \mu(O)$$

but this follows from $f \circ g \sqsubseteq \mathrm{id}_Q$, since $f^{-1}(g^{-1}(O)) = \{x \mid f(g(x)) \in O\} \subseteq O$ hence the above equation follows from the monotonicity of μ.

We define the new function space functor $\rightarrow_{\mathcal{V}}: \mathbf{Ipo}^E \times \mathbf{Ipo}^E \rightarrow \mathbf{Ipo}^E$ on objects as

$$P \rightarrow_{\mathcal{V}} Q = \{f: P \rightarrow \mathcal{V}Q \mid f \text{ Scott continuous}\}.$$

For morphisms we consider the composition of partial functions, $\mathcal{V}(g) \circ h \circ f^R$ where $f: P \lhd P'$ and $g: Q \lhd Q'$ and h is in $P \rightarrow_{\mathcal{V}} Q$. This is a partial function $k: P' \rightharpoonup \mathcal{V}(Q')$, but we can transform it to a total function by defining $\alpha(k)(x) = k(x)$ if $k(x)$ exists, otherwise $\alpha(k)(x) = O \mapsto 0$. So we set

$$(f \rightarrow_{\mathcal{V}} g)(h) = \alpha(\mathcal{V}(g) \circ h \circ f^R)$$

and similarly,

$$(f \rightarrow_{\mathcal{V}} g)^R(k) = \alpha(\mathcal{V}(g^R) \circ k \circ f)$$

since $O \mapsto 0$ is the least element of any $\mathcal{V}(X)$, it is clear that these definitions give continuous functions $(Q \rightarrow_{\mathcal{V}} P) \rightarrow (Q' \rightarrow_{\mathcal{V}} P')$ and vice versa. They are give a projection-pair since $(f \rightarrow_{\mathcal{V}} g)^R \circ (f \rightarrow_{\mathcal{V}} g)(h) = \alpha(\mathcal{V}(g^R \circ g) \circ h \circ f^R \circ f) = h$ and similarly $(f \to_{\mathcal{V}} g) \circ (f \to_{\mathcal{V}} g)^R(h) \sqsubseteq h$. Similarly, it is easy to show that $\to_{\mathcal{V}}$ is a functor.

We now want to show that $\to_{\mathcal{V}}$ is continuous in the sense that whenever

$$\rho\colon \langle P_n, f_n\rangle \lhd P$$

and

$$\nu\colon \langle Q_n, g_n\rangle \lhd Q$$

are universal cones, as defined in [36], then so is

$$\rho \to_{\mathcal{V}} \nu\colon \langle P_n \to_{\mathcal{V}} Q_n, f_n \to_{\mathcal{V}} g_n\rangle \lhd P \to_{\mathcal{V}} Q.$$

We call this *universal cone continuity.*

To prove the universal cone continuity we shall first prove a local continuity property for probabilistic function space, namely that given two increasing sequences of continuous, partial functions $f_n\colon P' \rightharpoonup P$ and $g_n\colon Q \rightharpoonup Q'$, then the obvious definition of $f_n \to_{\mathcal{V}} g_n$ (i.e. $f_n \to_{\mathcal{V}} g_n = \lambda h.\alpha(\mathcal{V}g_n \circ h \circ f_n)$) satisfies

$$\bigsqcup_n (f_n \to_{\mathcal{V}} g_n) = (\bigsqcup_n f_n) \to_{\mathcal{V}} (\bigsqcup_n g_n).$$

We first note that lubs of functions are given pointwise, so

$$\bigsqcup_n (f_n \to_{\mathcal{V}} g_n)(h)(x)(O) = \bigsqcup_n \mathcal{V}g_n(h(f_n(x))(O) = \bigsqcup_n h(f_n(x))(g_n^{-1}(O))$$

if $f_n(x)$ exists, 0 otherwise. But, assuming $f_n(x)$ exists (at least for sufficiently large n),

$$\begin{aligned}\bigsqcup_n h(f_n(x))(g_n^{-1}(O)) &= \bigsqcup_n(\bigsqcup_m h(f_n(x))(g_m^{-1}(O))).\\ = \bigsqcup_n(h(f_n(x))(\bigcup_m g_m^{-1}(O))) &= \bigsqcup_n(h(f_n(x))(\bigsqcup g_m)^{-1}(O))\end{aligned}$$

since if $g_n^{-1}(O) = A_n$ thus $A_n = \{x \mid g_n(x) \in O\}$, we can see $\bigcup A_n = (\bigsqcup g_n)^{-1}(O)$ since $x \in \bigcup A_n$ is equivalent to $\exists n$ s. t. $g_n(x) \in O$ which is clearly equivalent to $\bigsqcup g_n(x) \in O$ since O is open. Finally since h is continuous,

$$\begin{aligned}\bigsqcup_n(h(f_n(x))(\bigsqcup g_m)^{-1}(O)) &= h((\bigsqcup_n f_n)(x))(\bigsqcup g_m)^{-1}(O)\\ &= (\bigsqcup_n f_n) \to_{\mathcal{V}} (\bigsqcup_m g_m)(h)(x)(O).\end{aligned}$$

Now we can prove universal cone continuity exactly as in [39] by recalling that it is sufficient to prove that ρ is a universal cone to show $\bigsqcup_n \rho_n \circ \rho_n^R = \mathrm{id}_P$. Thus we need to prove that

$$\bigsqcup_n (\rho_n \to_{\mathcal{V}} \nu_n) \circ (\rho_n \to_{\mathcal{V}} \nu_n)^R = \mathrm{id}_{P \to_{\mathcal{V}} Q}$$

but $(\rho_n \to_{\mathcal{V}} \nu_n) \circ (\rho_n \to_{\mathcal{V}} \nu_n)^R = (\rho_n \circ \rho_n^R) \to_{\mathcal{V}} (\nu_n \circ \nu_n^R)$ so by the local continuity property we see that

$$\bigsqcup_n (\rho_n \to_{\mathcal{V}} \nu_n) \circ (\rho_n \to_{\mathcal{V}} \nu_n)^R = (\bigsqcup_n \rho_n \circ \rho_n^R) \to_{\mathcal{V}} (\bigsqcup_m \nu_m \circ \nu_m^R) = \mathrm{id}_P \to_{\mathcal{V}} \mathrm{id}_Q$$

since ρ is a universal cone, hence the result since $\mathrm{id}_P \to_{\mathcal{V}} \mathrm{id}_Q = \mathrm{id}_{P \to_{\mathcal{V}} Q}$.

Thus we have shown that, with this new function space construct, recursive domain equations can be solved as before; so all closed type expressions involving sums, products, probabilistic function space and recursion denote ipos in the category of embeddings. Note that for any closed recursive type $\mu P.\sigma(P)$, then $\sigma(P)$ is represented by a (cone) continuous functor F and $\mu P.\sigma(P)$ denotes the ipo X which is the colimit of the sequence $F^n(\emptyset)$ and there is an isomorphism $\theta_F \colon F(X) \cong X$. This isomorphism and its inverse will be used later in the semantics of recursive elements.

8.2 Some Functions

In this section we give some useful functions which we will need to express the semantics of our metalanguage, like the product and projection functions.

For products we need a function which gives a evaluation on the product of ipos $X_1, \ldots, X_n$ from evaluations on each X_i, and also projection functions which give a evaluation on X_i for each i from a evaluation on the product. We use the product as defined in Section 3.10, extended to products of finite numbers of ipos in the obvious way. The product also arises from the tensorial strength in Section 4.3, where it is shown to be a continuous function $\mathcal{V}(X_1) \times \ldots \times \mathcal{V}(X_n) \to \mathcal{V}(X_1 \times \ldots \times X_n)$. We also note that given a rectangle of open sets, say $(A_1 \times \ldots \times A_n)$, and evaluations μ_i in $\mathcal{V}(X_i)$ then

$$(\mu_1 \otimes \ldots \otimes \mu_n)(A_1 \times \ldots \times A_n) = \prod_{i=1}^{n} \mu_i(A_i).$$

We define the probabilistic projection function $\mathcal{V}\pi_i$ as the application of $\mathcal{V}$ to the usual projection function $\pi_i(\langle x_1, \ldots, x_n \rangle) = x_i$, this gives

$$\pi_i(\mu) = \lambda A.\mu(X_1 \times \ldots X_{i-1} \times A \times X_{i+1} \times \ldots \times X_n)$$

since $\pi_i^{-1}(A) = X_1 \times \ldots X_{i-1} \times A \times X_{i+1} \times \ldots \times X_n$.

Now for sums we will need functions for in_i and for the cases operation. For in we consider $\mathcal{V}\mathrm{in}_i$ where in_i is the function mapping x to (i, x) hence

$$\mathcal{V}\mathrm{in}_i(\mu) = \lambda O \in X_1 + \ldots + X_n.\mu(O|_{X_i})$$

since clearly $\mathrm{in}_i^{-1}(O) = O|_{X_i}$. For cases we will need to use the function $\Box$, which has the type $(X_1 \to Y) \times \ldots \times (X_n \to Y) \to (X_1 + \ldots + X_n) \to Y$ and is given by

$$\Box(f_1, \ldots, f_n)(\mathrm{in}_i(x)) = f_i(x)$$

where $\mathrm{in}_i(x)$ ranges over $X_1 + \ldots + X_n$ for varying i and x.

We will freely use the functor $\mathcal{V}$, the injection maps η_X and extensions of functions $f^\dagger$ in the semantics. All these are as defined in Chapter 4.

8.3 The Language

In this section we define the metalanguage, first giving the types, then the language and finally the typing rules.

We have the usual type constructions of product, sum, recursive types and a variant of function space, the "probabilistic" function space $\to_{\mathcal{V}}$. Lifting is omitted.The types are given by

$$\sigma ::= \times(\sigma_1, \ldots, \sigma_n) \mid +(\sigma_1, \ldots, \sigma_n) \mid \varphi \mid P \mid \mu P.\sigma$$

and function types

$$\varphi ::= \sigma \to_{\mathcal{V}} \sigma'.$$

We have two types of expressions defined in terms of each other, expressions and function expressions. Expressions include introduction and elimination constructions for products, sums, function space and recursion and application and function expressions include abstraction and recursion. We add the "probabilistic choice" operator or_p, thus giving:

$$\begin{array}{rcl} e & ::= & e_1 \,\mathrm{or}_p\, e_2 \mid \langle e_1, \ldots e_n \rangle \mid \pi_i(e) \mid \mathrm{in}_i(e) \mid \text{cases } e \text{ of } fe_1, \ldots, fe_n \mid \\ & & \lambda fe \mid \text{let } f{:}\,\varphi \text{ be } e \text{ in } e' \mid x \mid fe(e) \mid \mathrm{intro}(e) \mid \mathrm{elim}(e) \\ fe & ::= & (x \in \sigma).e \mid f \mid \mu f{:}\,\varphi.fe \end{array}$$

Expressions in the language are typed relative to assumptions of the form $x \in \sigma$ and $f{:}\,\varphi$. The rules are the same as those in [31], except for the rule for or_p, which is

$$\frac{e_1 \in \sigma \quad e_2 \in \sigma}{e_1 \,\mathrm{or}_p\, e_2 \in \sigma}$$

For products,

$$\frac{e_1 \in \sigma_1, \ldots, e_n \in \sigma_n}{\langle e_1, \ldots, e_n \rangle \in \sigma_1 \times \ldots \times \sigma_n}$$

$$\frac{e \in \sigma_1 \times \ldots \times \sigma_n}{\pi_i(e) \in \sigma_i}.$$

For sums,

$$\frac{e \in \sigma_i}{\mathrm{in}_i(e) \in \sigma_1 + \ldots + \sigma_n}$$

$$\frac{e \in \sigma_1 + \ldots + \sigma_n \ \ \mathit{fe}_i\colon \sigma_i \to_{\mathcal{V}} \sigma \ \ (i = 1, n)}{\text{cases } e \text{ of } \mathit{fe}_1, \ldots, \mathit{fe}_n \in \sigma}.$$

For functions,

$$\frac{\mathit{fe}\colon \varphi}{\lambda \mathit{fe} \in \varphi}$$

$$\frac{\begin{matrix} & (f\colon \varphi) \\ & \vdots \\ e \in \varphi & e' \in \sigma \end{matrix}}{\text{let } f\colon \varphi \text{ be } e \text{ in } e' \in \sigma}.$$

Application,

$$\frac{\mathit{fe}\colon \sigma \to_{\mathcal{V}} \sigma' \ \ e \in \sigma}{\mathit{fe}(e) \in \sigma'}.$$

Recursive Types,

$$\frac{e \in \sigma[\mu P.\sigma(P)]}{\mathrm{intro}(e) \in \mu P.\sigma(P)}$$

$$\frac{e \in \mu P.\sigma(P)}{\mathrm{elim}(e) \in \sigma[\mu P.\sigma(P)]}.$$

Then for the function expressions, abstraction,

$$\frac{\begin{matrix} (x \in \sigma) \\ \vdots \\ e \in \sigma' \end{matrix}}{(x \in \sigma).e\colon \sigma \to_{\mathcal{V}} \sigma'}$$

and recursion,

$$\frac{\begin{matrix} (f\colon \varphi) \\ \vdots \\ \mathit{fe}\colon \varphi \end{matrix}}{\mu f\colon \varphi.\mathit{fe}\colon \varphi}.$$

8.4 Defining Denotational Semantics

We now give the metalanguage a denotational semantics. In Section 8.1 we showed how closed type expressions denote ipos in the category of embeddings, for instance $\sigma \times \tau$ denotes the product of the ipos that σ and τ denote. We use the notation $[\![\sigma]\!]$ to mean the ipo which σ denotes. We will use type assignment functions α and β which assign types to variables x and function variables f. We say an expression e has (closed) type σ relative to α, β, if we can derive $e \in \sigma$ from the rules above plus assumptions of form $x \in \sigma'$ where $\alpha(x) = \sigma'$ and $f{:}\,\varphi$ where $\beta(f) = \varphi$, and similarly for function expressions. Denotations are given relative to an environment ρ where for any variable x if $\alpha(x) = \sigma$ then $\rho(x) \in [\![\sigma]\!]$ and if $\beta(f) = \varphi$ then $\rho(f) \in [\![\varphi]\!]$. If an expression has type σ relative to α and β, then its denotational semantics relative to α, β and ρ is an element of $\mathcal{V}[\![\sigma]\!]$. If a functional expression has type φ, its denotation is an element of $[\![\varphi]\!]$, that is a continuous function $[\![\sigma]\!] \to \mathcal{V}[\![\sigma']\!]$ where $\varphi = \sigma \to_{\mathcal{V}} \sigma'$. We use the notation $\rho[x = a]$ to mean the function $\rho[x = a](y) = \rho(y)$ (for $y \neq x$) and $\rho[x = a](x) = a$, that is ρ plus the added assumption that $\rho(x) = a$, similarly with α and β.

$$
\begin{array}{rcl}
[\![e_1 \text{ or}_p\, e_2; \alpha, \beta]\!]\rho & = & p[\![e_1; \alpha, \beta]\!]\rho + (1-p)[\![e_2; \alpha, \beta]\!]\rho \\
[\![\langle e_1, \ldots e_n\rangle; \alpha, \beta]\!]\rho & = & [\![e_1; \alpha, \beta]\!]\rho \otimes \ldots \otimes [\![e_n; \alpha, \beta]\!]\rho \\
[\![\pi_i(e); \alpha, \beta]\!]\rho & = & \mathcal{V}\pi_i[\![e; \alpha, \beta]\!]\rho \\
[\![\text{in}_i(e); \alpha, \beta]\!]\rho & = & \mathcal{V}\text{in}_i[\![e; \alpha, \beta]\!]\rho \\
[\![\text{cases } e \text{ of } fe_1, \ldots, fe_n; \alpha, \beta]\!]\rho & = & (\Box([\![fe_1; \alpha, \beta]\!]\rho, \ldots, [\![fe_n; \alpha, \beta]\!]\rho))^{\dagger}[\![e; \alpha, \beta]\!]\rho \\
[\![\lambda fe; \alpha, \beta]\!]\rho & = & \eta([\![fe; \alpha, \beta]\!]\rho) \\
[\![\text{let } f{:}\,\varphi \text{ be } e \text{ in } e'; \alpha, \beta]\!]\rho & = & (\lambda g{:}\,\varphi.[\![e'; \alpha, \beta[f = \varphi]]\!]\rho[f = g])^{\dagger}[\![e; \alpha, \beta]\!]\rho \\
[\![x; \alpha, \beta]\!]\rho & = & \eta(\rho(x)) \\
[\![fe(e); \alpha, \beta]\!]\rho & = & [\![fe; \alpha, \beta]\!]\rho^{\dagger}[\![e; \alpha, \beta]\!]\rho \\
[\![\text{intro}(e); \alpha, \beta]\!]\rho & = & \mathcal{V}\theta[\![e; \alpha, \beta]\!]\rho
\end{array}
$$

$$
\begin{array}{rcl}
[\![\mathrm{elim}(e);\alpha,\beta]\!]\rho & = & \mathcal{V}\theta^{-1}[\![e;\alpha,\beta]\!]\rho \\
[\![(x\in\sigma).e;\alpha,\beta]\!]\rho & = & \lambda a\in[\![\sigma]\!][\![e;\alpha[x=\sigma],\beta]\!]\rho[x=a] \\
[\![f;\alpha,\beta]\!]\rho & = & \rho(f) \\
[\![\mu f\colon\varphi.fe;\alpha,\beta]\!]\rho & = & \bigsqcup f_n \\
 & & \text{where } f_0=\lambda x.0,\, f_{n+1}=[\![fe;\alpha,\beta[f=\varphi]]\!]\rho[f=f_n]
\end{array}
$$

As usual, we will omit α, β and even ρ when they are understood. We note that the denotation of an expression is continuous in ρ, and furthermore for closed expressions and function expressions, their denotations are independent of ρ, α and β.

We will need a substitution and a recursion lemma.

Lemma 8.1 *(Substitution) For an expression $e[x]$, and a canonical expression $c \in \sigma$, where the denotation of the canonical expression is $\eta(a)$,*

$$[\![e[c];\alpha,\beta]\!]\rho = [\![e[x];\alpha[x=\sigma],\beta]\!]\rho[x=a]$$

and for any two function expressions fe$[f]$ and fe':φ,

$$[\![fe[fe'];\alpha,\beta]\!]\rho = [\![fe[f];\alpha,\beta[f=\varphi]]\!]\rho[f=[\![fe';\alpha,\beta]\!]\rho].$$

Proof Proof by structural induction on e and *fe*. See the next section for a definition of canonical expressions and remark that the denotation of a canonical expression is always a point evaluation.

Lemma 8.2 *(Recursion) For any function expression fe$[f]$,*

$$[\![\mu f\colon\varphi.fe[f];\alpha,\beta]\!]\rho = [\![fe[\mu f\colon\varphi.fe[f]];\alpha,\beta]\!]\rho.$$

Proof We know that

$$[\![\mu f\colon\varphi.fe[f];\alpha,\beta]\!]\rho = \bigsqcup_n f_n$$

where $f_0 = \lambda x.0$ and $f_{n+1} = [\![fe[f]; \alpha, \beta[f = \varphi]]\!]\rho[f = f_n]$ and by the substitution lemma we know that

$$[\![fe[\mu f\colon \varphi.fe]]\!] = [\![fe[f]]\!]\rho[f = [\![\mu f\colon \varphi.fe[f]]\!]\rho]$$

then by continuity of $[\![e]\!]\rho$ as a function of $\rho(f)$ we see that

$$[\![fe[\mu f\colon \varphi.fe]]\!] = \bigsqcup [\![fe]\!]\rho[f = f_n] = \bigsqcup f_{n+1}$$

so the lemma is proved. ∎

8.5 Operational Semantics

We define a set of deterministic canonical elements,

$$c ::= \langle c_1, \ldots, c_n \rangle \mid \mathrm{in}_i(c) \mid \lambda fe \mid \mathrm{intro}(c)$$

then we give the operational semantics in terms of a relationship written $e \Rightarrow \sum r_i c_i$ over expressions, where r_i is any finite set of positive real numbers such that $\sum r_i \leq 1$. A similar relationship for function expressions is written $\langle fe, c \rangle \Rightarrow \sum r_i c_i$. The rules are as follows :-

$$\frac{e_1 \Rightarrow \sum r_i c_i \quad e_2 \Rightarrow \sum r_j c_j}{e_1 \ \mathrm{or}_p \ e_2 \Rightarrow p \sum r_i c_i + (1-p) \sum r_j c_j}$$

$$\frac{e_i \Rightarrow \sum_{j=1}^{N_i} r_j^i c_j^i}{\langle e_1, \ldots, e_n \rangle \Rightarrow \sum_{j(i)} \left(\prod_{i=1}^{n} r_{j(i)}^i \right) \langle c_{j(1)}^1, \ldots, c_{j(n)}^n \rangle}$$

(here note that the summing over $j(i)$ meaning summing over all functions which map each i to an integer between 1 and N_i.)

$$\frac{e \Rightarrow \sum r_j \langle c_j^1, \ldots, c_j^n \rangle}{\pi_i(e) \Rightarrow \sum r_j c_j^i}$$

$$\frac{e \Rightarrow \sum r_j c_j}{\mathrm{in}_i(e) \Rightarrow \sum r_j \mathrm{in}_i(c_j)}$$

$$\frac{e \Rightarrow \sum r_j \mathrm{in}_{i(j)}(c_j) \quad \langle fe_{i(j)}, c_j\rangle \Rightarrow \sum_k r^j_k c^j_k}{\text{cases } e \text{ of } fe_1, \ldots, fe_n \Rightarrow \sum_{j,k} r_j r^j_k c^j_k}$$

$$\lambda fe \Rightarrow 1.\lambda fe$$

$$\frac{e \Rightarrow \sum_i r_i fe_i \quad e'[f/fe_i] \Rightarrow \sum_k r^i_k c^i_k}{\text{let } f{:}\,\varphi \text{ be } e \text{ in } e' \Rightarrow \sum_{i,k} r_i r^i_k c^i_k}$$

$$\frac{e \Rightarrow r_i c_i \quad \langle fe, c_i\rangle \Rightarrow \sum_j r^i_j c^i_j}{fe(e) \Rightarrow \sum_{i,j} r_i r^i_j c^i_j}$$

$$\frac{e \Rightarrow \sum r_i \mathrm{intro}(c_i)}{\mathrm{elim}(e) \Rightarrow \sum r_i c_i}$$

$$\frac{e \Rightarrow \sum r_i c_i}{\mathrm{intro}(e) \Rightarrow \sum r_i \mathrm{intro}(c_i)}$$

$$\frac{e[c/x] \Rightarrow \sum r_i c_i}{\langle (x \in \sigma)e, c\rangle \Rightarrow \sum r_i c_i}$$

$$\frac{\langle fe(\mu f{:}\,\varphi.fe), c\rangle \Rightarrow \sum r_i c_i}{\langle \mu f{:}\,\varphi.fe, c\rangle \Rightarrow \sum r_i c_i}$$

$$e \Rightarrow 0.c \qquad (c \in \sigma \text{ with } e \in \sigma)$$

$$\langle fe, c\rangle \Rightarrow 0.c' \quad (c \in \sigma,\ c' \in \tau \text{ and } fe{:}\,\sigma \rightarrow_\nu \tau)$$

These rules define a non-deterministic transition relation, non-deterministic in that one of the last two rules can be applied to any expression or function expression instead of the appropriate structure rule. These rules guarantee that every expression "evaluates", even ones whose denotational semantics is $\bot = O \mapsto 0$.

We can associate each $\sum r_i c_i$ with an element of $\mathcal{V}X$ via the semantics by $[\![\sum r_i c_i]\!] = \sum r_i [\![c_i]\!]$. In fact we can also show that the semantics of a canonical element always gives a point evaluation; this is simple since it is trivial for λfe, the product of point evaluations $\eta(x_i)$ is a point evaluation at $\langle x_1, \ldots, x_n \rangle$, for injections $\mathrm{in}_i(\eta(x)) = \eta(\mathrm{in}_i(x))$ and finally recall that $[\![\mathrm{intro}(c)]\!] = \mathcal{V}\theta[\![c]\!]$, so as $\mathcal{V}\theta(\eta(x)) = \eta(\theta(x))$, the result follows inductively. Thus we can also conclude that $[\![c]\!](X) = 1$, this will be useful later.

The equivalence of denotational and operational semantics for this metalanguage is clearly going to be of the form

$$\bigsqcup_{e \Rightarrow \sum r_i c_i} \sum_i r_i [\![c_i]\!] = [\![e]\!]$$

and for function expressions,

$$\bigsqcup_{\langle fe, c \rangle \Rightarrow \sum r_i c_i} \sum_i r_i [\![c_i]\!] = [\![fe]\!]^{\dagger}([\![c]\!]).$$

We can show that the left hand side is directed by an easy induction on the operational semantics.

Proving this equality requires the following theorem, proved by a straightforward induction on the length of proof of the $\Rightarrow$ relation.

Theorem 8.3 *For any expression e if $e \Rightarrow \sum r_i c_i$ then*

$$\sum r_i [\![c_i]\!] \sqsubseteq [\![e]\!]$$

and if $\langle fe, c \rangle \Rightarrow \sum r_i c_i$ then

$$\sum r_i [\![c_i]\!] \sqsubseteq [\![fe]\!]^{\dagger}([\![c]\!])$$

Proof Proof by induction on the size of the proof that e or $\langle fe, c\rangle \Rightarrow \sum r_i c_i$

Suppose the last rule applied is

$$\frac{e_1 \Rightarrow \sum r_i c_i \quad e_2 \Rightarrow \sum r_j c_j}{e_1 \text{ or}_p\, e_2 \Rightarrow p \sum r_i c_i + (1-p)\sum r_j c_j}.$$

By the inductive hypothesis, the theorem holds for $e_1 \Rightarrow \sum r_i c_i$, i.e. $\sum r_i [\![c_i]\!] \sqsubseteq [\![e_1]\!]$ and similarly for e_2, then

$$p\sum r_i c_i + (1-p)\sum r_j c_j \sqsubseteq p[\![e_1]\!] + (1-p)[\![e_2]\!] = [\![e_1 \text{ or}_p\, e_2]\!]$$

as required.

Now suppose the last rule applied was

$$\frac{e_i \Rightarrow \sum_{j=1}^{N_i} r^i_j c^i_j}{\langle e_1, \ldots, e_n\rangle \Rightarrow \sum_{j(i)} \prod_{i=1}^{n} r^i_{j(i)} \langle c^1_{j(1)}, \ldots, c^n_{j(n)}\rangle}$$

then we assume that for each i,

$$\sum_{j=1}^{N_i} r^i_j . [\![c^i_j]\!] \sqsubseteq [\![e_i]\!]$$

but then

$$[\![\langle e_1, \ldots, e_n\rangle]\!] = [\![e_1]\!] \otimes \ldots \otimes [\![e_n]\!]$$

hence

$$\left(\sum_{j=1}^{N_1} r^1_j . [\![c^1_j]\!]\right) \otimes \ldots \otimes \left(\sum_{j=1}^{N_n} r^n_j . [\![c^n_j]\!]\right) \sqsubseteq [\![\langle e_1, \ldots, e_n\rangle]\!]$$

but since $\otimes$ is bilinear we can rearrange the L.H.S. to be

$$\sum_{j(i)} \left(\prod_{i=1}^{n} r^i_{j(i)}\right) [\![\langle c^1_{j(1)}, \ldots, c^n_{j(n)}\rangle]\!] \sqsubseteq [\![\langle e_1, \ldots, e_n\rangle]\!]$$

as required.

Now suppose the last rule applied is

$$\frac{e \Rightarrow \sum r_j \langle c^1_{j(1)}, \ldots, c^n_{j(n)}\rangle}{\pi_i(e) \Rightarrow \sum r_j c^i_j}$$

so we know that

$$\sum r_j [\![\langle c^1_{j(1)}, \ldots, c^n_{j(n)}\rangle]\!] \sqsubseteq [\![e]\!]$$

by the inductive hypothesis. But then we know

$$\sum r_j [\![c^i_j]\!] \sqsubseteq [\![\pi_i(e)]\!]$$

since for any open set O, $\sum r_j [\![c^i_j]\!](O) = \sum_{c_j \in O} r_j$ and if we define O' as the open rectangle $X_1 \times \ldots \times X_{i-1} \times O \times X_{i+1} \times \ldots \times X_n$ (or $\pi_i^{-1}(O)$), then $[\![\pi_i(e)]\!](O) = [\![e]\!](O')$, and

$$\sum r_j [\![\langle c^1_{j(1)}, \ldots, c^n_{j(n)}\rangle]\!](O') = \sum_{c_j \in O} r_j.$$

Now suppose

$$\frac{e \Rightarrow \sum r_j c_j}{\mathrm{in}_i(e) \Rightarrow \sum r_j \mathrm{in}_i(c_j)}$$

is the last rule applied. Then the proof is trivial since $[\![\mathrm{in}_i(e)]\!] = \mathcal{V}\mathrm{in}_i[\![e]\!]$ and the function in_i is continuous.

Now suppose the last rule applied is

$$\frac{e \Rightarrow \sum r_j \mathrm{in}_{i(j)}(c_j) \quad \langle fe_{i(j)} c_j\rangle \Rightarrow r^j_k c^j_k}{\text{cases } e \text{ of } fe_1, \ldots, fe_n \Rightarrow \sum_{j,k} r_j r^j_k c^j_k}$$

we note that by the definition of $[\![\text{cases}]\!]$ we have

$$[\![\text{cases } e \text{ of } fe_1, \ldots, fe_n]\!] = (\Box([\![fe_1]\!], \ldots, [\![fe_n]\!]))^\dagger [\![e]\!]$$

by using the inductive hypothesis from the first part of the rule we get

$$(\Box([\![fe_1]\!], \ldots, [\![fe_n]\!]))^\dagger \sum r_j [\![\mathrm{in}_{i(j)}(c_j)]\!]) \sqsubseteq (\Box([\![fe_1]\!], \ldots, [\![fe_n]\!]))^\dagger [\![e]\!]$$

and by the definition of $\Box$,

$$(\Box([\![fe_1]\!], \ldots, [\![fe_n]\!]))^\dagger (\sum r_j [\![\mathrm{in}_{i(j)}(c_j)]\!]) = \sum r_j [\![fe_{i(j)}]\!]^\dagger ([\![c_j]\!])$$

and by the inductive hypothesis for each j,

$$\sum_k r^j_k c^j_k \sqsubseteq [\![fe_{i(j)}]\!]^\dagger [\![c_j]\!]$$

thus giving the required inequality.

Now in the case when the rule is

$$\lambda f e \Rightarrow 1.\lambda f e$$

then the theorem says we must have

$$[\![\lambda f e]\!] \sqsubseteq [\![\lambda f e]\!]$$

which is clearly true.

Now suppose the last rule was

$$\frac{e \Rightarrow \sum_i r_i f e_i \quad e'[f/fe_i] \Rightarrow \sum_k r_k^i c_k^i}{\text{let } f\colon \varphi \text{ be } e \text{ in } e' \Rightarrow \sum_{i,k} r_i r_k^i c_k^i}$$

then we need to prove that $\sum_{i,k} r_i r_k^i [\![c_k^i]\!] \sqsubseteq [\![\text{let } f\colon \varphi \text{ be } e \text{ in } e']\!]$. But by the definition of the denotational semantics,

$$[\![\text{let } f\colon \varphi \text{ be } e \text{ in } e']\!] = (\lambda g.[\![e'[f]]\!]\rho[f = g])^{\dagger}[\![e]\!]$$

then by the inductive hypothesis, $\sum_i r_i [\![fe_i]\!] \sqsubseteq [\![e]\!]$, so

$$\sum_i r_i [\![e'[f]]\!]\rho[f = [\![fe_i]\!]] \sqsubseteq [\![\text{let } f\colon \varphi \text{ be } e \text{ in } e']\!]$$

but then by the substitution lemma, $[\![e'[f]]\!]\rho[f = [\![fe_i]\!]] = [\![e'[fe_i]]\!]$ and again by the inductive hypothesis,

$$\sum_k r_k^i [\![c_k^i]\!] \sqsubseteq [\![e'[fe_i]]\!]$$

so

$$\sum_{i,k} r_i r_k^i [\![c_k^i]\!] \sqsubseteq [\![\text{let } f\colon \varphi \text{ be } e \text{ in } e']\!]$$

as required.

If

$$\frac{e \Rightarrow r_i c_i \quad \langle fe, c_i\rangle \Rightarrow \sum_j r_j^i c_j^i}{fe(e) \Rightarrow \sum_{i,j} r_i r_j^i c_j^i}$$

is the last rule applied then we recall

$$[\![fe(e)]\!] = [\![fe]\!]^{\dagger}[\![e]\!]$$

and by the inductive hypothesis and linearity

$$\sum r_i[\![fe]\!]^{\dagger}[\![c_i]\!] = [\![fe]\!]^{\dagger}\sum r_i[\![c_i]\!] \sqsubseteq [\![fe]\!]^{\dagger}[\![e]\!]$$

then by the inductive hypothesis on each term

$$\sum r_i \sum r^i_j[\![c^i_j]\!] \sqsubseteq \sum r_i[\![fe]\!]^{\dagger}[\![c_i]\!]$$

as required.

For the two rules

$$\frac{e \Rightarrow \sum r_i \mathrm{intro}(c_i)}{\mathrm{elim}(e) \Rightarrow \sum r_i c_i}$$

$$\frac{e \Rightarrow \sum r_i c_i}{\mathrm{intro}(e) \Rightarrow \sum r_i \mathrm{intro}(c_i)}$$

the proof is the straightforward application of the continuous functions θ and θ^{-1}.

Now suppose the last rule was

$$\frac{e[c/x] \Rightarrow \sum r_i c_i}{\langle (x \in \sigma)e, c\rangle \Rightarrow \sum r_i c_i}$$

we need to prove that

$$\sum r_i[\![c_i]\!] \sqsubseteq [\![(x \in \sigma)e]\!]^{\dagger}[\![c]\!]$$

from the definition of $[\![(x \in \sigma)e]\!]$ and $\dagger$, we get

$$[\![(x \in \sigma)e]\!]^{\dagger}[\![c]\!] = [\![e]\!]\rho[x = [\![c]\!]]$$

but by the substitution lemma we have that

$$[\![e]\!]\rho[x = [\![c]\!]] = [\![e[c/x]]\!]$$

and by the inductive hypothesis,

$$\sum r_i[\![c_i]\!] \sqsubseteq [\![e[c/x]]\!]$$

hence we are done.

Now suppose the last rule applied is

$$\frac{\langle fe(\mu f\colon \varphi.fe), c\rangle \Rightarrow \sum r_i c_i}{\langle \mu f\colon \varphi.fe, c\rangle \Rightarrow \sum r_i c_i}.$$

This case is easy because by the recursion lemma we know that

$$[\![fe(\mu f\colon \varphi.fe)]\!] = [\![\mu f\colon \varphi.fe]\!].$$

The final two rules are trivial. ∎

We now consider the more difficult part of the equivalence proof. We construct a relation $\precsim_\sigma$ between elements of type $[\![\sigma]\!]$ (note, not $\mathcal{V}[\![\sigma]\!]$) and closed canonical expressions of type σ and for function types between functions $f\colon [\![\sigma]\!] \to \mathcal{V}[\![\tau]\!]$. and closed function expressions of type φ where $\varphi = \sigma \to_{\mathcal{V}} \tau$. We require $\precsim$ to satisfy the following:

products $\langle x_1, \ldots, x_n\rangle \precsim_{\sigma_1 \times \ldots \times \sigma_n} \langle c_1, \ldots, c_n\rangle$ iff $\forall i, x_i \precsim_{\sigma_i} c_i$

sums $\mathrm{in}_i(x) \precsim_{\sigma_1 + \ldots + \sigma_n} \mathrm{in}_j(c)$ iff $i = j, x \precsim_{\sigma_i} c$

recursion $\theta(x) \precsim_{\mu P.\sigma[P]} \mathrm{intro}(c)$ iff $x \precsim_{\sigma[\mu P.\sigma(P)]} c$

functions $\eta(f) \precsim_\varphi \lambda fe$ iff $f \precsim_\varphi fe$ and $f \sqsubseteq [\![fe]\!]$

function types $f \precsim_{\sigma \to_{\mathcal{V}} \tau} fe$ iff $\forall x, c$ s.t. $x \precsim_\sigma c, f(x) \precsim_{\mathcal{V}\tau} fe(c)$

We define $\precsim_{\mathcal{V}\sigma}$ as follows, $\mu \precsim_{\mathcal{V}\sigma} e$ where $e \in \sigma$ and $\mu \in \mathcal{V}[\![\sigma]\!]$ iff μ is in the least Scott closed set containing all linear combinations of point evaluations $\sum_i r_i \eta(x_i)$ where $x_i \precsim_\sigma c_i$ and $e \Rightarrow \sum_i r_i c_i$. This least closed set can be formally defined as the intersection of all Scott closed sets containing these points, since the whole set is closed, there is always at least one such set. So trivially for any expression e, the set of μ such that $\mu \precsim_{\mathcal{V}\sigma} e$ is Scott closed. We also require the relation $x \precsim_\sigma c$

to imply that $\eta(x) \sqsubseteq [\![c]\!]$. This is sufficient to show that $\mu \lesssim_{\mathcal{V}\sigma} e$ implies $\mu \sqsubseteq [\![e]\!]$ since from Theorem 8.3, $e \Rightarrow \sum_i r_i c_i$ implies $\sum r_i [\![c_i]\!] \sqsubseteq [\![e]\!]$ and by the condition above $\sum r_i \eta(x_i) \sqsubseteq \sum r_i [\![c_i]\!]$. So the closed set consisting of all μ such that $\mu \sqsubseteq [\![e]\!]$ contains all the $\sum_i r_i \eta(x_i)$ hence $\mu \lesssim_{\mathcal{V}\sigma} e$ implies $\mu \sqsubseteq [\![e]\!]$. Finally we need that the set of x such that $x \lesssim_\sigma c$ is Scott closed for any canonical element c. We construct such a relation in appendix A.

We now give a lemma which we will use repeatedly in the proof of the main theorem. It gives a condition for when an evaluation μ is in the least set containing a certain collection of points. The proof uses the standard techniques for this kind of transfinite induction.

Lemma 8.4 *Suppose f is a continuous function $f: X_1 \times \ldots \times X_n \to Y$ and C is a closed subset of Y. Then if C_i is the least closed subset of X_i containing a set of points x^i_λ (for λ ranging over some index set) and $f(x^1_{\lambda_1}, \ldots, x^n_{\lambda_n})$ is in C for any $\lambda_1, \ldots, \lambda_n$, then for any set $x_1, \ldots, x_n$ with x_i in C_i, $f(x_1, \ldots, x_n)$ is in C.*

Proof We first consider the case where $n = 1$. We suppose that $f(x^1_\lambda)$ is contained in C for all λ. Then we consider the inverse image of C under f, this is a closed set and by the assumption it contains x^1_λ for all lambda. So since C_1 is the least closed set containing all x^1_λ, then $C_1 \subseteq f^{-1}(C)$, so for all x in C_1, we know $x \in f^{-1}(C)$ i.e. $f(x) \in C$.

For the case where $n > 1$ we consider the inverse image of C under f. Setting $X = X_1 \times \ldots \times X_n$ and applying the result in the previous paragraph shows that if $\underline{x}$ is in the least closed subset of $X_1 \times \ldots \times X_n$ containing $\langle x^1_{\lambda_1}, \ldots, x^n_{\lambda_n} \rangle$ for all $\lambda_1, \ldots, \lambda_n$, then $x \in f^{-1}(C)$, i.e. $f(x) \in C$. We thus just need to prove that the least set generated by the points $\langle x^1_{\lambda_1}, \ldots, x^n_{\lambda_n} \rangle$ is in fact $C_1 \times \ldots \times C_n$ to give the result.

So if we let D be the least closed set generated by $\langle x^1_{\lambda_1}, \ldots, x^n_{\lambda_n} \rangle$ for all $\lambda_1, \ldots, \lambda_n$ then it is trivial that $D \subseteq C_1 \times \ldots \times C_n$ since $C_1 \times \ldots \times C_n$ is Scott closed and clearly contains all the required points. We prove the opposite inclusion by induction on n, the case $n = 1$ is trivial, we assume the result for $n = k$ and try to prove it for $n = k + 1$. Clearly we can consider D as the least closed set generated by the points $\langle x_\lambda, \underline{x}_\gamma \rangle$ where the set of points $\underline{x}_\gamma$ is the set of points $\langle x^2_{\lambda_2}, \ldots, x^{k+1}_{\lambda_{k+1}} \rangle$. Then by the inductive hypothesis, we need to show that $D = C_1 \times D'$ where D' is the least closed set generated by $\underline{x}_\gamma$ and by the inductive hypothesis $D' = C_2 \times \ldots \times C_{k+1}$. To show that $C_1 \times D' \subseteq D$ it is sufficient to show that for all $x \in C_1$, the x cross-section of D, that is $D_x = \{\underline{x} \mid \langle x, \underline{x} \rangle \in D\}$ contains D'. But suppose we define K as the set of all $x \in X_1$ such that $D' \subseteq D_x$. We can see that K is Scott closed, since if $a \sqsubseteq b$ and $b \in K$, then since D is closed $D_b \subseteq D_a$ and similarly if a_λ is a directed subset of K, then

$$\bigcap_\lambda D_{a_\lambda} = D_{\bigsqcup_\lambda a_\lambda}$$

and $D' \subseteq D_{a_\lambda}$ for all λ implies that $D' \subseteq \bigcap_\lambda D_{a_\lambda}$. Furthermore K contains all x^1_λ, since if $a = x^1_\lambda$, then C_a is a closed set which contains all $\underline{x}_\gamma$, therefore it must contain D'. So K contains C_1, that is, for all $x \in C_1$, $D' \subseteq D_x$ as required. ∎

We will sometimes use this lemma with the set x^i_λ being itself a closed set, i.e. the set of functions $\precsim_\varphi fe$.

Theorem 8.5 *Given a relation $\precsim_\sigma$ with the properties as above, for an expression e with free variables $x_1, \ldots, x_n$ and $f_1, \ldots, f_m$, then for any $y_i \precsim_{\sigma_i} c_i$ and $g_k \precsim_{\varphi_k} fe_k$,*

$$[\![e]\!]\rho[x_i = y_i, f_k = g_k] \precsim_\sigma e[c_i/x_i, fe_k/f_k]$$

and for any function expression fe with free variables $x_1, \ldots, x_n$ and $f_1, \ldots, f_m$, then for any $y_i \precsim_{\sigma_i} c_i$ and $g_k \precsim_{\varphi_k} fe_k$

$$[\![fe]\!]\rho[x_i = y_i, f_k = g_k] \precsim_\varphi fe[c_i/x_i, fe_k/f_k].$$

Proof When the free variables and substitutions are obvious from the context we abbreviate $e[c_i/x_i, fe_k/f_k]$ to $\tilde{e}$, $fe[c_i/x_i, fe_k/f_k]$ to $\widetilde{fe}$ and $\rho[x_i = y_i, f_k = g_k]$ to $\tilde{\rho}$. The proof is by induction on the structure of expressions and function expressions.

probabilistic choice Suppose $e = e_1 \text{ or}_p\, e_2$ and by the inductive hypothesis we know $[\![e_1]\!]\tilde{\rho} \precsim_{\mathcal{V}_\sigma} \tilde{e}_1$ and $[\![e_2]\!]\tilde{\rho} \precsim_{\mathcal{V}_\sigma} \tilde{e}_2$. We need to show that $[\![e]\!]\tilde{\rho} \precsim_{\mathcal{V}_\sigma} \tilde{e}$. Consider the closed set C of $\mu \precsim_{\mathcal{V}_\sigma} \tilde{e}$, and the continuous function $(\mu_1, \mu_2) \mapsto p\mu_1 + (1-p)\mu_2$. By Lemma 8.4, we can see that $p\mu_1 + (1-p)\mu_2$ is in C for any $\mu_i \precsim_{\mathcal{V}_\sigma} \tilde{e}_i$, if it is in C whenever $\mu_1 = \sum_i r_i \eta(x_i)$ with $x_i \precsim_\sigma c_i$ and $e_1 \Rightarrow \sum_i r_i c_i$ and $\mu_2 = \sum_j r_j \eta(x_j)$ with $x_j \precsim_\sigma c_j$ and $e_2 \Rightarrow \sum_j r_j c_j$. But then $p\mu_1 + (1-p)\mu_2 = p \sum_i r_i \eta(x_i) + (1 - p)\sum_j r_j \eta(x_j)$ which is in C since $e \Rightarrow p\sum_i r_i c_i + (1-p)\sum_j r_j c_j$, $x_i \precsim_\sigma c_i$, and $x_j \precsim_\sigma c_j$.

products Suppose $e = \langle e_1, \ldots, e_n \rangle$. Let C be the set of evaluations $\precsim \tilde{e}$ and C_i for $\tilde{e}_i$. Consider the function $(\mu_1, \ldots, \mu_n) \mapsto \mu_1 \otimes \ldots \otimes \mu_n$, by Lemma 8.4 to show that $\mu_i \in C_i$ implies $\mu_1 \otimes \ldots \otimes \mu_n \in C$ we merely need to check it for μ_i of the form $\sum_j r^i_j \eta(x^i_j)$ where $x^i_j \precsim c^i_j$ and $\tilde{e}_i \Rightarrow \sum_j r^i_j c^i_j$. But then the image of a evaluation of this form is $\sum_{j(i)} (\prod_{i=1}^n r^i_{j(i)}) \langle \eta(x^1_{j(1)}), \ldots, \eta(x^n_{j(n)}) \rangle$ and we know that $\tilde{e} \Rightarrow \sum_{j(i)} (\prod_{i=1}^n r^i_{j(i)}) \langle c^1_{j(1)}, \ldots, c^n_{j(n)} \rangle$ and $\langle x^1_{j(1)}, \ldots, c^n_{j(n)} \rangle \precsim \langle c^1_{j(1)}, \ldots, c^n_{j(n)} \rangle$ by the condition for products on $\precsim$ so the image is in C. Hence, since by the inductive hypothesis, $[\![e_i]\!]\tilde{\rho} \in C_i$, we see that $[\![e]\!]\tilde{\rho} \in C$ as required.

Now we consider $\pi_k(e)$. We apply Lemma 8.4 to the function $\mathcal{V}\pi_k$, to show that when $[\![e]\!]\tilde{\rho}$ is in the smallest closed set containing all the evaluations of the form $\sum_i r_i \eta(\langle x^1_i, \ldots, x^n_i \rangle)$ such that $\langle x^1_i, \ldots, x^n_i \rangle \precsim \langle c^1_i, \ldots, c^n_i \rangle$ and $\tilde{e} \Rightarrow \sum_i r_i \langle c^1_i, \ldots, c^n_i \rangle$, then $\mathcal{V}\pi_k [\![e]\!]\tilde{\rho} = [\![\pi_k(e)]\!]\tilde{\rho}$ is in the corresponding closed set. But the image of such a point is a evaluation $\sum_i r_i \eta(x^k_i)$ where $x^k_i \precsim c^k_i$ and $\widetilde{\pi_k(e)} \Rightarrow \sum_i r_i c^k_i$, hence it is in this closed set and $[\![\pi_i(e)]\!]\tilde{\rho} \precsim \widetilde{\pi_k(e)}$ as required.

sums Consider $\text{in}_k(e)$. As in the previous example we use Lemma 8.4 with the function $\mathcal{V}\text{in}_k$ so we merely need to show that whenever $\tilde{e} \Rightarrow \sum_i r_i c_i$, and $x_i \precsim c_i$, then $\mathcal{V}\text{in}_k(\sum r_i \eta(x_i)) \precsim \widetilde{\text{in}_k(e)}$. But if this is true then $\widetilde{\text{in}_k(e)} \Rightarrow \sum_i r_i \text{in}_k(c_i)$ and $\text{in}_k(x_i) \precsim \text{in}_k(c_i)$ so $\sum_i r_i \eta(\text{in}_k(x_i)) = \mathcal{V}\text{in}_k \sum_i r_i \eta(x_i) \precsim \widetilde{\text{in}_k(e)}$ as required.

Consider $e' = \text{cases } e \text{ of } fe_1, \ldots fe_n$. Let C be the set of evaluations $\precsim \tilde{e'}$ and C' be the set $\precsim \tilde{e}$. Consider the function $(f_1, \ldots, f_n, \mu) \mapsto \Box(f_1, \ldots, f_n)^\dagger \mu$. We need to show that whenever $f_i \precsim fe_i$ and $\mu \in C$, that $\Box(f_1, \ldots, f_n)^\dagger \mu$ is in C. Since the set of $f_i \precsim fe_i$ is Scott closed, we can apply Lemma 8.4, to see that it is sufficient to check that $\Box(f_1, \ldots, f_n)^\dagger \mu$ is in C for all μ of the form $\sum_i r_i \eta(x_i)$ where $x_i \precsim c_i$ and $\tilde{e} \Rightarrow \sum_i r_i c_i$. Since $e \in \sigma_1 + \ldots + \sigma_n$ then each c_i is actually $\text{in}_{k(i)}(\bar{c}_i)$, and by the condition on the relation $x_i = \text{in}_{k(i)}(\bar{x}_i)$ and $\bar{x}_i \precsim \bar{c}_i$, then $\Box(f_1, \ldots, f_n)^\dagger \sum_i r_i \eta(x_i) = \sum_i r_i f_k(\bar{x}_i)$. But then since $\bar{x}_i \precsim \bar{c}_i$, by the condition on $\precsim$ for functions $f_k(\bar{x}_i) \precsim fe_k(\bar{c}_i)$, so we can apply Lemma 8.4 again, and we only need to show that $\sum_i r_i f_k(\bar{x}_i)$ is in C for $f_k(\bar{x}_i)$ equal to $\sum r^i_j \eta(x^i_j)$ where $\langle fe_k, \bar{c}_i \rangle \Rightarrow \sum r^i_j c^i_j$ and $x^i_j \precsim c^i_j$. But C contains $\sum r_i \sum_j r^i_j \eta(x^i_j)$, since we know that $e' \Rightarrow \sum_{i,j} r_i r^i_j c^i_j$ and $x^i_j \precsim c^i_j$ so we're done.

functions Consider $e = \lambda fe$. By the inductive hypothesis we know that $[\![fe]\!]\tilde{\rho} \precsim \widetilde{fe}$. Further $\lambda \widetilde{fe} \Rightarrow 1.\lambda \widetilde{fe}$ and by the condition on the relation for functions, $f \precsim \lambda \widetilde{fe}$ iff $f \precsim \widetilde{fe}$ hence $\eta([\![fe]\!]\tilde{\rho}) \precsim \widetilde{fe}$ so we're done since $\eta([\![fe]\!]\tilde{\rho}) = [\![\lambda fe]\!]\tilde{\rho}$.

Consider $e'' = \text{let } f{:}\,\varphi \text{ be } e \text{ in } e'[f]$. Let C be the set of $\mu \precsim \tilde{e''}$ and consider its inverse image under the function $(\mu, k) \mapsto k^\dagger(\mu)$ where $\mu \in \mathcal{V}[\![\varphi]\!]$ and $k{:}\, [\![\varphi]\!] \to_{\mathcal{V}} [\![\sigma]\!]$. We want to show that whenever μ and k satisfy the conditions of the inductive hypothesis, that is $\mu \precsim_\varphi e$ and for any $f \precsim fe$, then $k(f) \precsim_{\mathcal{V}\sigma} e'[fe]$, then $k^\dagger(\mu) \in C$. By Lemma 8.4, it is sufficient to show this for $\mu = \sum r_i \eta(x_i)$ where $x_i \precsim_\varphi c_i$ and $e \Rightarrow \sum_i r_i c_i$. But as c_i has type φ, in fact say $c_i = fe_i$ and $x_i = f_i$. Then $k^\dagger(\mu) = \sum_i r_i k(f_i)$, but by the second condition, since $f_i \precsim fe_i$, we know that $k(f_i) \precsim_{\mathcal{V}\sigma} \widetilde{e'[fe_i]}$. Again by applying Lemma 8.4, it is sufficient to show $\sum r_i f(k_i)$ is in C, by showing that whenever $\widetilde{e'[fe_i]} \Rightarrow \sum_j r^i_j c^i_j$ and $x^i_j \precsim c^i_j$, then $\sum_{i,j} r_i r^i_j \eta(x^i_j)$ is in C. But this is trivial by the definition of C since $\tilde{e}'' \Rightarrow \sum_{i,j} r_i r^i_j c^i_j$ from the definition of the operational semantics.

recursive types If $e' = \mathrm{intro}(e)$ then we apply Lemma 8.4 to the function $\mathcal{V}\theta$. We suppose that $\tilde{e} \Rightarrow \sum_i r_i c_i$ and $x_i \lesssim c_i$, then consider $\mathcal{V}\theta \sum_i r_i \eta(x_i) = \sum_i r_i \eta(\theta(x_i))$ and $\theta(x_i) \lesssim \mathrm{intro}(c_i)$ by the conditions on the relation and by the operational semantics, $\tilde{e}' \Rightarrow \sum_i r_i \mathrm{intro}(c_i)$. So then since $[\![e]\!]\tilde{\rho} \lesssim \tilde{e}$ clearly $\mathcal{V}\theta[\![e]\!]\tilde{\rho} = [\![e']\!]\tilde{\rho}$ is in C.

Similarly if $e' = \mathrm{elim}(e)$ then we apply Lemma 8.4 to the function $\mathcal{V}\theta^{-1}$. We suppose that $\tilde{e} \Rightarrow \sum_i r_i c_i$ and $x_i \lesssim c_i$. Then c_i must actually be $\mathrm{intro}(\bar{c}_i)$ and by the condition on the relation for recursion, $\theta^{-1}(x_i) \lesssim \bar{c}_i$. Then $\mathcal{V}\theta^{-1} \sum_i r_i \eta(x_i) = \sum_i r_i \eta(\theta^{-1}(x_i))$ and $\theta^{-1}(x_i) \lesssim \bar{c}_i$ and by the operational semantics, $\tilde{e}' \Rightarrow \sum_i r_i \bar{c}_i$. Hence $\mathcal{V}\theta^{-1}[\![e]\!]\tilde{\rho} = [\![e']\!]\tilde{\rho} \lesssim \mathrm{elim}(\tilde{e})$ as required since by the inductive hypothesis, $[\![e]\!]\tilde{\rho} \lesssim \tilde{e}$.

application Suppose $e' = fe(e)$, we let C be the set of evaluations $\lesssim \tilde{e'}$. We need to show that if $\mu \lesssim \tilde{e}$ and $f \lesssim \tilde{fe}$, then $f^\dagger(\mu) \in C$. By Lemma 8.4, it is sufficient to show this is true for $\mu = \sum_i r_i \eta(x_i)$ where $x_i \lesssim c_i$ and $\tilde{e} \Rightarrow \sum r_i c_i$. Then $f^\dagger(\mu) = \sum r_i f(x_i)$. But $f \lesssim \tilde{fe}$, by the conditions on the relation implies that $f(x_i) \lesssim \widetilde{fe}(c_i)$ and to show that $\sum r_i f(x_i)$ is in C it is sufficient to show that it is for $f(x_i) = \sum_j r^i_j \eta(x^i_j)$ where $x^i_j \lesssim c^i_j$ and $\langle \widetilde{fe}, c_i \rangle \Rightarrow \sum_j r^i_j c^i_j$. But then $\sum_i r_i f(x_i) = \sum_{i,j} r_i r^i_j \eta(x^i_j)$ and from the operational semantics we know that $fe(e) \Rightarrow \sum_{i,j} r_i r^i_j c^i_j$, so by the definition of C, $\sum_i r_i f(x_i) \in C$ as required.

function expressions: abstraction If $fe = (x \in \sigma)e[x]$ then we want to show that $[\![fe]\!]\tilde{\rho} \lesssim \widetilde{fe}$. By the condition on the relation for function expressions, this is true if for all $x \lesssim_\sigma c$, $[\![fe]\!]\tilde{\rho}(x) \lesssim_{\mathcal{V}\sigma} \widetilde{fe}(c)$. But by the inductive hypothesis, we already know that $[\![fe]\!]\tilde{\rho}(x) = [\![e[y]]\!]\tilde{\rho}[y = x] \lesssim_{\mathcal{V}\sigma} \widetilde{e[c]}$ and by the operational semantics, $\langle \widetilde{fe}, c \rangle \Rightarrow \sum_i r_i c_i$ implies $\widetilde{e[c]} \Rightarrow \sum_i r_i c_i$ so $[\![fe]\!]\tilde{\rho} \lesssim_{\mathcal{V}\sigma} \widetilde{fe}(c)$ as required.

recursion Consider $fe' = \mu f \colon \varphi . fe[f]$. By closure property of $\lesssim$ we only need to show that $f_n \lesssim \widetilde{fe}'$ for all n where $f_0 = \lambda x.W \mapsto 0$ and $f_{n+1} = [\![fe]\!]\tilde{\rho}[f_n = f]$.

The base case is trivial. Suppose we have $f_n \lesssim \widetilde{fe}'$. By the inductive hypothesis then we know $[\![fe]\!]\tilde{\rho}[f_n = f] \lesssim \widetilde{fe}[fe']$, and the left hand side is f_{n+1}. But also we know $\langle \widetilde{fe}[fe'], c\rangle \Rightarrow \sum_{i=1}^{n} r_i c_i$ implies $\langle \widetilde{fe}', c\rangle \Rightarrow \sum_{i=1}^{n} r_i c_i$ (by the operational semantics rule for recursion), so we see that $f_{n+1} \lesssim \widetilde{fe}'$ as required. ∎

Corollary 8.6 *For any closed, well-typed expression e,*

$$[\![e]\!] = \bigsqcup\{\sum_i r_i [\![c_i]\!] \mid e \Rightarrow \sum_i r_i c_i\}$$

Proof It has been noted that the set on the right is directed, let its lub be μ. By the theorem above, $[\![e]\!] \lesssim e$. So $[\![e]\!]$ is in the smallest closed set containing all $\sum r_i \eta(x_i)$ where $x_i \lesssim c_i$ and $e \Rightarrow \sum r_i c_i$. But consider the set of evaluations below μ, that is $\mu\!\downarrow$. This set is closed, and since $x_i \lesssim c_i$ implies $\eta(x_i) \sqsubseteq [\![c_i]\!]$ and $\sum r_i [\![c_i]\!] \sqsubseteq \mu$, it must contain all the evaluations of the form $\sum r_i \eta(x_i)$. So it must contain $[\![e]\!]$, hence $[\![e]\!] \sqsubseteq \bigsqcup\{\sum_i r_i [\![c_i]\!]\}$ hence with Theorem 8.3, $[\![e]\!] = \bigsqcup\{\sum_{i=1}^{n} r_i [\![c_i]\!]\}$. ∎

8.6 Concluding Remarks

In this chapter we have given a probabilistic metalanguage, its denotational and operation semantics and proved them equivalent. The categorical method for giving the relationship on types was based on one developed by Moggi [25], and part of it is sufficiently general to apply to any strong (and cone continuous) monad.

The metalanguage is purely call by value. It may not be very hard to add a lifting construct as in Plotkin's metalanguage, but this would only deal with one of the possible parameter passing mechanisms for probabilistic programs.

An example of using a similar probabilistic metalanguage to give the semantics of another probabilistic language is given in [18].

Chapter 9

Conclusions and Further Work

We have shown that the powerdomain of continuous evaluations forms a "good" basis for semantics of probabilistic languages. A powerdomain of measures might seem a more natural choice, because of their history as the foundation of probability theory, but using evaluations seems to avoid some of the problems in measure theory. We have used this basis in developing a probabilistic logic, equivalent to denotational semantics, for a simple while language and a denotational and operational semantics for a functional metalanguage.

We have seen that the set of continuous evaluations on the lattice of Scott open sets of an ipo forms an ipo itself. Then the obvious powerdomain functor $\mathcal{V}$ was shown to be a monadic functor on the category of ipos and also to satisfy the conditions for it to be a model of the λ_c-calculus as defined by Moggi. We then compared the $\mathcal{V}$-algebras arising from this monad to the category of abstract probabilistic domains as defined by Graham, which has a finitary equational presentation. We also gave various conditions for when certain types of measures were related in the partial order.

We then looked at the special case of when the ipo was continuous. We proved that any continuous evaluation was the lub of some directed set of linear combinations of point evaluations. We were then able to show that the continuous abstract probabilistic domains were exactly the $\mathcal{V}$-algebras.

This work has extended the work by Saheb-Djahromi and others on using Borel measures as a powerdomain. The problem with using Borel measures is that the set of measures under a suitable ordering does not in general seem to be an ipo while the set of evaluations naturally is. The theorem that every continuous evaluation is the lub of a directed set of linear combinations of point evaluations lets us conclude that measures on a continuous ipo which restrict to a continuous evaluation form an ipo, and that on second countable continuous ipos the set of all Borel measures forms an ipo.

We also defined a duality between sets of evaluations and functions, as suggested by the bilinearity and continuity of integration. Previous work in this area by Kozen [22] used a duality between measurable functions and measures, but upper continuous functions seemed more appropriate in the ipo setting. We first looked for a duality between upper continuous functions and measures, however linear and continuous functionals on the set of upper continuous functions on some ipo are naturally continuous evaluations rather than measures, which suggested using continuous evaluations from the start. In order to carry out this program it was also necessary to define integration of upper continuous functions with respect to continuous evaluations.

The duality then suggested defining a logic with terms as upper continuous functions. This idea led to the two program logics, one using upper continuous functions and the other using lower continuous functions and gave the appropriate interpretation via a denotational semantics.

Finally we defined a metalanguage, and gave it a denotational semantics and an operational semantics and proved their equivalence. The operational semantics was the last in a series of attempts to find a good solution of the problem of adding together canonical terms reached by possibly infinitely many different execution paths. The end result was a non-deterministic relation, effectively summing over any finite collection of execution paths. The other interesting part of this work was the relation between $\lesssim_\sigma$ and $\lesssim_{\mathcal{V}\sigma}$. This was rather complicated but the simpler ideas we considered proved to be insufficient.

There are several interesting unsolved problems which have arisen from this thesis. Firstly there is the question of whether products of evaluations are symmetric. We know this is the case for continuous ipos and conjecture that it is true for all ipos. In the metalanguage this corresponds to the order of evaluation of a pair of expressions being irrelevant to the final result.

Secondly is the question of whether the sets of measures on an ipo, with perhaps extra conditions imposed, form an ipo. We know that a lub of linear combinations of point measures exists, and the only proofs so far that all measures form an ipo come from results that over some class of ipos, every measure is a lub of point measures. Another problem is whether these results generalise i.e., if all measures were given by a directed lub of linear combinations of point measures, that would show that all sets of measures form an ipo. I do not think this is the case, but have been unable to find a counterexample. This is mainly due to the difficulty in constructing non-trivial measures, particularly on an ipo which is sufficiently intractable that it is not continuous.

Similar is the problem of whether all linear, continuous functions are super-linear. Linearity is a natural condition to impose on semantics, and super-linearity gives a categorical idea of linearity but is stronger (although perhaps equivalent).

Another problem is whether all evaluations on ipos can be extended to measures. If this was true then sets of measures on an ipo would always form an ipo and using measures would be equivalent to using evaluations. If it wasn't then the question would arise as to which best represented probabilistic computation. Also in Lawson's paper [23], he proves that any continuous evaluation extends to a regular Borel measure when the "patch" topology is compact; it would be interesting to see what this condition meant for ipos, and whether it corresponded to a useful subclass of ipos.

For further work, there are several directions in which this thesis may be expanded. One is to use the framework for deriving semantics of other probabilistic languages, e.g. a language with probabilistic concurrency. Perhaps one could use Moggi λ_c-calculus methods for giving semantics of languages with side-effects or exceptions.

Another problem to do with semantics is dealing with different parameter passing mechanisms. There are at least three different mechanisms, of which only pure call-by-value is dealt with here. It is not clear how they could be dealt with in the metalanguage, although perhaps call-by-name could be derived from lifting as in Plotkin's metalanguage. Here we think the appropriate model for an unevaluated computation of type σ is an element of $\mathcal{V}(1 \to \mathcal{V}[\![\sigma]\!]) = \mathcal{V}^2[\![\sigma]\!]$.

The integration with respect to evaluations could be developed further. We expect that it can be extended to perhaps all continuous (in the usual sense) functions. Then, for compact Hausdorff spaces by Riesz's theorem, this would imply that all continuous evaluations are regular measures.

Finally another obvious thing to do is to develop other probabilistic logics like in Chapter 7. The duality theorem suggests that functions are the natural choice for terms, but they are harder to understand than first-order logic terms; it would be better to represent the functions by some sort of first-order logic terms decorated with probabilities perhaps like in [11].

Appendix A

Defining a Relation

Here we will define the relation which we require to prove the equivalence of operational and denotational semantics in Chapter 8. Recall that this relation is between elements of $[\![\sigma]\!]$ and canonical expressions of type σ and it must satisfy the equations and properties given in Section 8.5.

The idea is that we find sequences with limit $[\![\sigma]\!]$ for all types σ which "uniformly" approximate the recursive types, so for a type σ with nested recursion, the nth term in the sequence is formed by taking the nth approximating ipo for the meaning of each recursive type. Furthermore these ipos have the property that $[\![\mu P.\sigma[P]]\!]^n = [\![\sigma[\mu P.\sigma]]\!]^{n-1}$ where $[\![\sigma]\!]^n$ is the nth approximating ipo for $[\![\sigma]\!]$. Then we can define a relation $\precsim^n_\sigma$ between canonical elements of type σ and elements of $[\![\sigma]\!]^n$, by induction on the structure of σ as in the conditions for $\precsim_\sigma$ except that for the recursion case, where $\precsim_{\mu P.\sigma[P]}$ is defined in terms of the relation on the more complicated type $\precsim_{\sigma[\mu P.\sigma]}$, we define $\precsim^{n+1}_{\mu P.\sigma[P]}$ in terms of $\precsim^n_{\sigma[\mu P.\sigma]}$. So the relation is defined by induction on σ and n. We can then define a relation on $[\![\sigma]\!]$ via the natural maps from $[\![\sigma]\!]^n$ to $[\![\sigma]\!]$ and show it has the desired properties by induction.

We will use the types given by:-

$$\sigma = \sigma_1 \times \ldots \times \sigma_n \mid \sigma_1 + \ldots + \sigma_n \mid \sigma \rightarrow_{\mathcal{V}} \tau \mid P \mid \mu P.\sigma[P]$$

For each type σ with free variables among $P_1, \ldots, P_n$ we will define a cone Δ_σ in the functor category of functors $(\mathbf{Ipo}^E)^n \rightarrow \mathbf{Ipo}^E$ (where by $\mathbf{C}^0$ we mean the category with one object and its identity morphism — i.e. if σ has no free variables the functor category is isomorphic to $\mathbf{Ipo}^E$). We will give the cone in terms of functors $K^1_\sigma, K^2_\sigma, \ldots, K_\sigma$ and natural transformations $\mu^n_\sigma : K^n_\sigma \rightarrow K^{n+1}_\sigma$ and $\eta^n_\sigma : K^n_\sigma \rightarrow K_\sigma$. For a σ with no free variables, note that K_σ will be $[\![\sigma]\!]$ as they will have the same construction.

For any σ, we define $K^0_\sigma(P_1, \ldots, P_n) = \emptyset$ and $K^0_\sigma(f_1, \ldots, f_n) = \mathrm{id}_\emptyset$. For any A in $\mathbf{Ipo}^E$ there is a unique embedding $\emptyset \lhd A$, this unique embedding gives us $\mu^0_\sigma(P_1, \ldots, P_n)$ and $\eta^0_\sigma(P_1, \ldots, P_n)$. Similarly, for the zero product, and $i > 0$ $K^i_{\langle\rangle}(P_1, \ldots, P_n) = \mathbb{1}$, (the ipo with one element) and $K^i_{\langle\rangle}(f_1, \ldots, f_n) = \mathrm{id}_{\mathbb{1}}$, (and similarly for $K_{\langle\rangle}$). Then we get a cone by setting $\mu^i(P_1, \ldots, P_n) = \eta^i(P_1, \ldots, P_n) = \mathrm{id}_{\mathbb{1}}$. So for $\sigma = \langle\rangle$ this defines a universal cone.

Similarly for $i > 0$, $K^i_{P_j}(P_1, \ldots, P_n) = P_j$ and $K^i_{P_j}(f_1, \ldots, f_n) = f_j$ and $\mu^i_{P_j}$ and $\eta^i_{P_j}$ are identity functions. Again this is clearly a universal cone.

For the type operators $\times, +$ and $\rightarrow_{\mathcal{V}}$ which are "cone-continuous", then for $i > 0$ we write

$$\begin{array}{rclrcl}
K^i_{\sigma_1 \times \ldots \times \sigma_n} & = & K^i_{\sigma_1} \times \ldots \times K^i_{\sigma_n} & K_{\sigma_1 \times \ldots \times \sigma_n} & = & K_{\sigma_1} \times \ldots \times K_{\sigma_n} \\
\mu^i_{\sigma_1 \times \ldots \times \sigma_n} & = & \mu^i_{\sigma_1} \times \ldots \times \mu^i_{\sigma_n} & \eta^i_{\sigma_1 \times \ldots \times \sigma_n} & = & \eta^i_{\sigma_1} \times \ldots \times \eta^i_{\sigma_n} \\
K^i_{\sigma_1 + \ldots + \sigma_n} & = & K^i_{\sigma_1} + \ldots + K^i_{\sigma_n} & K_{\sigma_1 + \ldots + \sigma_n} & = & K_{\sigma_1} + \ldots + K_{\sigma_n} \\
\mu^i_{\sigma_1 + \ldots + \sigma_n} & = & \mu^i_{\sigma_1} + \ldots + \mu^i_{\sigma_n} & \eta^i_{\sigma_1 + \ldots + \sigma_n} & = & \eta^i_{\sigma_1} + \ldots + \eta^i_{\sigma_n} \\
K^i_{\sigma \rightarrow_{\mathcal{V}} \tau} & = & K^i_\sigma \rightarrow_{\mathcal{V}} K^i_\tau & K_{\sigma \rightarrow_{\mathcal{V}} \tau} & = & K_\sigma \rightarrow_{\mathcal{V}} K_\tau \\
\mu^i_{\sigma \rightarrow_{\mathcal{V}} \tau} & = & \mu^i_\sigma \rightarrow_{\mathcal{V}} \mu^i_\tau & \eta^i_{\sigma \rightarrow_{\mathcal{V}} \tau} & = & \eta^i_\sigma \rightarrow_{\mathcal{V}} \eta^i_\tau
\end{array}$$

where we take products, sums and function space of natural transformations and functors pointwise. Clearly, by the cone-continuity of the product, sum and probabilistic function space, given universal cones Δ_{σ_i} and Δ_τ, the cones $\Delta_{\sigma_1\times\ldots\times\sigma_n}, \Delta_{\sigma_1+\ldots+\sigma_n}$ and $\Delta_{\sigma\to_{\mathcal{V}}\tau}$ are all universal cones.

Now we only need to define $\Delta_{\mu P.\sigma[P]}$ and show that it is a universal cone. We first define $K_{\mu P.\sigma[P]}(P_1,\ldots,P_n)$ to be the colimit of the diagram $\langle Q_m, f_m\rangle$ in $\mathbf{Ipo}^E$ where $Q_0=\emptyset$, $Q_{m+1}=K_\sigma(P_1,\ldots,P_n,Q_m)$, f_0 is the unique embedding $\emptyset \lhd Q_1$ and $f_{m+1}=K_\sigma(\mathrm{id}_{P_1},\ldots,\mathrm{id}_{P_n},f_m)$. Note that this is exactly how $[\![\mu P.\sigma[P]]\!]$ is defined in terms of the functor given by considering $[\![\sigma[P]]\!]$ as a function of P. Similarly we define $K_{\mu P.\sigma[P]}(f_1,\ldots,f_n)$ on morphisms, as the "colimit" of the functions $h_0{:}\,\emptyset\lhd\emptyset$ and $h_{i+1}=K_{\sigma[P]}(f_1,\ldots,f_n,h_i)$ (actually it is a mediating map constructed using the composition of the h_i and the limit morphisms to the colimit, as in Lemma A.1 below). Note that we have an isomorphism $K_{\mu P.\sigma[P]}\cong K_\sigma(K_{\mu P.\sigma[P]})$. Then we define $K^{n+1}_{\mu P.\sigma[P]}(P_1,\ldots,P_n)=K^n_\sigma(P_1,\ldots,P_n,K^n_{\mu P.\sigma[P]}(P_1,\ldots,P_n))$ and $K^{n+1}_{\mu P.\sigma[P]}(f_1,\ldots,f_n)=K^n_\sigma(f_1,\ldots,f_n,K^n_{\mu P.\sigma[P]}(f_1,\ldots,f_n))$. Then, by the naturality of μ^n_σ we can see that

$$\mu^n_\sigma(P_1,\ldots,P_n,K^{n+1}_{\mu P.\sigma[P]}(P_1,\ldots,P_n))\circ K^n_\sigma(\mathrm{id}_{P_1},\ldots,\mathrm{id}_{P_n},\mu^n_{\mu P.\sigma[P]}(P_1,\ldots,P_n))$$

$$=K^{n+1}_\sigma(\mathrm{id}_{P_1},\ldots,\mathrm{id}_{P_n},\mu^n_{\mu P.\sigma[P]}(P_1,\ldots,P_n))\circ\mu^n_\sigma(P_1,\ldots,P_n,K^n_{\mu P.\sigma[P]}(P_1,\ldots,P_n))$$

which is what we define $\mu^{n+1}_{\mu P.\sigma[P]}$ to be and similarly for η.

We will show that for closed types σ, the above functors and natural transformations (which are actually just objects and morphisms in $\mathbf{Ipo}^E$) form a cone. We will do this by defining the notion of a functor cone, which on closed types is just that the functors and natural transformations form a cone. Firstly we prove a purely categorical lemma.

Lemma A.1 *In a co-complete category $\boldsymbol{C}$, consider the diagram of an infinite square grid of objects P^i_j with i,j ranging over positive integers and morphisms $f^i_j{:}\,P^i_j\to P^{i+1}_j$ and $g^i_j{:}\,P^i_j\to P^i_{j+1}$ where each square commutes. Then the colimit of each column gives a unique sequence of objects P^i and morphisms $f^i{:}\,P^i\to P^{i+1}$ which commute with the original diagram and whose limit is the limit of the* *entire diagram, and also the limit of the diagonal subdiagram, of objects P_n^n and morphisms given by composition of appropriate f and g.*

Proof This situation is shown in the diagram

$$\begin{tikzcd}
P^0 \arrow[r, "f^0"] & P^1 \arrow[r, "f^1"] & P^2 & & P \\
\vdots & \vdots & \vdots & \iddots & \\
P_2^0 \arrow[r, "f_2^0"] & P_2^1 \arrow[r, "f_2^1"] & P_2^2 & \cdots & \\
P_1^0 \arrow[u, "g_1^0"'] \arrow[r, "f_1^0"] & P_1^1 \arrow[u, "g_1^1"'] \arrow[r, "f_1^1"] \arrow[ur, dashed] & P_1^2 \arrow[u, "g_1^2"'] & \cdots & \\
P_0^0 \arrow[u, "g_0^0"'] \arrow[r, "f_0^0"] \arrow[ur, dashed] & P_0^1 \arrow[u, "g_0^1"'] \arrow[r, "f_0^1"] & P_0^2 \arrow[u, "g_0^2"'] & \cdots &
\end{tikzcd}$$

If the limit morphisms to P^i are $\rho_j^i \colon P_j^i \to P^i$ then we derive f^i as the unique mediating morphism from the cone on the i^th column to P^{i+1} given by the morphisms $\rho_j^{i+1} \circ f_j^i$. Then the colimit of P^i, f^i, which has limit morphisms say $\rho^i \colon P^i \to P$, is the colimit of the entire diagram since we have maps $P_j^i \to P$ given by $\rho^i \circ \rho_j^i$ and clearly all the triangles commute. Furthermore if X is any object with morphisms $\sigma_j^i \colon P_j^i \to X$ such that all triangles commute, then there are unique mediating morphisms $\theta^i \colon P^i \to X$, such that σ_j^i commutes with $\theta^i \circ \rho_j^i$ and by this uniqueness σ^i commutes with $\sigma^{i+1} \circ f^i$ hence there is a unique morphism $\theta \colon P \to X$ and θ^i commutes with $\theta \circ \rho^i$ hence we see that σ_j^i commutes with $\theta \circ \rho^i \circ \rho_j^i$ as required.

Then we note that the functor $L \colon \mathbf{N} \to \mathbf{N}^2$ (where $\mathbf{N}$ is the linearly ordered set of natural numbers) given by $i \mapsto (i, i)$ is clearly final (see [36, page 213]), hence by theorem 1 on [36, page 213] the limit of the diagonal subdiagram is isomorphic to the limit of the whole diagram. ∎

We call a set of functors and natural transformations K, K^n, μ^n, η^n a *functor cone* if for any variables $P_1, \ldots, P_n$, the sequence of objects $K^n(P_1, \ldots, P_n)$ and morphisms, $\mu^n(P_1, \ldots, P_n)$ form a cone with limit $K(P_1, \ldots, P_n)$ by $\eta^n(P_1, \ldots, P_n)$. Clearly we hope that the functors associated with each type σ as defined above are functor cones.

Lemma A.2 *Given a co-cone in* $\textbf{\textit{Ipo}}^E$*, say* $Q_1, \ldots, f_n: Q_n \to Q_{n+1}$*, with limit* Q *and morphisms* $\rho_n: Q_n \to Q$ *and some type* σ *with a least one free variable* P*, then if* K, K^n, μ^n, η^n *is a functor cone, and* K *is cone continuous then the sequence* $P_n = K^n(Q_n)$ *with morphisms* $k_n: P_n \to P_{n+1}$ *given by* $\mu^n(K^n(Q_n)) \circ K^{n+1}(f_n) = K^n(f_n) \circ \mu^n(K^{n+1}(Q_{n+1}))$ *is a cone with limit* $K(Q)$ *and maps* $\nu_n: P_n \to K(Q)$ *given by* $\eta^n(K^n(Q_n)) \circ K^{n+1}(\rho_n) = K^n(\rho_n) \circ \eta^n(K^{n+1}(Q_{n+1}))$.

Proof Define $P_m^n = K^n(Q_m)$ and

$$f_m^n: P_m^n \to P_m^{n+1} = \eta^n(P_m^n)$$

$$g_m^n: P_m^n \to P_{m+1}^n = K^n(f_m)$$

then by naturality of η^n each square commutes. Clearly $P_n = P_n^n$ and the morphisms k_n are given by composition along the diagonal. So by Lemma A.1 above, the limit of the diagonal is equal to the limit of the limits of each column. But clearly along the m column, the limit must be $K(Q_m)$, since K^n is a functor cone, and the unique functor $K(Q_m) \to K(Q_{m+1})$ which commutes must be $K(f_m)$, then as K is cone continuous, the limit along this cone is $K(Q)$ as required. ∎

We now prove that $K_\sigma, K_\sigma^n, \mu_\sigma^n, \eta_\sigma^n$ form a functor cone by induction on the structure of σ. The base case where σ is a variable is trivial. Since the type operators $+$, $\times$ and $\to_\mathcal{V}$ are cone-continuous these cases are trivial. Finally we need to show that given $K_{\sigma[P]}, K_{\sigma[P]}^n, \mu_{\sigma[P]}^n, \eta_{\sigma[P]}^n$ is a functor cone, then the corresponding functors and natural transformations for $\mu P.\sigma[P]$ are also a functor cone.

Theorem A.3 *If $\sigma[P]$ gives rise to a functor cone $K_\sigma, K_\sigma^n, \mu_\sigma^n, \eta_\sigma^n$ where σ has at least one free variable P, then $\mu P.\sigma[P]$ also gives rise to a functor cone.*

Proof Pick any $P_1, \ldots, P_n$ from $\mathbf{Ipo}^E$ where $\mu P.\sigma[P]$ has free variables $P_1, \ldots, P_n$. We will always consider the cones relative to these variables but drop them for clarity.

We define a square grid as follows.

$$P_0^n = P_m^0 = \emptyset$$

$$P_{m+1}^{n+1} = K_\sigma^n(P_m^n)$$

and embeddings $f_m^n : P_m^n \to P_m^{n+1}$ and $g_m^n : P_m^n \to P_{m+1}^n$. For n or m equal to zero we define f and g to be the unique embeddings $\emptyset \lhd A$ for the appropriate ipo A. Otherwise inductively

$$f_{m+1}^{n+1} = \mu_\sigma^n(P_m^{n+1}) \circ K_\sigma^n(f_m^n) = K_\sigma^{n+1}(f_m^n) \circ \mu_\sigma^n(P_m^n)$$

(by naturality of μ_σ^n) and

$$g_{m+1}^{n+1} = K_\sigma^n(g_m^n).$$

All the edge squares commute by the uniqueness of embeddings $\emptyset \lhd A$, and an easy calculation shows that the square on P_{m+1}^{n+1} commutes if the square on P_m^n does, hence all the squares commute.

The zeroth column (for varying n) is the trivial cone $\emptyset \lhd \emptyset \ldots$ with limit $\emptyset$. It is clear that the $(m+1)$th column can be obtained by applying the functor K_σ^n to the nth element of the mth column and then shifting the column up by one and putting $\emptyset$ at the bottom. Also the functions f_m^n are the same as the diagonal morphisms in Lemma A.2. So by this lemma we see that if P_m is the limit of the mth column, then $P_0 = \emptyset$ and $P_{m+1} = K_\sigma(P_m)$. Also, the commuting embeddings from P_m to P_{m+1} are as in the definition of $K_{\mu P.\sigma[P]}$. So we see that the limit of this square is $K_{\mu P.\sigma[P]}$. Then by Lemma A.1, we know that the limit along the diagonal of this diagram is isomorphic to $K_{\mu P.\sigma[P]}$. But the diagonal satisfies the relation, $P_0^0 = \emptyset$ and $P_{n+1}^{n+1} = K_\sigma^n(P_n^n)$ so $P_n^n = K_{\mu P.\sigma[P]}^n$ and the diagonal maps are equal to $\mu_{\mu P.\sigma[P]}^n$ and also the limit morphisms are $\eta_{\mu P.\sigma[P]}^n$. ∎

Now since for every type we have a functor cone $K_\sigma, K_\sigma^n, \mu_\sigma^n, \eta_\sigma^n$ then for closed types, $\langle K_\sigma^n, \mu_\sigma^n \rangle$ is a cone with limit K_σ and limit morphisms η_σ^n. So can we proceed to define our relation.

We shall denote the embedding part of μ_σ^n by f_σ^n which is a continuous function $K_\sigma^n \to K_\sigma^{n+1}$ and the retract by $g_\sigma^n\colon K_\sigma^{n+1} \to K_\sigma^n$ and similarly denote the parts of η_σ^n by then functions $\varphi_\sigma^n\colon K_\sigma^n \to K_\sigma$ and $\psi_\sigma^n\colon K_\sigma \to K_\sigma^n$.

We will need to use the operational semantics given in Chapter 8, and the denotational semantics. Recall that the semantics of a canonical form $[\![c]\!]$ for c of closed type σ, is a point measure in $\mathcal{V}[\![\sigma]\!] = \mathcal{V}K_\sigma$. We will let $(\!(c)\!)$ denote this point, i.e. $(\!(c)\!) \in [\![\sigma]\!]$. Note that then $(\!(\lambda fe)\!) = [\![fe]\!]$, since $[\![\lambda fe]\!] = \eta([\![fe]\!])$.

Now we will define a relation $\precsim_\sigma^n$ between K_σ^n and closed canonical expressions of type σ as follows. For $n = 0$, we take any relation since one side is the empty set. Then, by induction on n and structure of σ we define

- $\langle x_1, \ldots, x_n \rangle \precsim_{\sigma_1 \times \ldots \times \sigma_n}^n c_1 \times \ldots \times c_n$ iff for all i, $x_i \precsim_\sigma^n c_i$
- $\mathrm{in}_i(x) \precsim_{\sigma_1 + \ldots + \sigma_n}^n \mathrm{in}_j(c)$ iff $i = j$ and $x \precsim_{\sigma_i}^n c$
- $x \precsim_{\mu P.\sigma[P]}^n \mathrm{intro}(c)$ iff $x \precsim_{\sigma[\mu P.\sigma]}^{n-1} c$
- $f \precsim_{\sigma \to_\mathcal{V} \tau}^n \lambda fe$ iff $f \sqsubseteq \psi_{\sigma \to_\mathcal{V} \tau}^n(\!(fe)\!)$ and $x \precsim_\sigma^n c$ implies $f(x)$ is in the least Scott closed set containing all $\mu = \sum_i r_i \eta(x_i)$ where $x_i \precsim_\tau^n c_i$ and $\langle fe, c \rangle \Rightarrow \sum_i r_i c_i$

We will now give three properties of this relation, firstly that $x \precsim_\sigma^n c$ implies x is smaller that the semantics of c, secondly that $x \precsim_\sigma^n c$ implies that $f_\sigma^n(x) \precsim_\sigma^{n+1} c$, thirdly, that the relation is closed below. All these properties hold trivially for $n = 0$, so we can prove them by induction on n and the structure of σ.

Lemma A.4 *With the relations defined above, for any type σ and integer n,*

1. $x \precsim^n_\sigma c$ *implies* $x \sqsubseteq\!\!\sim \psi^n_\sigma(\!(c)\!)$.

2. $x \precsim^n_\sigma c$ *iff* $f^n_\sigma(x) \precsim^{n+1}_\sigma c$.

3. *The set of x s.t. $x \precsim^n_\sigma c$ is Scott closed.*

Proof We use induction on n and the structure of σ.

For the first, we note the base case $n = 0$ is trivial. We then consider $n = k+1$ and use induction on σ. For products we simply note that if $x_i \precsim^n_\sigma c_i$ for $i = 1, \ldots n$, so inductively, $(x_1, \ldots, x_n) \sqsubseteq\!\!\sim (\psi^n_{\sigma_1}(\!(c_1)\!), \ldots \psi^n_{\sigma_n}(\!(c_n)\!)) = \psi^n_{\sigma_1 \times \ldots \times \sigma_n}(\!(c_1 \times \ldots \times c_n)\!)$ and similarly for sums. For function space it is part of the premise and finally for recursion it is trivial.

For the second, the base case $n = 0$ is (as always) trivial. For $n = k+1$, we use induction on the structure of σ. $(x_1, \ldots, x_n) \precsim^{k+1}_{\sigma_1 \times \ldots \times \sigma_n} c_1 \times \ldots \times c_n$ iff $x_i \precsim^{k+1}_{\sigma_i} c_i$ and by inductive hypothesis, this is iff $f^{k+1}(x_i) \precsim^{k+2}_{\sigma_i} c_i$, hence since $f^{k+1}_{\sigma_1 \times \ldots \times \sigma_n}(x_1, \ldots, x_n) = f^{k+1}_{\sigma_1}(x_1) \times \ldots \times f^{k+1}_{\sigma_n}(x_n)$ we're done. Similarly for sums, since $f^{k+1}_{\sigma_1 + \ldots + \sigma_n}(\mathrm{in}_i(x)) = \mathrm{in}_i(f^{k+1}_{\sigma_i}(x))$. For function space, the order condition is trivial from the commuting condition of φ, ψ, f and g. We suppose that $f \precsim^n_{\sigma \to_{\mathcal{V}} \tau} \lambda fe$ and try to prove $f^n_{\sigma \to_{\mathcal{V}} \tau}(f) \precsim^{n+1}_{\sigma \to_{\mathcal{V}} \tau} \lambda fe$. Given any $x \precsim^{n+1}_\sigma c$ then $f^n_{\sigma \to_{\mathcal{V}} \tau}(f)(x) = \mathcal{V} f^n_\tau(f(g^n_\sigma(x)))$. If $g^n_\sigma(x)$ does not exist then $f^n_{\sigma \to_{\mathcal{V}} \tau}(f)(x)$ is the zero measure, hence it is trivial, if it does exist, say it equals y, then $f^n_\sigma(y) = x$ so $y \precsim^n_\sigma c$, then we have $f(y)$ is in the least closed set containing all $\sum_i r_i \eta(x_i)$ etc., but by applying Lemma 8.4 with the function $\mathcal{V} f^n_\tau$, we only need to check that $\mathcal{V} f^n_\tau(\sum_i r_i \eta(x_i))$ is in the closed set, but $x_i \precsim^n_\tau c_i$ implies that $f^n_\tau(x_i) \precsim^{n+1}_\tau c_i$ so we can see that $f^n_{\sigma \to_{\mathcal{V}} \tau}(f) \precsim^{n+1}_{\sigma \to_{\mathcal{V}} \tau} \lambda fe$ as required. For the converse we need to show that if $f^n_{\sigma \to_{\mathcal{V}} \tau}(f) \precsim^{n+1}_{\sigma \to_{\mathcal{V}} \tau} \lambda fe$, then $f \precsim^n_{\sigma \to_{\mathcal{V}} \tau} \lambda fe$. Suppose that $x \precsim^n_\sigma c$, then by the induction hypothesis $f^n_\sigma(x) \precsim^{n+1}_\sigma c$ so we know that $f^n_{\sigma \to_{\mathcal{V}} \tau}(f)(f^n_\sigma(x)) = \mathcal{V} f^n_\tau(f(x))$ is in the least closed set containing all $\sum_i r_i \eta(x_i)$ where $x_i \precsim^{n+1}_\tau c_i$ and $\langle fe, c \rangle \Rightarrow$

$\sum_i r_i c_i$. But as f and g form a projection pair, also $f(x) = \mathcal{V} g^n_\tau(\mathcal{V} f^n_\tau(f(x)))$, so by Lemma 8.4, it is sufficient to show that $\mathcal{V} g^n_\tau(\sum_i r_i \eta(x_i))$ is in this set whenever $x_i \precsim^{n+1}_\tau c_i$ and $\langle fe, c\rangle \Rightarrow \sum_i r_i c_i$. But $\mathcal{V} g^n_\tau(\sum_i r_i \eta(x_i)) = \sum_i r_i \eta(g^n_\tau(x_i))$ and by the inductive hypothesis $g^n_\tau(x_i) \precsim^{n+1}_\tau c_i$ so clearly $\mathcal{V} g^n_\tau(\sum_i r_i \eta(x_i))$ is is the set as required. Finally for recursive types this is simple since $x \precsim^n_{\mu P.\sigma} \text{intro}(c)$ iff $x \precsim^{n-1}_{\sigma[\mu p.\sigma]} c$ iff $f^{n-1}_{\sigma[\mu P.\sigma]}(x) \precsim^n_{\sigma[\mu p.\sigma]} c$ iff $f^n_{\mu P.\sigma[P]}(x) \precsim^{n+1}_{\mu P.\sigma[P]} \text{intro}(c)$ as required.

For the final part we similarly use induction on n, with trivial base case and then work by structural induction on types. For product and sum types it is obvious, since the product of closed sets is closed and so is the injection. For function spaces, we first note that the set of $f \sqsubseteq \psi^n_{\sigma \to_\mathcal{V} \tau}(\!(fe)\!)$ is Scott closed, and by definition the set of f satisfying the second condition is also closed, therefore the set of $f \precsim^n_{\sigma \to_\mathcal{V} \tau} \lambda fe$ which is the intersection of these sets is also closed. For recursive types, this is again trivial. ∎

We now give our final relationship by $x \precsim_\sigma c$ iff for all n s.t. $\psi^n_\sigma(x)$ exists, then $\psi^n_\sigma(x) \precsim^n_\sigma c$. Note that since ψ^n_σ is part of the limit morphisms of a cone, then for all x, $\psi^n_\sigma(x)$ exists for sufficiently large n.

Theorem A.5 *The relation given above satisfies the following.*

- $(x_1, \ldots, x_n) \precsim_{\sigma_1 \times \ldots \times \sigma_n} c_1 \times \ldots \times c_n$ *iff* $x_i \precsim_\sigma c_i$ *for* $i = 1, \ldots, n$
- $in_i(x) \precsim_{\sigma_1 + \ldots + \sigma_n} in_j(c)$ *iff* $i = j$ *and* $x \precsim_{\sigma_i} c$
- $x \precsim_{\mu P.\sigma[P]} intro(c)$ *iff* $\theta^{-1}(x) \precsim_{\sigma[\mu P.\sigma]} c$
- $f \precsim_{\sigma \to_\mathcal{V} \tau} \lambda fe$ *iff* $f \sqsubseteq (\!(\lambda fe)\!)$ *and* $x \precsim_\sigma c$ *implies* $f(x)$ *is in the least closed set containing all evaluations of the form* $\sum_i r_i \eta(x_i)$ *where* $x_i \precsim_\tau c_i$ *and* $\langle fe, c\rangle \Rightarrow \sum_i r_i c_i$

Proof For the product condition we know that $(x_1, \ldots, x_n) \precsim^n_{\sigma_1 \times \ldots \times \sigma_n} c_1 \times \ldots \times c_n$ iff $x_i \precsim^n_\sigma c_i$ for $i = 1, \ldots, n$. So $(x_1, \ldots, x_n) \precsim_{\sigma_1 \times \ldots \times \sigma_n} c_1 \times \ldots \times c_n$ iff for all

n, $\psi^n_{\sigma_1\times\ldots\times\sigma_n}(x_1,\ldots,x_n) \precsim^n_{\sigma_1\times\ldots\times\sigma_n} c_1 \times \ldots \times c_n$ but since $\psi^n_{\sigma_1\times\ldots\times\sigma_n}(x_1,\ldots,x_n) = \psi^n_{\sigma}(x_1) \times \ldots \times \psi^n_{\sigma_n}(x_n)$ this is iff $\psi^n_\sigma(x_i) \precsim^n_\sigma c_i$ for $i = 1,\ldots,n$ i.e. iff $x_i \precsim_\sigma c_i$ as required.

For sums, $\mathrm{in}_j(x) \precsim_{\sigma_1+\ldots+\sigma_n} \mathrm{in}_j(c)$ iff for all n, $\psi^n_{\sigma_1+\ldots+\sigma_n}(\mathrm{in}_j(x)) \precsim^n_{\sigma_1+\ldots+\sigma_n} \mathrm{in}_j(c)$, but $\psi^n_{\sigma_1+\ldots+\sigma_n}(\mathrm{in}_j(x)) = \mathrm{in}_j(\psi^n_{\sigma_j}(x))$ and $\mathrm{in}_j(y) \precsim^n_{\sigma_1+\ldots+\sigma_n} \mathrm{in}_i(c)$ iff $i = j$ and $y \precsim_{\sigma_i} c$, hence $i = j$ and $\psi^n_{\sigma_j}(x) \precsim^n_{\sigma_i} c$ so $x \precsim_{\sigma_i} c$ as required.

For recursion, $x \precsim_{\mu P.\sigma[P]} \mathrm{intro}(c)$ iff for all n, $\psi^n_{\mu P.\sigma[P]}(x) \precsim^n_{\mu P.\sigma[P]} \mathrm{intro}(c)$, iff $\psi^n_{\mu P.\sigma[P]}(x) \precsim^{n-1}_{\sigma[\mu P.\sigma]} c$ but $\psi^n_{\mu P.\sigma[P]}(x) = \psi^{n-1}_{\sigma[\mu P.\sigma]}(\theta^{-1}(x))$ so for all n, we have that $\psi^n_{\sigma[\mu P.\sigma]}(\theta^{-1}(x)) \precsim_{\sigma[\mu P.\sigma]}{}^n c$ i.e. $\theta^{-1}(x) \precsim_{\sigma[\mu P.\sigma]} c$ as required.

Finally for functions we first suppose that $f \precsim_{\sigma\to_{\mathcal{V}}\tau} \lambda fe$ and that $x \precsim_\sigma c$. Then $\psi^n(f) \precsim^n_{\sigma\to_{\mathcal{V}}\tau} \lambda fe$ implies $\psi^n(f) \sqsubseteq \psi^n(\llparenthesis\lambda fe\rrparenthesis)$ and so as ψ,φ form a cone, $f \sqsubseteq \llparenthesis\lambda fe\rrparenthesis$. Further $\psi^n_\sigma(x) \precsim^n_\sigma c$ (over the n for which it exists) and $\psi^n_{\sigma\to_{\mathcal{V}}\tau}(f) \precsim^n_{\sigma\to_{\mathcal{V}}\tau} \lambda fe$ for all n. So $\psi^n_{\sigma\to_{\mathcal{V}}\tau}(f)(\psi^n_\sigma(x)) = \mathcal{V}\psi^n_\tau(f(x))$ is in the least closed set generated by the points $\sum_i r_i\eta(x_i)$ where $x_i \precsim^n_\tau c_i$ and $\langle fe, c\rangle \Rightarrow \sum_i r_i c_i$. We need to prove that $f(x)$ is in the least closed set containing all $\sum_i r_i\eta(x_i)$ where $x_i \precsim_\tau c_i$. Since this set is closed and ψ,φ form a cone, it is sufficient to show that $\mathcal{V}\varphi^n_\tau(\mathcal{V}\psi^n_\tau(f(x)))$ is in this closed set for all n. But the inverse image of this closed set under the function $\mathcal{V}\psi^n_\tau$ is closed and must contain all the necessary points since $x_i \precsim^n_\tau c_i$ implies $\varphi^n_\tau(x_i) \precsim_\tau c_i$ (by the lemma on f^n_τ), so we are done. To show the converse we want to prove that if the condition above is satisfied, then for any n, $\psi^n_{\sigma\to_{\mathcal{V}}\tau}(f) \precsim^n_{\sigma\to_{\mathcal{V}}\tau} \lambda fe$, (which always exists). First, we note that $f \sqsubseteq \llparenthesis\lambda fe\rrparenthesis$, so as ψ is continuous, $\psi^n_{\sigma\to_{\mathcal{V}}\tau}(f) \sqsubseteq \psi^n_{\sigma\to_{\mathcal{V}}\tau}(\llparenthesis\lambda fe\rrparenthesis)$, this gives the first condition for $\precsim^n_{\sigma\to_{\mathcal{V}}\tau}$. For the second we suppose $x \precsim^n_\sigma c$, then $\varphi^n_\sigma(x) \precsim_\sigma c$ so by the condition above we know that $f(\varphi^n_\sigma(x))$ is in the least closed set containing all $\sum_i r_i\eta(x_i)$ with $x_i \precsim_\tau c_i$. and $\langle fe, c\rangle \Rightarrow \sum_i r_i c_i$. We need to prove that $\psi^n_{\sigma\to_{\mathcal{V}}\tau}(f)(x)$ is in the least closed set containing all $\sum_i r_i\eta(y_i)$ with $y_i \precsim^n_\tau c_i$. But we know $\psi^n_{\sigma\to_{\mathcal{V}}\tau}(f)(x) = \mathcal{V}\psi^n_\tau(f(\varphi^n_\sigma(x)))$ and by considering the inverse image under $\mathcal{V}\psi^n_\tau$ of the second closed set, we just need to show that $\mathcal{V}\psi^n_\tau(\sum_i r_i\eta(x_i))$ is in the second closed set, but this is trivial since $x_i \precsim_\tau c_i$ implies $\psi^n_\tau(x_i) \precsim^n_\tau c_i$. ∎

From the Lemma A.4 above, we can show that our final relation has the two other desired properties. Firstly the set of $x \precsim_\sigma c$ is Scott closed, since if we set $C = \{x \mid x \precsim_\sigma c\}$ and $C_n = \{x \mid x \precsim^n_\sigma c\}$, then it is easy to see that $C = \bigcap_n (\psi^n_\sigma)^{-1}(C_n)$, hence as C_n is closed, so is C. Secondly, $x \precsim_\sigma c$ implies $x \sqsubseteq (\!(c)\!)$ since if we take $x \precsim_\sigma c$, then $\psi^n_\sigma(x) \precsim^n_\sigma c$ so $\psi^n_\sigma(x) \sqsubseteq \psi^n_\sigma(\!(c)\!)$ and so applying φ^n_σ we see for all n, $\varphi^n_\sigma(\psi^n_\sigma(x)) \sqsubseteq \varphi^n_\sigma(\psi^n_\sigma(\!(c)\!))$, so $\bigsqcup_n \varphi^n_\sigma(\psi^n_\sigma(x)) = x \sqsubseteq \bigsqcup_n \varphi^n_\sigma(\psi^n_\sigma(\!(c)\!)) = (\!(c)\!)$.

We finally get the relationship we want for Chapter 8 by taking the relationship described above for expressions and then a relation on function expressions given by $f \precsim_\varphi fe$, where $\varphi = \sigma \rightarrow_\mathcal{V} \tau$, iff for all $x \precsim_\sigma c$, $f(x)$ is in the least closed set containing all $\sum_i r_i \eta(x_i)$ with $x_i \precsim_\tau c_i$ and $\langle fe, c\rangle \Rightarrow \sum_i r_i c_i$. It is clear that this relation satisfies the required conditions given in Chapter 8, and has the properties required.

Bibliography

[1] S. Abramsky. *Domain Theory and the Logic of Observable Properties.* PhD thesis, Queen Mary College, University of London, 1987.

[2] D. Angluin and L. G. Valiant. Fast probabilistic algorithms for hamiltonian circuits and matchings. *J. Computing System Science*, 18(2):155–193, 1979.

[3] K. R. Apt. Ten years of hoare's logic: A survey — part 1. *ACM Transactions on Programming Languages and Systems*, 3(4):431–483, October 1981.

[4] E. R. Berlekamp. Factoring polynomials over large finite fields. *Mathematics of Computing*, 24(111):713–735, 1970.

[5] G. Birkhoff. *Lattice Theory.* Amer. Math. Soc. Colloq. Publ., 3rd edition, 1967.

[6] B. Bollobas. Unpublished lecture notes on measure theory. Department of Pure Mathematics and Mathematical Statistics, Cambridge University.

[7] B. Bollobas. *Graph Theory, an Introductory Course.* Springer-Verlag, 1979.

[8] K. de Leeuw, E. F. Moore, C. E. Shannon, and N. Shapiro. Computability by probabilistic machines. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, pages 183–212. Princeton University Press, 1956. Annals of Mathematics Studies 34.

[9] D. Frutos Escrig. Some probabilistic powerdomains in the category SFP. In *STACS 86, Lecture Notes in Computer Science 210*, pages 49–59. Springer-Verlag, 1986.

[10] R. Fagin, J. Y. Halpern, and N. Megiddo. A logic for reasoning about probabilities. In *Proceeding of the IEEE Third Annual Symposium on Logic in Computer Science*, pages 410–421. Computer Society Press, July 1988.

[11] Y. Feldman. A decidable propositional probabilistic dynamic logic with explicit probabilities. *Information and Control*, 63:11–38, 1984.

[12] L. R. Ford and D. R. Fulkerson. Maximal flow through a network. *Canad. J. Math*, 8:339–404, 1959.

[13] J. T. Gill. Computational complexity of probabilistic turing machines. *SIAM Journal of Computing*, 6:675–695, 1977.

[14] S. K. Graham. Closure properties of a probabilistic domain construction. In M. Main, A. Melton, et al., editors, *Mathematical Foundations of Programming Language Semantics, Lecture Notes in Computer Science 298*, pages 213–233. Springer-Verlag, 1987.

[15] C. A. Gunter and D. S. Scott. *Handbook of Theoretical Computer Science*, chapter Semantic Domains. North-Holland, To Appear.

[16] P. R. Halmos. *Measure Theory*. Van Nostrand, 1965.

[17] P. T. Johnstone. *Stone Spaces*, volume 3 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1982.

[18] C. Jones and G. D. Plotkin. A probabilistic powerdomain of evaluations. In *Proceeding of the IEEE Fourth Annual Symposium on Logic in Computer Science*, pages 186–195. Computer Society Press, 1989.

[19] T. Kamimura and A. Tang. Retracts of SFP objects — projective limits, solving domain equations. In *Workshop on Mathematical Foundations of Programming Semantics, Manhatten, Kansas*, 1986.

[20] R. M. Karp. The probabilistic analysis of some combinatorial search algorithms. In J. F. Traub, editor, *Algorithms and Complexity*, pages 1–20. Academic Press, 1976.

[21] D. Kozen. Semantics of probabilistic programs. *Journal of Computer and System Sciences*, 22:328–350, 1981.

[22] D. Kozen. A probabilistic PDL. In *15th ACM Symposium on Theory of Computing*, 1983.

[23] J. D. Lawson. Valuations on continuous lattices. In Rudolf-Eberhard Hoffman, editor, *Continuous Lattices and Related Topics*, pages 204–225. Universitat Bremen, 1982. Mathematik Arbeitspapiere 27.

[24] E. G. Manes. *Algebraic Theories*, volume 26 of *Graduate Texts in Mathematics*. Springer-Verlag, 1976.

[25] E. Moggi. The relationship between denotational and operational semantics. Part of PG Theory Course, Edinburgh, 1987.

[26] E. Moggi. Computational lambda-calculus and monads. In *Proceeding of the IEEE Fourth Annual Symposium on Logic in Computer Science*, pages 14–23. Computer Society Press, 1989.

[27] N. J. Nilsson. Probabilistic logic. *Artificial Intelligence*, 28:71–87, 1986.

[28] B. J. Pettis. On the extension of measures. *Ann. of Math.*, 54:186–197, 1951.

[29] G. D. Plotkin. Domains. Lecture Notes for PG course, Edinburgh University, 1980.

[30] G. D. Plotkin. Probabilistic powerdomains. In *Proceedings CAAP*, 1982.

[31] G. D. Plotkin. Notes on metalanguage. Stanford, 1985.

[32] M. O. Rabin. Probabilistic algorithms. In *Algorithms and Complexity*, pages 21–40. Academic Press, 1976.

[33] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 1974.

[34] N. Saheb-Djahromi. Probabilistic LCF. In *Mathematical Foundations of Computer Science, 64*. Springer, 1978.

[35] N. Saheb-Djahromi. CPO's of measures for non-determinism. *Theoretical Computer Science*, 12(1):19–37, 1980.

[36] S.MacLane. *Categories for the Working Mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer-Verlag, 1971.

[37] M. B. Smyth. Powerdomains and predicate transformers: a topological veiw. In J. Diaz, editor, *Proceedings ICALP 1983*, volume 154 of *Lecture Notes in Computer Science*, pages 662–675. Springer-Verlag, 1983.

[38] M. B. Smyth. Finite approximation of spaces. In *Catgeory Theory and Computer Programming*, volume 260 of *Lecture Notes in Computer Science*, pages 225–241. Springer-Verlag, 1988.

[39] M. B. Smyth and G. D. Plotkin. The category-theoretic solution of recursive domain equation solutions. *SIAM Journal of Computing*, 11(4):761–783, November 1982.

[40] R. Solovay and V. Strassen. A fast monte-carlo test for primality. *SIAM Journal of Computing*, 6:84–85, 1977.

[41] D. J. A. Welsh. Randomised algorithms. *Discrete Applied Mathematics*, 5:133–145, 1983.

[42] R. J. Wilson. *An Introducton to Graph Theory*. Longman, 1972.

[43] S. Yamada. A mathematical theory of semantic domains of randomized algorithms. Waseda University, Tokyo, 1987.

Index of Notation

Index of Definitions

Index