A Minkowski type duality mediating between state and predicate transformer semantics for a probabilistic nondeterministic language

K. Keimel, A. Rosenbusch, T. Streicher

Fachbereich Mathematik, Technische Universität
Schloßgartenstr. 7, D-64289 Darmstadt, Germany

1 Introduction

We consider a basic imperative programming language $\mathcal{L}_p$ whose syntax is given (in BNF-form) by

$$P ::= a \mid P; P \mid \mathbf{cond}(b, P, P) \mid \mathbf{while}(b, P) \mid P_p\oplus P \mid P[\!]P$$

where b ranges over a set BExp of *boolean expressions*, a ranges over a set Act of *basic actions* and p is a real number with $0 < p < 1$. We write $\mathbf{cond}(b, P, Q)$ for the *conditional* usually denoted as **if** b **then** P **else** Q **fi** and $\mathbf{while}(b, P)$ for the *while loop* usually denoted as **while** b **do** P **od** The program $P[\!]Q$ *nondeterministically* executes either P or Q. The program $P_p\oplus Q$ executes P with probability p and Q with probability $1-p$.

A variant of this language has been considered by A. McIver and C. Morgan in [MMa,MMb,MM] together with a *state transformer* and a *predicate transformer* semantics associating with every program P a state transformer $[\![P]\!] : S \to \mathcal{P}_U\mathcal{V}(S)$ and a predicate transformer $\mathsf{wp}(P) : \mathbb{I}^S \to \mathbb{I}^S$, respectively, where $\mathbb{I}$ is the unit interval $[0, 1]$, S is a set of states and $\mathcal{P}_U$ is a kind of powerdomain suitable for total correctness over the space $\mathcal{V}(S)$ of subprobability distributions over S. Moreover, these two semantics are related by a function $\mathsf{Wp} : \Big[S \to \mathcal{P}_U\mathcal{V}(S)\Big] \to \Big[\mathbb{I}^S \to \mathbb{I}^S\Big]$ such that $\mathsf{wp}(P) = \mathsf{Wp}\Big([\![P]\!]\Big)$ for all programs P.

In their work, McIver and Morgan considered mainly finite, occasionally countably infinite, sets S of states and the result sketched in the previous paragraph cannot be found, stated and proved completely in one single paper but is scattered over various of their publications. Besides treating the problem over arbitrary infinite sets of states, we consider as the main achievement of this

paper the identification of a *Minkowksi type duality* between $\mathcal{P}_U\mathcal{V}(S)$ and a class $\mathcal{G}(S)$ of "good" functionals $G : \mathbb{I}^S \to \mathbb{I}$ where $Q \in \mathcal{P}_U\mathcal{V}(S)$ is represented by its "Minkowski functional"

$$\Phi_Q : \mathbb{I}^S \to \mathbb{I} : \gamma \mapsto \min_{\mu \in Q} \langle \mu, \gamma \rangle$$

and from such a functional G the set Q may be reconstructed as

$$\Psi_G = \{\mu \in \mathcal{V}(S) \mid \forall \gamma \in \mathbb{I}^S.\ G(\gamma) \leq \langle \mu, \gamma \rangle\}$$

where $\langle \mu, \gamma \rangle$ stands for $\sum_{s \in S} \mu(s)\gamma(s)$, the integral of γ w.r.t. μ.

We show that $\Phi : \mathcal{P}_U\mathcal{V}(S) \to \mathcal{G}(S)$ is a bijection preserving all relevant structure and exploit this fact for showing that $\mathsf{wp}(P) = \mathsf{Wp}([\![P]\!])$. Thus *Minkowski duality* is the mathematical principle from which the correspondence between state and predicate transformer semantics follows for nondeterministic probabilistic languages like $\mathcal{L}_p$.

The Minkowski duality has been worked out in [KP] for the general framework of dcpo-cones. As $\mathcal{V}(S)$ is not a cone but only a kind of a truncated cone, we cannot apply those results directly.

We slightly deviate from McIver and Morgan's approach in restricting values of predicates to $\mathbb{I}$ instead of $\mathbb{R}_+$ for systematic reasons and in a different, but equivalent, formulation of the healthiness conditions adapting their terminology to a well-established one in mathematics.

In state transformer semantics, a crucial point is the correct definition of the semantics of the composition of programs. McIver and Morgan are guessing the correct formula from a good intuition. We derive this formula in a natural way with the help of the Minkowski duality.

2 Preliminaries

We denote by $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{I}$ the reals, the nonnegative reals, and the unit interval $[0, 1]$, respectively, endowed with their usual topology and linear order.

For an arbitrary set S, the spaces $\mathbb{R}^S$, $\mathbb{R}_+^S$ and $\mathbb{I}^S$ of all functions γ from S into $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{I}$, respectively, are endowed with the pointwise defined order

$$\gamma \leq \beta \iff \gamma(s) \leq \beta(s) \text{ for all } s \in S$$

and the topology of pointwise convergence, also called the product topology. The sets

$$U_{r,s} = \{\gamma \mid \gamma(s) > r\} \text{ and } L_{t,s} = \{\gamma \mid \gamma(s) < t\}$$

where s ranges over S and r, t over real numbers, form a subbasis for the open sets of the product topology. Note that $\mathbb{I}^S$ is a compact space by Tychonoff's theorem. With respect to the order, we have a pointwise defined meet operation in all of our three function spaces

$$(\gamma \wedge \beta)(s) = \min(\gamma(s), \beta(s))$$

Also, $\mathbb{R}^S$ is a real vector space for pointwise defined addition and scalar multiplication and $\mathbb{I}^S$ is a convex subset. More generally, a *subconvex combination* of γ and β is an element of the form $r\gamma + t\beta$ with $r, t \in \mathbb{I}$ and $r + t \leq 1$. If $t = 1 - r$, then we have a *convex combination*. The space $\mathbb{I}^S$ is closed under subconvex and, in particular, under convex combinations. For $A \subseteq \mathbb{I}^S$, we write $\mathsf{conv}(A)$ for its *convex hull*, the smallest convex set containing A, which can be obtained from A by closing up under convex combinations.

We need two basic concepts from domain theory. (For an extensive treatment of domain theory one may consult [GHK$^+$].)

A *bounded directed complete partially ordered set* (a *bdcpo*, for short) is a partially ordered set L in which every directed family $(d_i)_i$, which has an upper bound, has a least upper bound $\sup_i d_i$. If every directed family in L has a least upper bound, then L is called *directed complete* (or a *dcpo*, for short). We also suppose that our dcpos always have a smallest element. An *upper set* in a (b)dcpo is a subset A with the property that $x \geq a \in A$ implies $x \in A$. *Lower sets* are defined dually. Upper subsets are also called *saturated*; for any subset A, its *saturation* is

$$\uparrow A = \{b \mid a \leq b \text{ for some } a \in A\}$$

A map f from a (b)dcpo L to another (b)dcpo M is said to be *Scott-continuous* if it preserves the order (i.e., $a \leq b \implies f(a) \leq f(b)$) and suprema of (bounded) directed sets (i.e., $f(\sup_i d_i) = \sup_i f(d_i)$ for every (bounded) directed family $(d_i)_i$ in L). The set $[L \to M]$ of all Scott-continuous maps from L to M with the pointwise defined order is again a (b)dcpo with directed suprema being defined pointwise.

$\mathbb{R}$ and $\mathbb{R}_+$ and the function spaces $\mathbb{R}^S$ and $\mathbb{R}_+^S$ are examples of bdcpos, and $\mathbb{I}$ and $\mathbb{I}^S$ are dcpos. Addition $(\gamma, \beta) \mapsto \gamma + \beta$ and the meet operation $(\gamma, \beta) \mapsto \gamma \wedge \beta$ is continuous as well as Scott-continuous on $\mathbb{R}$ and $\mathbb{R}^S$; scalar multiplication $(r, \gamma) \mapsto r \cdot \gamma$ is continuous, but Scott-continuous only if we restrict to $r \geq 0$ and $\gamma \geq 0$. It follows that subconvex combinations $(r, t, \gamma, \beta) \mapsto r\gamma + t\beta$ depend continuously on all of their arguments simultaneously, and that Scott continuity is guaranteed, when $\gamma, \beta \geq 0$.

Let us stress that topological notions, like *closed set*, *continuous function* al-

ways refer to the Hausdorff topologies considered at the beginning of these preliminaries, whilst the term *Scott-continuous* refers to the order theoretical notion of preservation of directed suprema.

Notice, moreover, that dcpos form a cartesian closed category (with exponential objects $[L \to M]$ as described above) and thus provides a model for typed λ-calculus (see e.g. [Plo,Str]). This has the consequence that every λ-definable function is automatically Scott-continuous. This fact will be used later on in a crucial way for simplifying arguments. Occasionally we will informally use the notation of λ-calculus, where $\lambda x.E(x)$ stands for $x \mapsto E(x)$.

3 State transformer semantics for $\mathcal{L}_p$

Let S be some unspecified (countable) set of states. Basic actions are interpreted as (and identified with) certain functions $a\colon S \to S$. We now have to build our powerdomain.

The set $\mathcal{V}(S)$ of *subprobability distributions* on S consists of all $\mu : S \to \mathbb{I}$ with $\sum_{s\in S} \mu(s) \leq 1$. We may put $\mu(\bot) = 1 - \sum_{s\in S} \mu(s)$ giving rise to a *probability measure* μ on $S_\bot = S \cup \{\bot\}$ with $\mu(A) = \sum_{s\in A} \mu(s)$ for arbitrary $A \subseteq S_\bot$. Note that $\mathcal{V}(S)$ is a closed lower subset of $\mathbb{I}^S$, also closed under subconvex combinations. In particular, $\mathcal{V}(S)$ is a compact convex ordered space.

There is a canonical inclusion

$$\eta\colon S_\bot \to \mathcal{V}(S)$$

sending $\bot$ to the constant map with value 0 and $s \in S$ to the Dirac measure $\eta(s)$ defined by $\eta(s)(t) = 1$ if $s = t$ and $\eta(s)(t) = 0$ otherwise.

The upper powerdomain $\mathcal{P}_U\mathcal{V}(S)$ consists of all subsets Q of $\mathcal{V}(S)$ which are **nonempty**, **compact**, **convex** and **saturated** and is ordered by reverse inclusion $Q_1 \sqsubseteq Q_2$ iff $Q_1 \supseteq Q_2$. For a directed family $(Q_i)_{i\in I}$ in $\mathcal{P}_U\mathcal{V}(S)$ its intersection $\bigcap_{i\in I} Q_i$ is again in $\mathcal{P}_U\mathcal{V}(S)$ whence $\left(\mathcal{P}_U\mathcal{V}(S), \sqsubseteq\right)$ is a dcpo with $\bigsqcup_i Q_i = \bigcap_{i\in I} Q_i$. There is a canonical inclusion

$$i = (\mu \mapsto \uparrow\mu)\colon \mathcal{V}(S) \to \mathcal{P}_U\mathcal{V}(S)$$

which is easily seen to be Scott-continuous. Composing the two canonical maps we obtain a canonical map

$$\varepsilon = i \circ \eta = \Big(s \mapsto \uparrow\eta(s)\Big)\colon S \to \mathcal{P}_U\mathcal{V}(S)$$

The semantics we will define for $\mathcal{L}_p$ will associate with every program P a

function $[\![P]\!] : S \to \mathcal{P}_U\mathcal{V}(S)$. For interpreting probabilistic choice ${}_p\oplus$ we need the following lemma.

Lemma 1 *For $Q_1, Q_2 \in \mathcal{P}_U(\mathcal{V}(S))$ and $0 < p < 1$, the convex combination*

$$Q_1 \ {}_p\oplus Q_2 = pQ_1 + (1-p)Q_2 = \{p\mu_1 + (1-p)\mu_2 \mid \mu_1 \in Q_1, \mu_2 \in Q_2\}$$

is again a member of $\mathcal{P}_U\mathcal{V}(S)$.

PROOF. [1] Being the image of the compact convex set $Q_1 \times Q_2$ under the continuous affine map $(\mu_1, \mu_2) \mapsto p\mu_1 + (1-p)\mu_2$, the set $pQ_1 + (1-p)Q_2$ is also compact and convex. In order to prove that it is saturated, let $p\mu_1 + (1-p)\mu_2 \leq \mu \in \mathcal{V}(S)$ for some $\mu_1 \in Q_1$, $\mu_2 \in Q_2$. Let $r_s = \frac{\mu(s)}{p\mu_1(s) + (1-p)\mu_2(s)}$ whenever the denominator is not 0. Clearly $r(s) \geq 1$. Define $\mu_1'(s) = r_s\mu_1(s)$ and $\mu_2'(s) = r_s\mu_2(s)$ for all $s \in S$ for which r_s was defined, and let $\mu_1'(s) = \mu_2'(s) = \mu(s)$ for all other $s \in S$. Then $\sum_{s\in S} \mu_1'(s) \leq \sum_{s\in S} \mu(s) \leq 1$, whence $\mu_1' \in \mathcal{V}(S)$, and similarly $\mu_2' \in \mathcal{V}(S)$. Further, $\mu_1 \leq \mu_1'$ and $\mu_2 \leq \mu_2'$, whence $\mu_1' \in Q_1$ and $\mu_2' \in Q_2$, and $\mu = p\mu_1' + (1-p)\mu_2' \in pQ_1 + (1-p)Q_2$.

For interpreting $\sqcap$ we need the existence of binary infima in $\mathcal{P}_U\mathcal{V}(S)$ as guaranteed by the following lemma.

Lemma 2 *For any two members Q_1, Q_2 of $\mathcal{P}_U\mathcal{V}(S)$, the convex hull*

$$Q_1 \sqcap Q_2 = \mathsf{conv}(Q_1 \cup Q_2)$$

is again compact, convex and saturated and, hence, the smallest member of $\mathcal{P}_U\mathcal{V}(S)$ containing Q_1 and Q_2.

PROOF. The convex hull of $Q_1 \cup Q_2$ is equal to $\bigcup_{p\in\mathbb{I}} pQ_1 + (1-p)Q_2$. Being the union of sets that are saturated by the previous lemma, $\mathsf{conv}(Q_1 \cup Q_2)$

[1] Added November 23, 2008: This proof is completely wrong. The mistake occurs where we claim that $\sum_{s\in S} \mu_1'(s) \leq \sum_{s\in S} \mu(s) \leq 1$, whence $\mu_1' \in \mathcal{V}(S)$. But the lemma is true. For proving that $pQ_1 + (1-p)Q_2$ is saturated, it suffices to prove this in the case where $Q_1 = \uparrow\mu_1$ and $Q_2 = \uparrow\mu_2$. In this case $p \cdot \uparrow\mu_1 + (1-p) \cdot \mu_2 = \uparrow(p\mu_1 + (1-p)\mu_2)$.

PROOF. We first note that $r\cdot\mathcal{V}(S) + s\cdot\mathcal{V}(S) = (r+s)\cdot\mathcal{V}(S)$. Secondly we note that $\uparrow\mu_1 = \mu_1 + (1 - ||\mu_1||_1)\mathcal{V}(S)$, where $||\mu_1||_1 = \sum_s \mu_1(s)$. Hence, $p\cdot\uparrow\mu_1 + (1-p)\cdot\uparrow\mu_2 = p\mu_1 + p(1 - ||\mu_1||_1)\mathcal{V}(S) + (1-p)\mu_2 + (1-p)(1 - ||\mu_2||_1)\mathcal{V}(S) = p\mu_1 + (1-p)\mu_2 + \big(p(1 - ||\mu_1||_1) + (1-p)(1 - ||\mu_2||_1)\big)\mathcal{V}(S) = p\mu_1 + (1-p)\mu_2 + \big(1 - (p||\mu_1||_1 + (1-p)||\mu_2||_1)\big)\mathcal{V}(S) = p\mu_1 + (1-p)\mu_2 + \big(1 - ||p\mu_1 + (1-p)\mu_2||_1\big)\mathcal{V}(S) = \uparrow(p\mu_1 + (1-p)\mu_2)$.

is saturated, too. It also is compact and convex, as it is the image of the compact set $[0,1]\times Q_1\times Q_2$ under under the continuous affine map $(p,\mu_1,\mu_2)\mapsto p\mu_1+(1-p)\mu_2$.

In order to define the semantics of composition we have to be able to lift a function $f\colon S\to\mathcal{P}_U\mathcal{V}(S)$ to a Scott-continuous function $f^\dagger\colon\mathcal{P}_U\mathcal{V}(S)\to\mathcal{P}_U\mathcal{V}(S)$ because then we may define $[\![P_1;P_2]\!]$ as $[\![P_2]\!]^\dagger\circ[\![P_1]\!]$. Moreover, in order to define the semantics of recursive programs it is necessary that the lifting operation

$$(-)^\dagger\colon[S\to\mathcal{P}_U\mathcal{V}(S)]\to[\mathcal{P}_U\mathcal{V}(S)\to\mathcal{P}_U\mathcal{V}(S)]$$

itself is Scott-continuous. For this purpose it is helpful to exploit the fact that $\mathcal{P}_U\mathcal{V}(S)$ is isomorphic to a certain set $\mathcal{G}(S)$ of Scott-continuous functions from $\mathbb{I}^S$ to $\mathbb{I}$ (see Appendix A for details).

Let $\mathcal{G}(S)$ be the set of all Scott-continuous $G\colon\mathbb{I}^S\to\mathbb{I}$ such that for all $\gamma,\beta\in\mathbb{I}^S$ and $r,t\in\mathbb{I}$ with $r+t\leq 1$ it holds that

$$G(r\gamma+t\beta)\geq rG(\gamma)+tG(\beta)$$
$$G(r\gamma+t\mathbf{1})\leq rG(\gamma)+t$$

where $\mathbf{1}$ denotes the constant function with value 1. For $t=0$, these two equations imply that $G(r\gamma)=rG(\gamma)$ for all $r\in\mathbb{I}$; this means that G is superlinear and satisfies condition (*) in the terminology of Appendix A. Due to lack of a better name we will refer to the functionals in $\mathcal{G}(S)$ as "good" functionals.

By definition, $\mathcal{G}(S)$ is a subset of the set $\mathbb{I}^{\mathbb{I}^S}$ of all functions $G\colon\mathbb{I}^S\to\mathbb{I}$. From our preliminaries, replacing there S by $\mathbb{I}^S$, we know that $\mathbb{I}^{\mathbb{I}^S}$ is a dcpo with a Scott-continuous meet operation and subconvex combinations which are Scott-continuous in each argument, where the order relation, directed suprema, binary meets and subconvex combinations are defined pointwise. It is straightforward to verify that $\mathcal{G}(S)$ is closed under all of these operations:

Lemma 3

(a) *For every directed family $(G_i)_i$ in $\mathcal{G}(S)$, the (pointwise) supremum $G(\gamma)=\sup_i G_i(\gamma)$ is again a member of $\mathcal{G}(S)$.*
(b) *For G_1 and G_2 in $\mathcal{G}(S)$, the (pointwise) meet $G_1\wedge G_2$ is again a member of $\mathcal{G}(S)$.*
(c) *For G_1 and G_2 in $\mathcal{G}(S)$, the (pointwise defined) subconvex combination rG_1+tG_2 is again a member of $\mathcal{G}(S)$, where $r,t\in\mathbb{I}$ with $r+t\leq 1$.*

Thus $\mathcal{G}(S)$ is a dcpo with Scott-continuous binary meets and Scott-continuous subconvex combinations.

By Proposition 20 in the Appendix there is an order isomorphism $\Phi\colon \mathcal{P}_U\mathcal{V}(S) \to \mathcal{G}(S)$. Using the notation $\langle \mu, \gamma\rangle = \sum_s \mu(s)\gamma(s)$, it is given by

$$\Phi_Q(\gamma) = \min_{\mu\in Q}\langle \mu, \gamma\rangle \quad \text{for all } \gamma \in \mathbb{I}^S$$

the inverse being the map $\Psi\colon \mathcal{G}(S) \to \mathcal{P}_U\mathcal{V}(S)$ given by

$$\Psi_G = \{\mu \in \mathcal{V}(S) \mid \langle \mu, \gamma\rangle \geq G(\gamma) \text{ for all } \gamma \in \mathbb{I}^S\}$$

Next we show that Φ and Ψ preserve all relevant structure.

Lemma 4

(a) Φ *and* Ψ *are Scott-continuous.*
(b) Φ *and* Ψ *preserve binary meets, i.e.* $\Phi_{Q_1\sqcap Q_2} = \Phi_{Q_1} \wedge \Phi_{Q_2}$ *and* $\Psi_{G_1\wedge G_2} = \Psi_{G_1} \sqcap \Psi_{G_2}$.
(c) Φ *and* Ψ *preserve convex combinations, i.e.* $\Phi_{Q_1\ {}_p\!\oplus\ Q_2} = p\Phi_{Q_1} + (1-p)\Phi_{Q_2}$ *and* $\Psi_{pG_1+(1-p)G_2} = \Psi_{G_1\ {}_p\!\oplus\ G_2}$.

PROOF. (a) and (b) are just consequences of the order isomorphism property. Claim (c) is shown by the following calculation

$$\begin{aligned}
p\Phi_{Q_1}(\gamma) + (1-p)\Phi_{Q_2}(\gamma) &= p\min_{\mu_1\in Q_1}\langle \mu_1, \gamma\rangle + (1-p)\min_{\mu_2\in Q_2}\langle \mu_2, \gamma\rangle \\
&= \min_{\mu_1\in Q_1}\langle p\mu_1, \gamma\rangle + \min_{\mu_2\in Q_2}\langle (1-p)\mu_2, \gamma\rangle \\
&= \min_{\mu_1\in pQ_1}\langle \mu_1, \gamma\rangle + \min_{\mu_2\in (1-p)Q_2}\langle \mu_2, \gamma\rangle \\
&= \min_{\mu_1\in pQ_1, \mu_2\in (1-p)Q_2}(\langle \mu_1, \gamma\rangle + \langle \mu_2, \gamma\rangle) \\
&= \min_{\mu_1\in pQ_1, \mu_2\in (1-p)Q_2}\langle \mu_1 + \mu_2, \gamma\rangle \\
&= \min_{\mu\in pQ_1+(1-p)Q_2}\langle \mu, \gamma\rangle \\
&= \Phi_{pQ_1+(1-p)Q_2}(\gamma)
\end{aligned}$$

where $\gamma \in \mathbb{I}^S$ and $p \in \mathbb{I}$.

As the meet operation and subconvex combinations are Scott-continuous in $\mathcal{G}(S)$, the preceding lemma allows us to conclude:

Corollary 5 *The operations* $\sqcap$ *and* ${}_p\!\oplus$ *are Scott-continuous on* $\mathcal{P}_U\mathcal{V}(S)$.

Recall that in continuation semantics (see e.g. [BHM]) a function $f : S \to \mathbb{I}^{\mathbb{I}^S}$ is lifted to the function

$$f^{\#} = \Big(G \mapsto \lambda\gamma.G(\lambda s.f(s)(\gamma)\Big) \colon [\mathbb{I}^S \to \mathbb{I}] \to [\mathbb{I}^S \to \mathbb{I}]$$

which is Scott-continuous since it is λ-definable. The so defined lifting operation $(-)^{\#}$ validates the laws

$$f^{\#} \circ \eta = f \qquad\qquad g^{\#} \circ f^{\#} = (g^{\#} \circ f)^{\#}$$

where $\eta = \lambda s.\lambda\gamma.\gamma(s) \colon S \to \mathbb{I}^{\mathbb{I}^S}$. These laws guarantee that $[S \to \mathbb{I}^{\mathbb{I}^S}]$ is a monoid w.r.t. (Kleisli) composition $f; g = g^{\#} \circ f$ with unit η. The next lemma tells us that this lifting restricts to $\mathcal{G}(S)$ in the following sense.

Lemma 6 *For $f \colon S \to \mathcal{G}(S)$ its lifting $f^{\#}$ restricts to a Scott-continuous endomap on $\mathcal{G}(S)$ which preserves subconvex combinations and binary meets. Moreover, the restricted lifting map*

$$f \mapsto f^{\#} \colon [S \to \mathcal{G}(S)] \to [\mathcal{G}(S) \to \mathcal{G}(S))]$$

is itself Scott-continuous.

PROOF. For $f : S \to \mathcal{G}(S) \subseteq [\mathbb{I}^S \to \mathbb{I}]$ its lifting $f^{\#}$ is λ-definable and thus Scott-continuous. Moreover, the map

$$(-)^{\#} : [S \to [\mathbb{I}^S \to \mathbb{I}]] \to [[\mathbb{I}^S \to \mathbb{I}] \to [\mathbb{I}^S \to \mathbb{I}]]$$

itself is also λ-definable and thus Scott-continuous.

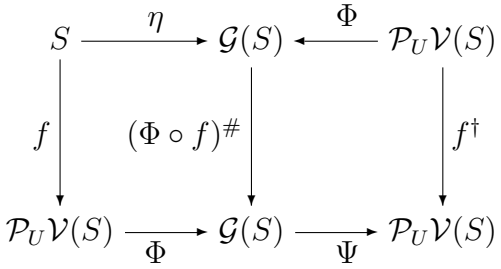
Since $f^{\#}(G)(\gamma) = G(\lambda s.f(s)(\gamma))$ it validates all inequalities holding for G. Thus $f^{\#}$ sends elements of $\mathcal{G}(S)$ to elements of $\mathcal{G}(S)$ and preserves the operations $\wedge$ and ${}_p\oplus$ since they are defined pointwise.

Using the isomorphism $\Phi \colon \mathcal{P}_U\mathcal{V}(S) \to \mathcal{G}(S)$ of Proposition 20 we can define the lifting of maps $S \to \mathcal{P}_U\mathcal{V}(S)$ as follows

Definition 7 *For $f : S \to \mathcal{P}_U\mathcal{V}(S)$ let*

$$f^{\dagger} = \Psi \circ (\Phi \circ f)^{\#} \circ \Phi : \mathcal{P}_U\mathcal{V}(S) \to \mathcal{P}_U\mathcal{V}(S)$$

as illustrated by



where $\eta(s) = \lambda\gamma.\gamma(s)$.

The so defined $f^\dagger$ is Scott-continuous and preserves ${}_p\oplus$ and $\sqcap$. Moreover, this lifting operation $(-)^\dagger$ is again Scott-continuous and satisfies the laws

$$f^\dagger \circ \eta = f \qquad\qquad g^\dagger \circ f^\dagger = (g^\dagger \circ f)^\dagger$$

for all $f, g : S \to \mathcal{P}_U\mathcal{V}(S)$.

Now we are ready to give the clauses for the direct semantics for $\mathcal{L}_p$.

Definition 8 *Let* Act *be some set of endofunctions on* S *and* BExp *be some set of functions from* S *to* $\{0,1\}$. *The* direct semantics *associating to every* $\mathcal{L}_p$ *program* P *a function*

$$[\![P]\!] \colon S \to \mathcal{P}_U\mathcal{V}(S)$$

is defined inductively by the following semantic clauses

$$\begin{aligned}
&[\![a]\!] = \eta \circ a\\
&[\![P_1; P_2]\!] = [\![P_2]\!]^\dagger \circ [\![P_1]\!]\\
&[\![P_1 {}_p\!\oplus P_2]\!](s) = p \cdot [\![P_1]\!](s) + (1-p) \cdot [\![P_2]\!](s)\\
&[\![P_1 [\!] P_2]\!](s) = [\![P_1]\!](s) \sqcap [\![P_2]\!](s)\\
&[\![\mathbf{cond}(b, P_1, P_2)]\!](s) = b(s) \cdot [\![P_1]\!](s) + \neg b(s) \cdot [\![P_2]\!](s)\\
&[\![\mathbf{while}(b, P)]\!] = \mathsf{Minfix} f.\lambda s.\ b(s) \cdot f^\dagger([\![P]\!](s)) + \neg b(s) \cdot \varepsilon(s))
\end{aligned}$$

where s *ranges over* S, b *over* BExp *and* $\neg b(s) = 1-b(s)$. *Further,* $\mathsf{Minfix}X.E(X)$ *denotes the least fixed point of the map* $X \mapsto E(X)$ *which is well defined, if* X *ranges over a dcpo with a smallest element and if the map* $X \mapsto E(X)$ *is Scott-continuous, which is the case in our setting.*

Next we give an explicit construction of $f^\dagger : \mathcal{P}_U\mathcal{V}(S) \to \mathcal{P}_U\mathcal{V}(S)$ from $f : S \to \mathcal{P}_U\mathcal{V}(S)$ which has an immediate intuitive operational reading.

Lemma 9 *For* $f : S \to \mathcal{P}_U\mathcal{V}(S)$ *its lifting* $f^\dagger : \mathcal{P}_U\mathcal{V}(S) \to \mathcal{P}_U\mathcal{V}(S))$ *is given*

by

$$f^\dagger(Q) = \uparrow\left\{\sum_s \mu(s)h(s) \mid h \in \prod_{s\in S} f(s) \text{ and } \mu \in Q\right\}$$

for $Q \in \mathcal{P}_U\mathcal{V}(S)$. In particular, we have

$$f^\dagger(\uparrow\mu) = \uparrow\left\{\sum_s \mu(s)h(s) \mid h \in \prod_{s\in S} f(s)\right\}$$

for $\mu \in \mathcal{V}(S)$.

PROOF. Let $Q \in \mathcal{P}_U\mathcal{V}(S)$. The set

$$M_f = \left\{\sum_s \mu(s)h(s) \mid h \in \prod_s f(s) \text{ and } \mu \in Q\right\}$$

is convex and nonempty and, moreover, compact since it arises as image under a continuous function of the compact set $Q \times \prod_s f(s)$. Thus its upward closure is an element of $\mathcal{P}_U\mathcal{V}(S)$. Thus, for showing the desired equality by Definition 7 it suffices to show that

$$(\Phi \circ f)^\#(\Phi_Q) = \Phi_{\uparrow M_f}$$

For this purpose for $\gamma \in \mathbb{I}^S$ we calculate as follows

$$\begin{aligned}
(\Phi \circ f)^\#(\Phi_Q(\gamma)) &= \Phi_Q(\lambda s.\,(\Phi \circ f)(s)(\gamma)) = \Phi_Q(\lambda s.\,\Phi_{f(s)}(\gamma)) \\
&= \min_{\mu\in Q} \textstyle\sum_s \mu(s) \cdot \Phi_{f(s)}(\gamma) \\
&= \min_{\mu\in Q} \textstyle\sum_s \mu(s) \cdot \min_{\nu\in f(s)} \langle \nu, \gamma\rangle \\
&\overset{(*)}{=} \min_{\mu\in Q} \min_{h\in \Pi_s f(s)} \textstyle\sum_s \mu(s) \cdot \langle h(s), \gamma\rangle \\
&= \min_{\mu\in Q} \min_{h\in \Pi_s f(s)} \langle \textstyle\sum_s \mu(s)h(s), \gamma\rangle \\
&= \min_{\nu\in M_f} \langle \nu, \gamma\rangle = \min_{\nu\in \uparrow M_f} \langle \nu, \gamma\rangle \\
&= \Phi_{\uparrow M_f}(\gamma)
\end{aligned}$$

where $(*)$ follows from the fact that for every $s \in S$ we may choose an $h(s) \in f(s)$ with $\langle h(s), \gamma\rangle = \min_{\nu\in f(s)}\langle \nu, \gamma\rangle$.
The particular case follows from the fact that $\sum_s \mu(s)h(s) \leq \sum_s \nu(s)h(s)$ whenever $\mu \leq \nu$.

Thus, according to this lemma $\mu' \in f^\dagger(Q)$ iff its is above some subconvex combination $\sum_s \mu(s) \cdot h(s)$ of possible results $h(s) \in f(s)$ where the weights are given by some $\mu \in Q$.

An immediate consequence of Lemma 9 is the following explicitation of the semantics of sequential compositions of programs

$$[\![P_1; P_2]\!](s) = [\![P_2]\!]^\dagger([\![P_1]\!](s)) = \bigg\uparrow \left\{ \sum_s \mu(s)h(s) \mid \mu \in [\![P_1]\!](s), h \in \prod_s [\![P_2]\!](s) \right\}$$

which is taken as a defining clause of the state transformer semantics presented in [MM] (Def. 5.4.7 on p. 140).

4 From state to predicate transformer semantics

The idea of *predicate transformer semantics* is to consider instead of $f : S \to S$ (thought of as the meaning of a total deterministic program) the function $\Sigma^f : \Sigma^S \to \Sigma^S$ where $\Sigma = \{0, 1\}$ is the Sierpiński space of truth values. The advantage of such a view is that it is closer to reasoning about programs than its direct semantics because $A \subseteq \Sigma^f(B)$ iff for all $s \in A$ the result $f(s) \in B$, usually denoted as $\{A\}f\{B\}$. Thus $\Sigma^f(B)$ is the *weakest precondition* guaranteeing that the execution of the program denoting f results in a state which is an element of B. Predicate transformers $T : \Sigma^S \to \Sigma^S$ with $T = \Sigma^f$ for some $f : S \to S$ can be characterised as those maps $\Sigma^S \to \Sigma^S$ such that Φ preserves arbitrary suprema and finite infima.

The intention of this section is to study predicate transformer semantics for programs in the language $\mathcal{L}_p$ and how to derive it from its direct semantics. Recall from the previous section that the interpretation of an $\mathcal{L}_p$ program is a function $S \to \mathcal{P}_U\mathcal{V}(S)$ which by Proposition 20 may be identified with a function $f : S \to \mathcal{G}(S) \subseteq [\mathbb{I}^S \to \mathbb{I}]$ which uniquely corresponds to a Scott-continuous function $\mathsf{Wp}(f) : \mathbb{I}^S \to \mathbb{I}^S$ as described in the following theorem.

Theorem 10 *The function* $\mathsf{Wp} : [S \to \mathcal{G}(S)] \to [\mathbb{I}^S \to \mathbb{I}^S]$ *with* $\mathsf{Wp}(f)(\gamma)(s) = f(s)(\gamma)$ *is Scott-continuous and one-to-one. The image of* Wp *consists precisely of those Scott-continuous functions* $T : \mathbb{I}^S \to \mathbb{I}^S$ *satisfying the conditions*

(1) $T(r\gamma + t\beta) \geq rT(\gamma) + tT(\beta)$
(2) $T(r\gamma + t\mathbf{1}) \leq rT(\gamma) + t\mathbf{1}$

for all $\gamma, \beta \in \mathbb{I}^S$ *and* $r, t \in \mathbb{I}$ *with* $r + t \leq 1$. *We write* PT *for the image of* Wp.

PROOF. The function $\mathsf{Wp}(f)$ is Scott-continuous since it is λ-definable and for the same reason the function Wp itself is Scott-continuous, too.

A function $T : \mathbb{I}^S \to \mathbb{I}^S$ is in the image of Wp iff for all $s \in S$ the function $\lambda\gamma.T(\gamma)(s) \in \mathcal{G}(S)$ which is equivalent to the conditions (1) and (2) which

express precisely this requirement.

For $f : S \to \mathcal{P}_U\mathcal{V}(S)$ the associated predicate transformer $\mathsf{Wp}(\Phi \circ f)$ is also denoted as $\mathsf{Wp}(f)$ and can be described explicitly as follows.

Corollary 11 *For $f : S \to \mathcal{P}_U\mathcal{V}(S)$ we have*

$$\mathsf{Wp}(f)(\gamma)(s) = \mathsf{Wp}(\Phi \circ f)(\gamma)(s) = \Phi_{f(s)}(\gamma) = \inf_{\mu \in f(s)} \langle \mu, \gamma \rangle$$

for all $\gamma \in \mathbb{I}^S$ and $s \in S$.

In [MM] the elements of $\mathbb{I}^S$ are called *expectations*. We do not follow this terminology because they have nothing to do with *expectation values* in the sense of probability theory. Nor should elements of $\mathbb{I}^S$ be thought of as probability distributions since in general they do not sum up to a number less or equal 1. In our opinion only elements of the form $\mathsf{Wp}(f)(b)$ with $b \in \{0,1\}^S$ have an intuitive meaning [2] whereas the more general $\mathsf{Wp}(f)(\beta)$ with $\beta \in \mathbb{I}^S$ have a merely auxiliary status because they arise as intermediate steps when computing $\mathsf{Wp}([\![P;Q]\!])$ as in the subsequent Theorem 13.

The following Theorem 13 gives rise to a definition of a predicate transformer semantics for $\mathcal{L}_p$ which fully avoids any kind of powerdomains and thus makes it easier to reason about $\mathcal{L}_p$ programs and was introduced for this purpose in [MM] (for a variant of $\mathcal{L} - p$ called pGCL). We think that the following Theorem 13 gives a kind of "rational reconstruction" of this predicate transformer semantics because it shows how it can be derived from the direct semantics and the function Wp which are both well-motivated.

Definition 12 *For an $\mathcal{L}_p$ program P let $\mathsf{wp}(P) = \mathsf{Wp}([\![P]\!])$ be the* predicate transformer *associated with P.*

Theorem 13 *The following equations hold for wp and characterise it uniquely*

$$\begin{aligned}
&\mathsf{wp}(a)(\gamma) = \gamma \circ a \\
&\mathsf{wp}(P_1; P_2) = \mathsf{wp}(P_1) \circ \mathsf{wp}(P_2) \\
&\mathsf{wp}(P_1 {}_p\!\oplus P_2)(\gamma) = p \cdot \mathsf{wp}(P_1)(\gamma) + (1-p) \cdot \mathsf{wp}(P_2)(\gamma) \\
&\mathsf{wp}(P_1 [\!] P_2)(\gamma) = \mathsf{wp}(P_1)(\gamma) \wedge \mathsf{wp}(P_2)(\gamma) \\
&\mathsf{wp}(\mathbf{cond}(b, P_1, P_2))(\gamma) = (b \wedge \mathsf{wp}(P_1)(\gamma)) \vee (\neg b \wedge \mathsf{wp}(P_2)(\gamma)) \\
&\mathsf{wp}(\mathbf{while}(b, P))(\gamma) = \mathsf{Minfix}\ \beta.\ (b \wedge \mathsf{wp}(P)(\beta)) \vee (\neg b \wedge \gamma)
\end{aligned}$$

[2] namely as the probability that the program with direct semantics f terminates with final state in b when started in state s

where $\wedge$ and $\vee$ stand for the pointwise infimum and supremum on $\mathbb{I}^S$, respectively, $(\neg b)(s) = 1 - b(s)$ and $\mathsf{Minfix}X.E[X]$ stands for the least fixed point of the Scott-continuous function $\lambda X.E[X]$.

PROOF. The crucial cases are composition and the while-loop whereas all other cases are straightforward and left to the reader. For sake of simplicity we work rather on the side of $\mathcal{G}(S)$ than on the side of the more complicated $\mathcal{P}_U\mathcal{V}(S)$ which does not do any damage since they are isomorphic by our crucial Proposition 20.

For $\mathcal{L}_p$-programs P_1 and P_2 we have

$$\begin{aligned}
\mathsf{wp}(P_1;P_2)(\gamma)(s) &= \mathsf{Wp}([\![P_1;P_2]\!])(\gamma)(s) = \mathsf{Wp}([\![P_2]\!]^{\#} \circ [\![P_1]\!])(\gamma)(s) \\
&= ([\![P_2]\!]^{\#} \circ [\![P_1]\!])(s)(\gamma) = [\![P_2]\!]^{\#}([\![P_1]\!](s))(\gamma) \\
&= [\![P_1]\!](s)(\lambda s.[\![P_2]\!](s)(\gamma)) \qquad \text{(by def. of } (-)^{\#}) \\
&= \mathsf{Wp}([\![P_1]\!])(\mathsf{Wp}([\![P_2]\!])(\gamma))(s) \\
&= \Big(\mathsf{Wp}([\![P_1]\!]) \circ \mathsf{Wp}([\![P_2]\!])\Big)(\gamma)(s)
\end{aligned}$$

Next we consider the case of while-loops. For $\gamma \in \mathbb{I}^S$ we define the auxiliary functions

$$\begin{aligned}
&h_\gamma(f) := \mathsf{Wp}(f)(\gamma) \\
&k(f) := \lambda s{:}S.\, b(s) \cdot f^{\#}([\![P]\!](s)) + \neg b(s) \cdot \eta(s) \\
&g(\beta) := (b \wedge \mathsf{wp}(P)(\beta)) \vee (\neg b \wedge \gamma)
\end{aligned}$$

One easily checks that h_γ is strict (i.e. preserves the least element) and the diagram

$$\begin{array}{ccc}
\mathcal{G}(S)^S & \xrightarrow{h_\gamma} & \mathbb{I}^S \\
{\scriptstyle k}\downarrow & & \downarrow{\scriptstyle g} \\
\mathcal{G}(S)^S & \xrightarrow[h_\gamma]{} & \mathbb{I}^S
\end{array}$$

commutes from which it follows by Plotkin's Lemma on least fixed point operators (see [Plo] or [GHK$^+$] II-2.4) that

$$\mathsf{wp}(\mathbf{while}(b,P)) = h_\gamma(\mu(k)) = \mu(g) = \mu\beta.\,(b \wedge \mathsf{wp}(P)(\beta)) \vee (\neg b \wedge \gamma)$$

as desired.

A Appendix: The Minkowski duality

Although for semantics the case of a countable set of states is the most relevant one, the following developments hold for any set S considered as a discrete set without any topology or order.

We consider the linear subspace ℓ^∞ of the vector space $\mathbb{R}^S$ consisting of all bounded functions $\gamma\colon S \to \mathbb{R}$. We equip this linear subspace with the topology of pointwise convergence, that is, the topology induced from the product topology on $\mathbb{R}^S$ as in the preliminary section 2, and with the pointwise defined order $\beta \leq \gamma$ iff $\beta(s) \leq \gamma(s)$ for all $s \in S$. The graph of this order is closed. The positive cone, i.e. the set of all nonnegative functions $\gamma \in \ell^\infty$, is denoted by ℓ^∞_+.

As before, we use the notation $\mathbf{1}$ for the constant function with value 1. Then $\mathbb{I}^S = \{\gamma \in \ell^\infty_+ \mid \gamma \leq \mathbf{1}\}$ is a compact convex subset of the vector space ℓ^∞ which is also closed under subconvex combinations.

In ℓ^∞ we consider the linear subspace ℓ^1 of all functions $\mu\colon S \to \mathbb{R}$ such that $\sum_{s\in S} |\mu(s)| < +\infty$ with the norm $|\mu|_1 = \sum_{s\in S} |\mu(s)|$. However, we will not consider the norm topology but instead the topology of pointwise convergence as above. The positive cone, i.e. the set of all nonnegative functions $\gamma \in \ell^1$, is denoted by ℓ^1_+.

The set $\mathcal{V}(S)$ of subprobability distributions on S consists of all $\mu \in \ell^1_+$ with $|\mu|_1 \leq 1$.

Lemma 14 *If $A \subseteq \ell^\infty$ is convex (resp. compact) then its saturation $\uparrow A = \{\gamma \in \ell^\infty \mid \gamma \geq \alpha$ for some $\alpha \in A\}$ is also convex (resp. closed).*

PROOF. First consider a convex set A. If $\gamma_1, \gamma_2 \in \uparrow A$, then there are $\beta_1, \beta_2 \in A$ with $\gamma_1 \geq \beta_1$ and $\gamma_2 \geq \beta_2$. For $0 < p < 1$, we then have $p\gamma_1 + (1-p)\gamma_2 \geq p\beta_1 + (1-p)\beta_2 \in A$, whence $p\gamma_1 + (1-p)\gamma_2 \in \uparrow A$.

When A is compact, consider a generalised sequence γ_i in $\uparrow A$ (indexed by some directed set I) converging to some γ. There are $\beta_i \in A$ with $\beta_i \leq \gamma_i$ for every i. In the compact set A, the β_i have a subsequence β_{i_j} converging to some $\beta \in A$. As $\beta_{i_j} \leq \gamma_{i_j}$ and as the graph of the order is closed, we conclude that $\beta = \lim_j \beta_{i_j} \leq \lim_j \gamma_{i_j} = \gamma$, whence $\gamma \in \uparrow A$.

Recall that a function f from a topological space X into $\mathbb{R}$ (or into a subset of $\mathbb{R}$) is *lower semicontinuous* if the set of all $x \in X$ such that $f(x) > r$ is open in X for every $r \in \mathbb{R}$. Next we show that, in our setting, lower semicontinuity

is equivalent to Scott continuity, a fact that will be used subsequently without further mention.

Lemma 15 *An function f from ℓ^∞_+ $\left(\ell^1_+, \mathbb{I}^S, \mathcal{V}(S)\text{, respectively}\right)$ to $\mathbb{R}_+$ is order preserving and lower semicontinuous if and only if it is Scott-continuous.*

PROOF. For the forward direction it suffices to notice that, if γ is the pointwise supremum of a directed family γ_i and if U is a basic open neighborhood of γ, then $\gamma_i \in U$ for some i. For the reverse direction fix γ and consider an arbitrary $r < f(\gamma)$. For every finite subset F of S of cardinality n define $\gamma_F(s) = \max(\gamma(s) - \frac{1}{n}, 0)$, whenever $s \in F$, and $\gamma_F(s) = 0$, else. Then $\gamma = \sup_F \gamma_F$. Since f is assumed to be Scott-continuous we have $f(\gamma) = \sup_F f(\gamma_F)$. Thus, there is an F such that $f(\gamma_F) > r$. For all $\beta \geq \gamma_F$, we then have $f(\beta) > r$, as f is supposed to be order preserving, and the set of all these β is a neighborhood of γ.

Let V_+ be any of the positive cones $\mathbb{R}_+$, ℓ^∞_+, ℓ^1_+. Recall that a function $f\colon V_+ \to \mathbb{R}_+$ is called

homogeneous if $f(r\gamma) = rf(\gamma)$ for all $r \in \mathbb{R}_+$,
superadditive if $f(\gamma + \beta) \geq f(\gamma) + f(\beta)$,
superlinear if it is homogeneous and superadditive,
linear if $f(r\gamma + t\beta) = rf(\gamma) + tf(\beta)$ whenever $r, t \in \mathbb{R}_+$.

In V_+, consider the subset $K = \mathbb{I}, \mathbb{I}^S, \mathcal{V}(S)$, respectively, which is a closed lower set also closed under subconvex combinations.

We want to apply the above terminology to functions $g\colon K \to \mathbb{I}$. As addition and scalar multiplication lead out of K, we have to modify the definition in the following way: A function $g\colon K \to \mathbb{I}$ is called

homogeneous if $g(r\gamma) = rg(\gamma)$ for all $r \in \mathbb{I}$,
0-concave if $g(r\gamma + t\beta) \geq rg(\gamma) + tg(\beta)$ whenever $r, t \in \mathbb{I}$ and $r + t \leq 1$,
superlinear if it is 0-concave and homogeneous,
linear if $g(r\gamma + t\beta) = rg(\gamma) + tg(\beta)$ whenever $r, t \in \mathbb{I}$ and $r + t \leq 1$.

Each homogeneous functional $g\colon K \to \mathbb{I}$ has a unique extension to a homogeneous functional $\widehat{g}\colon V_+ \to \mathbb{R}_+$: for $\beta \in V_+$ there is a $\gamma \in K$ such that $\beta = r\gamma$ for some $r \in \mathbb{R}_+$, and if we set $\widehat{g}(\beta) = rg(\gamma)$ this value is independent of the choice of γ in K because of homogeneity.

The extension $\widehat{g}\colon V_+ \to \mathbb{R}_+$ of a homogeneous functional $g\colon K \to \mathbb{I}$ is superlinear, linear, and Scott-continuous, respectively, if g is.

Note that any superlinear functional $f\colon V_+ \to \mathbb{R}_+$ is order preserving. Indeed, if $\beta \leq \gamma$ in V_+, then $\gamma - \beta \in V_+$ and $f(\beta) \leq f(\beta) + f(\gamma - \beta) \leq f(\beta + (\gamma - \beta)) = f(\gamma)$ by superlinearity. We conclude that every superlinear functional $g\colon K \to \mathbb{I}$ is order preserving, too, as such a g can be extended to a superlinear functional on V_+. All of this applies in particular to linear functionals.

Although being defined as a subspace of ℓ^∞_+, the cone ℓ^1_+ should rather be considered as the dual of ℓ^∞_+ and vice versa as described in Lemma 16. In order to formulate this duality for every $\mu \in \ell^1_+$ and every $\gamma \in \ell^\infty_+$ consider

$$\langle \mu, \gamma \rangle = \sum_s \mu(s)\gamma(s)$$

which converges (absolutely) since $|\mu(s)\gamma(s)| \leq \mu(s)$ and $\sum_s \mu(s)$ converges by definition.

Lemma 16

(a) *The mapping*

$$(\mu, \gamma) \mapsto \langle \mu, \gamma \rangle\colon \ell^1_+ \times \ell^\infty_+ \to \mathbb{R}_+$$

is bilinear and Scott-continuous.

(b) *For every Scott-continuous linear functional $f\colon \ell^\infty_+ \to \mathbb{R}_+$ there is a (unique) $\mu \in \ell^1$ such that $f(\gamma) = \langle \mu, \gamma \rangle$ and $|\mu|_1 = f(\mathbf{1})$, and for every Scott-continuous linear functional $g\colon \ell^1_+ \to \mathbb{R}_+$ there is a (unique) $\gamma \in \ell^\infty_+$ such that $g(\mu) = \langle \mu, \gamma \rangle$.*

PROOF. (a) Bilinearity is straightforward. For $\gamma \in \ell^\infty_+$ and every finite subset $F \subseteq S$, we define $\gamma|_F(s) = \gamma(s)$, whenever $s \in F$, and $\gamma|_F(s) = 0$, else. Similarly, we define $\mu|_F$ for $\mu \in \ell^1_+$. The function $(\mu, \gamma) \mapsto \langle \mu|_F, \gamma|_F \rangle = \sum_{s \in F} \mu(s) \cdot \gamma(s)$ is continuous on $\ell^1_+ \times \ell^\infty_+$. As $\langle \mu, \gamma \rangle = \sum_{s \in S} \mu(s)\gamma(s) = \sup_F \sum_{s \in F} \mu(s) \cdot \gamma(s) = \sup_F \langle \mu|_F, \gamma|_F \rangle$, where F ranges over all finite subsets of S, the map $(\mu, \gamma) \mapsto \langle \mu, \gamma \rangle\colon \ell^1_+ \times \ell^\infty_+ \to \mathbb{R}_+$ is the (pointwise) supremum of a directed family of continuous functions and hence lower semicontinuous, whence Scott-continuous by Lemma 15.

(b) For every $s \in S$, we denote by $\eta(s)$ the Dirac measure $\eta(s)(s) = 1$ and $\eta(s)(t) = 0$ for all $t \neq s$. Every $\gamma \in \ell^\infty_+$ can be written in the form $\gamma = \sup_F \sum_{s \in F} \gamma(s)\eta(s)$, where F ranges over all finite subsets of S, and similarly for every $\mu \in \ell^1_+$

Let f be a Scott-continuous linear functional on ℓ^∞_+. We define $\mu(s) = f(\eta(s))$. Then $f(\gamma) = f(\sup_F \sum_{s \in F} \gamma(s)\eta(s)) = \sup_F \sum_{s \in F} \gamma(s)\mu(s) = \sum_{s \in S} \gamma(s)\mu(s) = \langle \mu, \gamma \rangle$, where we have used the Scott-continuity and the linearity of f. If we choose $\gamma = \mathbf{1}$, then $|\mu|_1 = \sum_{s \in S} \mu(s) = \langle \mu, \mathbf{1} \rangle = f(\mathbf{1}) < +\infty$, which shows that $\mu \in \ell^\infty_+$.

Now let g be any Scott-continuous linear functional on ℓ^1_+. We define $\gamma(s) = g(\eta(s))$. Then $\gamma(s) \geq 0$ and, for all $\mu \in \ell^1_+$, we have $g(\mu) = g(\sup_F \sum_{s\in F} \mu(s)\eta(s)) = \sup_F \sum_{s\in F} \mu(s)\gamma(s) = \sum_{s\in S} \mu(s)\gamma(s) = \langle \mu, \gamma\rangle$, where we have used the Scott-continuity and the linearity of g. Note that $\gamma \in \ell^\infty$. Indeed, if we had $\sup_s \gamma(s) = +\infty$, we could choose a sequence s_n in S such that $\gamma(s_n) \geq 2^{2n}$ and, for μ with $\mu(s_n) = 2^{-n}$ and $\mu(s) = 0$ for s different from all s_n, we would have $g(\mu) = \sum_n \mu(s_n)\gamma(s_n) \geq \sum_n 2^{-n} \cdot 2^{2n} \geq +\infty$, a contradiction.

¿From Lemma 16 we deduce a duality between $\mathbb{I}^S$ and $\mathcal{V}(S)$ as follows.

Corollary 17

(a) *The function*
$$(\mu, \gamma) \mapsto \langle \mu, \gamma\rangle \colon \mathcal{V}(S) \times \mathbb{I}^S \to \mathbb{I}$$
is bilinear and Scott-continuous.

(b) *For every Scott-continuous linear functional $f \colon \mathbb{I}^S \to \mathbb{I}$, there is a (unique) $\mu \in \mathcal{V}(S)$ such that $f(\gamma) = \langle \mu, \gamma\rangle$ for all $\gamma \in \mathbb{I}^S$ and for every Scott-continuous linear functional $g \colon \mathcal{V}(S) \to \mathbb{I}$ there is a (unique) $\gamma \in \mathbb{I}^S$ such that $g(\mu) = \langle \mu, \gamma\rangle$ for all $\mu \in \mathcal{V}(S)$.*

Note. Every Scott-continuous linear functional $f \colon \mathbb{I}^S \to \mathbb{I}$ is not only lower semicontinuous by Lemma 15, but also upper semicontinuous and hence continuous.

Indeed, let $f \colon \mathbb{I}^S \to \mathbb{I}$ be linear. If $f(\mathbf{1}) = 0$ then $f = 0$ and there is nothing to be shown. If $f(\mathbf{1}) > 0$ we replace f by $\frac{1}{f(\mathbf{1})} \cdot f$ and thus may suppose w.l.o.g. that $f(\mathbf{1}) = 1$. We then have $f(\gamma) = 1 - f(\mathbf{1} - \gamma)$ by linearity for all $\gamma \in \mathbb{I}^S$. Let γ_i be any downdirected family in $\mathbb{I}^S$ and $\gamma = \inf_i \gamma_i$. Then $\mathbf{1} - \gamma_i$ is an updirected family with supremum $\mathbf{1} - \gamma$. Since f is assumed as Scott-continuous we conclude $f(\mathbf{1} - \gamma) = \sup_i f(\mathbf{1} - \gamma_i)$. Hence $f(\gamma) = 1 - f(\mathbf{1} - \gamma) = \inf_i(1 - f(\mathbf{1} - \gamma_i)) = \inf_i f(\gamma_i)$. Thus lower semicontinuity implies upper semicontinuity.

There are discontinuous linear functionals $f \colon \mathbb{I}^S \to \mathbb{I}$, whenever S is infinite. Indeed, if $\mathcal{U}$ is a non-principal ultrafilter on S, then every $\gamma \in \mathbb{I}^S$ has a limit along $\mathcal{U}$ and $\gamma \mapsto \lim_{\mathcal{U}} \gamma$ is a linear functional which is not lower semicontinuous; indeed, $\lim_{\mathcal{U}} \mathbf{1}|_F = 0$ for every finite subset F, whence $0 = \sup_F \lim_{\mathcal{U}} \mathbf{1}|_F \neq \lim_{\mathcal{U}} \sup_F \mathbf{1}|_F = \lim_{\mathcal{U}} \mathbf{1} = 1$.

In contrast most Scott-continuous linear functionals on $\mathcal{V}(S)$ are not upper semicontinuous.

We now consider the powerdomain $\mathcal{P}_U\mathcal{V}(S)$ of all subsets of $\mathcal{V}(S)$ which are nonempty, compact, convex and saturated (i.e. upper sets in $\mathcal{V}(S)$).

We want to extend the correspondence between elements of $\mathcal{V}(S)$ and Scott-continuous linear functionals $f\colon \mathbb{I}^S \to \mathbb{I}$ (see corollary 17) to a Minkowski type correspondence between the sets $Q \in \mathcal{P}_U\mathcal{V}(S)$ and certain functionals $G\colon \mathbb{I}^S \to \mathbb{I}$. In his seminal paper [Min], H. Minkowski has established a one-to-one correspondence between compact convex sets in $\mathbb{R}^3$ and superlinear functionals on $\mathbb{R}^3$. There are many generalisations of Minkowski's duality to much more general situations (see e.g. [Tol]). But we could not find the result that we need in the literature. In the following we proceed quite along the same lines as Minkowski's original result.

To every $Q \in \mathcal{P}_U\mathcal{V}(S)$ we associate the functional $\Phi_Q =: \mathbb{I}^S \to \mathbb{I}$ defined by

$$\Phi_Q(\gamma) = \inf_{\mu\in Q}\langle\mu,\gamma\rangle$$

As $\mu \mapsto \langle\mu,\gamma\rangle$ is lower semicontinuous this function attains its minimum on the compact set Q so that we can write

$$\Phi_Q(\gamma) = \min_{\mu\in Q}\langle\mu,\gamma\rangle$$

Being the (pointwise) infimum of linear functionals, Φ_Q is superlinear and, hence, order preserving. Moreover it has the following property:

$(*)\quad G(r\gamma + t\mathbf{1}) \leq rG(\gamma) + t \quad$ for all $\beta,\gamma \in \mathbb{I}^S$ and $r,t \in \mathbb{I}$ with $r + t \leq 1$

since $\Phi_Q(r\gamma+t\mathbf{1}) = \min_{\mu\in Q}(r\langle\gamma,\mu\rangle+t\langle\mathbf{1},\mu\rangle) \leq \min_{\mu\in Q} r\langle\gamma,\mu\rangle+t = \Phi_Q(\gamma)+t.$

Lemma 18 *The functional Φ_Q is Scott-continuous.*

PROOF. We show that Φ_Q is lower semicontinuous. Fix γ and consider any r such that $\Phi_Q(\gamma) > r$. Then, $\langle\mu,\gamma\rangle > r$ for every $\mu \in Q$. As $(\mu,\gamma) \mapsto \langle\mu,\gamma\rangle$ is lower semicontinuous, for every $\mu \in Q$, there are neighborhoods U_μ of γ and V_μ of μ such that $\langle\nu,\beta\rangle > r$ for all $\beta \in U_\mu$ and all $\nu \in V_\mu$. As Q is compact, it is covered by finitely many $V_{\mu_1},\ldots,V_{\mu_n}$. For the neighborhood $U = U_{\mu_1} \cap \cdots \cap U_{\mu_n}$ of γ we then have $\langle\mu,\beta\rangle > r$ for all $\beta \in U$ and all $\mu \in Q$, whence $\Phi_Q(\beta) = \min_{\mu\in Q}\langle\mu,\beta\rangle > r$ for all $\beta \in U$.

Conversely, for a Scott-continuous superlinear functional $G\colon \mathbb{I}^S \to \mathbb{I}$ satisfying $(*)$ let

$$\Psi_G = \{\mu \in \mathcal{V}(S) \mid \langle\mu,\gamma\rangle \geq G(\gamma) \text{ for all } \gamma \in \mathbb{I}^S\}$$

Lemma 19 *Ψ_G is a compact, convex and saturated subset of $\mathcal{V}(S)$.*

PROOF. Clearly, Ψ_G is saturated. As G is superlinear, Ψ_G is convex. For the compactness, it suffices to prove that Ψ_G is closed in $\mathcal{V}(S)$. Thus, take a net μ_i

in Ψ_G which converges to some $\mu \in \mathcal{V}(S)$, that is, $\mu(s) = \lim_i \mu_i(s)$ for every s. We want to show that $\langle \mu, \varphi \rangle \geq G(\varphi)$ for all $\varphi \in \mathbb{I}^S$. Consider first elements φ with finite support, i.e., $\varphi = \sum_{s \in F} r_s \eta(s)$ for a finite subset F of S and $r_s \in \mathbb{I}$. Then $\langle \mu, \varphi \rangle = \sum_{s \in F} \mu(s) r_s = \lim_i \sum_{s \in F} \mu_i(s) r_s = \lim_i \langle \mu_i, \varphi \rangle \geq G(\varphi)$. An arbitrary element $\varphi \in \mathbb{I}^S$ can be represented as the pointwise supremum of the directed family $\varphi|_F$ of its restrictions to finite subsets: $\varphi|_F = \sum_{s \in F} \varphi(s) \eta(s)$. As the functions G and $\varphi \mapsto \langle \mu, \varphi \rangle$ are Scott-continuous, we obtain $\langle \mu, \varphi \rangle = \langle \mu, \sup_F \varphi|_F \rangle = \sup_F \langle \mu, \varphi|_F \rangle \geq \sup_F G(\varphi|_F) = G(\sup_F \varphi|_F) = G(\varphi)$.

We are now ready for the main result of this Appendix. In its proof we will use the following two standard Hahn-Banach separation theorems (a convenient reference is e.g. [DS, Theorem V.2.8 ff.]):

(a) If A is a closed convex and B an open convex subset disjoint from A in a topological vector space V then there exists a continuous linear functional $f\colon V \to \mathbb{R}$ and a real number r such that $f(a) \leq r$ for all $a \in A$ and $r < f(b)$ for all $b \in B$.
(b) If A is a closed convex subset of a locally convex topological vector space V then for every $b \in V \setminus A$ there is a continuous linear functional $f\colon V \to \mathbb{R}$ such that $f(b) < f(a)$ for all $a \in A$.

Proposition 20 *$Q \mapsto \Phi_Q$ and $G \mapsto \Psi_G$ are mutually inverse order isomorphisms between the collection $\mathcal{P}_U\mathcal{V}(S)$ of all nonempty compact convex upper subsets of $\mathcal{V}(S)$ and the set $\mathcal{G}(S)$ of all Scott-continuous superlinear functionals $G\colon \mathbb{I}^S \to \mathbb{I}$ satisfying condition $(*)$.*

PROOF. We first prove that $Q = \Psi(\Phi_Q)$ for every $Q \in \mathcal{P}_U\mathcal{V}(S)$. Clearly, $Q \subseteq \Psi(\Phi_Q)$. For the converse inclusion suppose that $\nu \notin Q$. The upper set $\uparrow Q$ generated by Q in ℓ^1 is closed and convex by Lemma 14. By the above mentioned Hahn-Banach separation theorem (b), there is a continuous linear functional f on ℓ^1 such that $f(\nu) < f(\mu)$ for all $\mu \in \uparrow Q$.

We now show that f maps ℓ^1_+ to $\mathbb{R}_+$. We choose a fixed $\mu \in Q$. For every $s \in S$ we have $\mu + r\eta(s) \in \uparrow Q$ and consequently $f(\nu) < f(\mu + r\eta(s)) = f(\mu) + r f(\eta(s))$ for all $r \geq 0$ which implies $f(\eta(s)) \geq 0$. By linearity it follows that $f\Big(\sum_{s \in F} r_s \eta(s)\Big) \geq 0$ for every finite subset F of S and $r_s \geq 0$ for $s \in F$. By continuity of f we conclude that $f(\mu) \geq 0$ for all $\mu \in \ell^1_+$. As every continuous linear functional on ℓ^1_+ is order preserving and lower semicontinuous, hence Scott-continuous, Lemma 16 tells us that there is a $\gamma \in \ell^\infty_+$ such that $f(\mu) = \langle \mu, \gamma \rangle$ for all $\mu \in \ell^1_+$. Replacing γ by $\frac{1}{m}\gamma$ (for a sufficiently big $m \in \mathbb{N}$) we may suppose that $\gamma \in \mathbb{I}^S$ and we have $\langle \nu, \gamma \rangle < \langle \mu, \gamma \rangle$ for all $\mu \in Q$ and thus $\nu \notin \Psi(\Phi_Q)$ as desired.

Now suppose $G\colon \mathbb{I}^S \to \mathbb{I}$ is a Scott-continuous superlinear functional satisfying condition $(*)$. We will show that $G = \Phi(\Psi_G)$, i.e. $G(\gamma) = \inf_{\mu\in\Psi_G}\langle\mu,\gamma\rangle = \inf\{\langle\mu,\gamma\rangle \mid \mu \in \mathcal{V}(S), G \leq \langle\mu,-\rangle \text{ on } \mathbb{I}^S\}$ for all $\gamma \in \mathbb{I}^S$. This obviously holds if G is constantly 0. Thus, let us assume w.l.o.g. that G is not constantly 0. As by Corollary 17 the elements $\mu \in \mathcal{V}(S)$ are in a one-to-one correspondence with the Scott-continuous linear functionals $f\colon \mathbb{I}^S \to \mathbb{I}$, we have to show that, for all $\gamma \in \mathbb{I}^S$

$$G(\gamma) = \inf\{f(\gamma) \mid f\colon \mathbb{I}^S \to \mathbb{I} \text{ linear, Scott-continuous, and } G \leq f \text{ on } \mathbb{I}^S\}$$

For the proof we fix a $\gamma \in \mathbb{I}^S$. We will show that every real number $r > 0$ with $G(\gamma) \leq r \leq 1$ there is a Scott-continuous linear functional $f\colon \mathbb{I}^S \to \mathbb{I}$ such that $G \leq f$ on $\mathbb{I}^S$ and $f(\gamma) \leq r$.

For this purpose we consider the unique homogeneous extension $\widehat{G}\colon \ell^\infty_+ \to \mathbb{R}_+$ of G. The extended functional $\widehat{G}$ is Scott-continuous, superlinear and satisfies the inequality $(*)$. We search for a linear functional $f\colon \ell^\infty_+ \to \mathbb{R}_+$ such that $\widehat{G} \leq f$, $f(\gamma) \leq r$ and $f(\mathbb{I}^S) \subseteq \mathbb{I}$. Replacing γ by $\frac{1}{r}\gamma$, we may suppose that $r = 1$.

We now form the line segment $A = \{(1-t)\gamma + t\mathbf{1} \mid t \in \mathbb{I}\}$ connecting γ and $\mathbf{1}$, which is closed and convex, and the set $U = \{\beta \in \ell^\infty_+ \mid \widehat{G}(\beta) > 1\}$ which is nonempty since G and thus $\widehat{G}$ is not constantly 0. As $\widehat{G}$ is lower semicontinuous, superlinear and, hence, order preserving U is a nonempty open convex upper set. Furthermore, the set U is disjoint from A : indeed, if $\beta \in A$ then $\beta = (1-t)\gamma + t\mathbf{1}$ for some $t \in \mathbb{I}$ for which it holds that

$$\begin{aligned}
\widehat{G}(\beta) &= \widehat{G}\Big((1-t)\gamma + t\cdot\mathbf{1}\Big) \\
&\leq (1-t)\widehat{G}(\gamma) + t \qquad \text{by condition } (*) \\
&\leq 1 \qquad\qquad\qquad\quad \text{as } \widehat{G}(\gamma) \leq 1
\end{aligned}$$

whence $\beta \notin U$. In ℓ^∞ we may apply the Hahn-Banach separation theorem (a) cited above for disjoint open and closed convex sets and separation theorem (a) cited above for disjoint open and closed convex sets and obtain a continuous linear functional f on ℓ^∞ and a real number s such that $f(\beta) \leq s$ for all $\beta \in A$ and $f(\beta) > s$ for all $\beta \in U$. As in the first part of this proof one shows that $f(\beta) \geq 0$ for every $\beta \geq 0$. Thus, the functional f is also monotonic on ℓ^∞_+. For $\beta \in \mathbb{I}^S$ we have $\beta \leq \mathbf{1} \in A$, whence $f(\beta) \leq f(\mathbf{1}) \leq s$. Since $0 \leq f(\mathbf{1}) \leq s$ we have $0 \leq s$. If s were 0 then f would be constantly 0 on $\mathbb{I}^S$ and thus also constantly 0 on ℓ^∞_+ which is impossible since U is nonempty. Thus, we have shown that $s > 0$ and w.l.o.g. we may assume that $s = 1$. Finally we show that $G(\beta) \leq f(\beta)$ for all $\beta \in \mathbb{I}^S$. We proceed by showing something stronger, namely that $\widehat{G}(\beta) \leq f(\beta)$ for all $\beta \in \ell^\infty_+$. Indeed, whenever $\widehat{G}(\beta) > \varepsilon > 0$ then $\widehat{G}(\frac{\beta}{\varepsilon}) > 1$, whence $\frac{\beta}{\varepsilon} \in U$ and consequently $f(\frac{\beta}{\varepsilon}) > 1$, that is $f(\beta) > \varepsilon$.

Thus, we have also shown the existence of a linear $f : \mathbb{I}^S \to \mathbb{I}$ with $G \leq f$ from which it follows that Ψ_G is nonempty. For G constantly 0 we have that $\Psi_G = \mathcal{V}(S)$ and thus nonempty. From this observation together with Lemma 19 it follows that $\Psi_G \in \mathcal{P}_U\mathcal{V}(S)$ whenever $G \in \mathcal{G}(S)$.

Note. As for linear functionals, we have for superlinear functionals $G\colon \mathbb{I}^S \to \mathbb{I}$ satisfying condition $(*)$: If G is Scott-continuous then G is also upper semicontinuous, hence continuous.

Indeed, for a Scott-continuous superlinear functional $G\colon \mathbb{I}^S \to \mathbb{I}$ satisfying condition $(*)$ we have by Proposition 20 that $G(\gamma) = \inf_{\mu\in\Psi_G}\langle\mu, \gamma\rangle$. Thus G is the (pointwise) infimum of the set of linear functionals $\gamma \mapsto \langle\mu, \gamma\rangle$ where $\mu \in \Psi_G$. As these functionals are continuous their infimum is upper semicontinuous.

Note. For understanding condition (*) it might be helpful to notice that it is equivalent to

$$(**) \qquad G(r\gamma + t\beta) \leq rG(\gamma) + t|\beta|_\infty$$

where $|\beta|_\infty = \sup_s \beta(s)$ is the sup-norm of β; for $r = 0, t = 1$ this implies $G(\beta) \leq |\beta|_\infty$, that is, G is dominated by the sup-norm functional. (Indeed, $(**)$ implies $(*)$ by considering the special case $\beta = \mathbf{1}$. Conversely, as $\beta \leq |\beta|_\infty\mathbf{1}$, we have $G(r\gamma + t\beta) \leq G(r\gamma + t|\beta|_\infty\mathbf{1}) \leq rG(\gamma) + t|\beta|_\infty$ by $(*)$.)

References

[BHM] N. Benton, J. Hughes, E. Moggi *Monads and Effects* pp.42-122 in *Applied Semantics* Lecture Notes in Computer Science **2395**, Springer Verlag (2002).

[DS] N. Dunford, J. T. Schwartz *Linear Operators, Part I: General Theory.* Interscience Publ., New York (1967).

[GHK+] G. Gierz, K. H. Hofmann, J. D. Lawson, M. W. Mislove, D. S. Scott *Continuous Lattices and Domains.* Cambridge University Press (2003).

[Hec] R. Heckmann, Power domains and second-order predicates, *Theoretical Computer Science*, **111**, 59–88, 1993.

[KP] K. Keimel, G. Plotkin *Predicate Transformers for Convex Powerdomains* submitted (2007).

[MM] A. McIver, C. Morgan *Abstraction, refinement and proof for probabilistic systems.* Monographs in Computer Science, Springer (2005).

[MMa] McIver, A. and Morgan, C. Demonic, angelic and unbounded probabilistic choices in sequential programs. *Acta Informatica* **37**, 329–354 (2001).

[MMb] McIver, A. and Morgan, C. Partial correctness for probabilistic demonic programs. *Theoretical Computer Science* 266, 513–541 (2001).

[Min] H. Minkowski Volumen und Oberfläche. *Mathematische Annalen* **57**, 447–495 (1903).

[Plo] G. Plotkin *Domain Theory.* Lecture Notes available from the author's homepage.

[Smy] M. B. Smyth Power domains and predicate transformers: a topological view, *Proc. 10th ICALP* (ed. J. Díaz), Lecture Notes in Computer Science, **154**, 662–675, Springer-Verlag (1983).

[Str] T. Streicher *Domain-theoretic Foundations of Functional Programming* World Scientific (2006).

[TKP] R. Tix, K. Keimel, G. Plotkin *Semantic domains for combining probability and nondeterminism.* Electronic Notes in Theoretical Computer Science **129**, Elsevier (2005).

[Tol] A. A. Tolstogonov, *Support functions of convex compacta*, (Russian) Matematicheskie Zametki **22**, 203–213 (1977). English translation in: Mathematical Notes **22**, 604–609, (1977).