Continuous Previsions*

Jean Goubault-Larrecq

LSV, ENS Cachan, CNRS, INRIA Futurs
61, avenue du président-Wilson, 94235 Cachan, France
goubault@lsv.ens-cachan.fr

Abstract. We define strong monads of *continuous (lower, upper) previsions*, and of *forks*, modeling both probabilistic and non-deterministic choice. This is an elegant alternative to recent proposals by Mislove, Tix, Keimel, and Plotkin. We show that our monads are sound and complete, in the sense that they model exactly the interaction between probabilistic and (demonic, angelic, chaotic) choice.

1 Introduction

Moggi's computational λ-calculus [17] has proved useful to define various notions of computations on top of the lambda-calculus: side-effects, input-output, continuations, non-determinism [27], probabilistic computation [20] in particular. But mixing monads is hard, and finding the "right" monad that would combine both non-determinism and probabilistic choice has taken quite some effort. (We review recent progress below.)

The purpose of this paper is to introduce simple monads that do the job well. These are monads of *continuous previsions*, which can be seen as continuation-style monads. The idea of considering previsions comes from economics and statistics [4, 12].

Outline. After stating some required preliminaries in Section 2, we recall the notion of *game* introduced in [5], arguing why these are natural extensions of notions of continuous valuations ($\sim$ measures) that also accommodate demonic and angelic non-deterministic choice. These notions induce functors on $\mathbf{Top}$, $\mathbf{Cpo}$, $\mathbf{Pcpo}$ (*pointed* cpos), but fail to yield monads. We analyze this failure in Section 4 by moving, through a Riesz-like representation theorem, to the new notions of collinear previsions, and previsions. We then show that indeed previsions yield strong monads, giving a simple semantics to a rich λ-calculus [17] with both probabilistic and non-deterministic choice. Finally, we show in Section 5 that our monad model is not only sound but complete.

This work is a summary of most of Chapters 10-12 of [6].

Related Work. Finding a monad combining both probabilistic and non-deterministic choice can be done by using general monad combination principles. The right way to combine monads in general is open to discussion. Lüth [11] proposes to combine monads by taking their coproduct in the category of monads. This coproduct exists under relatively mild assumptions [10]. However, in general the coproduct of two monads is an inscrutable object. A simpler, explicit description can be found in specific cases. For example, when taking coproducts of two *ideal* monads [2]. In particular, combining

* Partially supported by the INRIA ARC ProNoBis. Part of Section 4 was done while the author was invited at U. Laval, Québec City, Québec, July 2004. We acknowledge their support.

non-blocking non-determinism and probabilistic choice falls into this case. The resulting monad is relatively unenlightening, though: it is the monad of all sequences of choices, both probabilistic and non-deterministic [2, exemple 4.3].

Varacca [25, 26] also proposed a monad combining non-determinism with probabilistic choice. Ghani and Uustalu [2] note that the above coproduct monad is close to Varacca's synchronization trees. The works closer to ours in computer science are those of Mislove [16] and Tix [23, 24]. While this won't be entirely obvious from our definitions, we will establish a formal connection between their models and ours (Section 5). Outside computer science, previsions have their roots in economics and statistics [28]. However, we consider previsions on topological spaces, not just on sets.

This paper can be seen also be seen as a followup to [5], inasmuch as previsions are strongly tied to notions of convex and concave games. Our previsions can also be seen as predicate transformers; as such, and modulo minor details, they were discovered independently by Keimel and Plotkin [9] (K. Keimel, personal communication).

Acknowledgment. Thanks to Nadia Tawbi for inviting me in July 2004 in Québec City. Thanks to Vincent Danos, Josée Desharnais and François Laviolette, through whom I came to study related problems, namely [5]. Thanks to Prakash Panangaden for his warm support, to all people who showed interest in this work [6], to Michael Mislove, and to Klaus Keimel.

2 Preliminaries

We assume the reader to be familiar with (point-set) topology, in particular topology of T_0 but not necessarily Hausdorff spaces, as encountered in domain theory. See [3, 1, 15] for background. Let $int(A)$ denote the interior of A, $cl(A)$ its closure. The *Scott topology* on a poset X, with ordering $\leq$, has as opens the upward-closed subsets U (i.e., $x \in U$ and $x \leq y$ imply $y \in U$) such that for every directed family $(x_i)_{i \in I}$ having a least upper bound $\sup_{i \in I} x_i$ inside U, some x_i is already in U. The *way-below* relation $\ll$ is defined by $x \ll y$ iff for any directed family $(z_i)_{i \in I}$ with a least upper bound z such that $y \leq z$, then $x \leq z_i$ for some $i \in I$. A poset is *continuous* iff $\twoheaddownarrow y = \{x \in X | x \ll y\}$ is directed, and has y as least upper bound. Then every open U can be written $\bigcup_{x \in U} \twoheaduparrow x$, where $\twoheaduparrow x = \{y \in X | x \ll y\}$, and each of these sets is open.

Every topological space X has a specialization quasi-ordering $\leq$, defined by: $x \leq y$ iff every open that contains x contains y. X is T_0 iff $\leq$ is a (partial) ordering. The specialization ordering of the Scott topology of a quasi-ordering $\leq$ is $\leq$ itself. Every open is upward-closed, and a subset $A \subseteq X$ is *saturated* if and only if A is the intersection of all opens that contain it; equivalently, iff A is upward-closed in $\leq$. Let $\uparrow A$ denote the upward-closure of A under a quasi-ordering $\leq$, $\downarrow A$ its downward-closure. A T_0 space is *sober* iff every irreducible closed subset is the closure $cl\{x\} = \downarrow x$ of a (unique) point x. The Hofmann-Mislove Theorem implies that every sober space is *well-filtered* [8]: given any filtered family of saturated compacts $(Q_i)_{i \in I}$ in X, and any open U, $\bigcap_{i \in I} Q_i \subseteq U$ iff $Q_i \subseteq U$ for some $i \in I$. In particular, $\bigcap_{i \in I} Q_i$ is saturated compact. X is *locally compact* iff whenever $x \in U$ (U open) there is a saturated compact Q such that $x \in int(Q) \subseteq Q \subseteq U$. Every continuous cpo is sober and locally compact in its Scott topology. X is *coherent* iff the intersection of any two saturated compacts is compact. (Some authors take X to be coherent iff finite intersections of saturated compacts are compact. This would imply compactness, which we don't assume.) A coherent, well-filtered locally compact space is called *stably locally compact*. *Stably compact* spaces are those that are additionally compact, and have a wonderful theory (see, e.g., [8]). We consider the space $\mathbb{R}$ of all reals with the Scott topology of its natural ordering $\leq$. Its opens are $\emptyset$, $\mathbb{R}$, and the intervals $(t, +\infty)$, $t \in \mathbb{R}$. Any closed interval of $\mathbb{R}$, e.g., $[0, 1]$, is a stably locally compact, continuous cpo with the Scott topology. Because we equip $\mathbb{R}$ with the Scott topology, our *continuous* functions $f : X \to \mathbb{R}$ are those usually called *lower semi-continuous* in the mathematical literature.

By a *capacity* on X, we mean any function ν from $\mathcal{O}(X)$, the set of all opens of X, to $\mathbb{R}^+$, such that $\nu(\emptyset) = 0$ (a.k.a., a *set function*). A *game* ν is a *monotonic capacity*, i.e., $U \subseteq V$ implies $\nu(U) \leq \nu(V)$[1]. A *valuation* is a *modular* game ν, i.e., one such that $\nu(U \cup V) + \nu(U \cap V) = \nu(U) + \nu(V)$ for every opens U, V. A game is *continuous* iff $\nu(\bigcup_{i \in I} U_i) = \sup_{i \in I} \nu(U_i)$ for every directed family $(U_i)_{i \in I}$ of opens, and *normalized* iff $\nu(X) = 1$. Continuous valuations have a nice theory that fits topology well [7, 8].

The *Dirac valuation* δ_x at $x \in X$ is the continuous valuation mapping each open U to 1 if $x \in U$, to 0 otherwise. Continuous valuations are canonically ordered by $\nu \leq \nu'$ iff $\nu(U) \leq \nu'(U)$ for every open U of X.

A *monad* on a category $\mathbb{C}$ may be presented in several different ways. One is based on triples $(\boldsymbol{T}, \boldsymbol{\eta}, \boldsymbol{\mu})$ of an endofunctor on $\mathbb{C}$, a unit, and a multiplication natural transformation. A presentation that is easier to grasp is in terms of Kleisli triples [14]. A *Kleisli triple* is a triple $(\boldsymbol{T}, \boldsymbol{\eta}, \_^\dagger)$, where $\boldsymbol{T}$ maps objects X of $\mathbb{C}$ to objects $\boldsymbol{T}X$ of $\mathbb{C}$, $\boldsymbol{\eta}_X$ is a morphism from X to $\boldsymbol{T}X$ for each X, and $f^\dagger$ (the *extension* of f) is a morphism from $\boldsymbol{T}X$ to $\boldsymbol{T}Y$ for each morphism $f : X \to \boldsymbol{T}Y$, satisfying: (1) ${\boldsymbol{\eta}_X}^\dagger = \mathrm{id}_{\boldsymbol{T}X}$; (2) for every $f : X \to \boldsymbol{T}Y$, $f^\dagger \circ \boldsymbol{\eta}_X = f$; (3) for every $g : X \to \boldsymbol{T}Y$, $f : Y \to \boldsymbol{T}Z$, then $f^\dagger \circ g^\dagger = (f^\dagger \circ g)^\dagger$. Kleisli triples and monads are equivalent.

3 Continuous Games, Convexity, Concavity

We follow [5]. A game ν on X is *convex* iff $\nu(U \cup V) + \nu(U \cap V) \geq \nu(U) + \nu(V)$ for every opens U, V. It is *concave* if the opposite inequality holds. Convex games are a cornerstone of economic theory [4, 18].

One fundamental example of a game that is not a valuation is the *unanimity game* $\mathfrak{u}_A$ ($A \neq \emptyset$), defined by $\mathfrak{u}_A(U) = 1$ if $A \subseteq U$, $\mathfrak{u}_A(U) = 0$ otherwise. As we argue in [5], $\mathfrak{u}_A$ is a natural "probability-like" description of demonic non-deterministic choice, in the sense that drawing "at random" according to $\mathfrak{u}_A$ means that some malicious adversary C will choose an element of A for you. This is perhaps best conveyed by a thought experiment. You, the honest player P, would like to draw some element x from X with distribution ν (a game). Imagine you would like to know your chances of getting one from some (open) subset U of X. If ν is a probability distribution, then your chances will be equal to $\nu(U)$. This is standard. For general ν, continue to define

[1] The name "game" is unfortunate, as there is no obvious relationship between this and games as they are usually handled in computer science, in particular with stochastic games. The notion stems from (cooperative) games in economics, where X is the set of players, not of states.

your chances as $\nu(U)$. If $\nu = \mathfrak{u}_A$, and U does not contain A, then $\nu(U) = 0$, and your chances are zero: intuitively, C will pick an element in A, but outside U—on purpose. The only case where C is forced to pick an element in A which will suit P (i.e., be in U, too), is when $A \subseteq U$—and then P will be pleased with probability one.

It is clear that $\mathfrak{u}_A$ is convex. It is in fact more. Call a game ν *totally convex* iff:

$$\nu\left(\bigcup_{i=1}^{n} U_i\right) \geq \sum_{I \subseteq \{1,\ldots,n\}, I \neq \emptyset} (-1)^{|I|+1} \nu\left(\bigcap_{i \in I} U_i\right) \tag{1}$$

for every finite family $(U_i)_{i=1}^n$ of opens ($n \geq 1$), where $|I|$ denotes the cardinality of I. A *belief function* is a totally convex game. The dual notion of *total concavity* is obtained by replacing $\bigcup$ by $\bigcap$ and conversely in (1), and turning $\geq$ into $\leq$. A *plausibility* is a totally concave game. If $\geq$ is replaced by $=$ in (1), then we retrieve the familiar inclusion-exclusion principle from statistics. In particular any (continuous) valuation is a (continuous) belief function. Any belief function is a convex game, but not conversely [4, 5].

Every game of the form $\sum_{i=1}^n a_i \mathfrak{u}_{Q_i}$, with $a_i \in \mathbb{R}^+$, and Q_i compact saturated and non-empty, is a continuous belief function, which we call *simple* belief function in [5]. When $\sum_{i=1}^n a_i = 1$, drawing an element from X "at random" (in the sense illustrated above) according to the simple belief function $\nu = \sum_{i=1}^n a_i \mathfrak{u}_{Q_i}$ intuitively corresponds to drawing one compact Q_i at random with probability a_i, then to let the malicious adversary C draw some element, demonically, from Q_i [5].

Let us turn to integration. Let ν be a game on X, and f be continuous from X to $\mathbb{R}^+$ (i.e., lower semi-continuous: $\mathbb{R}^+$ comes with the Scott topology). Assume f bounded, too, i.e., $\sup_{x \in X} f(x) < +\infty$. The *Choquet integral* of f along ν is:

$$\oint_{x \in X} f(x) d\nu = \int_0^{+\infty} \nu(f^{-1}(t, +\infty)) dt \tag{2}$$

where the right hand side is an improper Riemann integral. This is well-defined, since $f^{-1}(t, +\infty)$ is open for every $t \in \mathbb{R}^+$ by assumption, and ν measures opens. Also, since f is bounded, the improper integrals above really are ordinary Riemann integrals over some closed intervals. The function $t \mapsto \nu(f^{-1}(t, +\infty))$ is decreasing, and every decreasing (even non-continuous, in the usual sense) function is Riemann-integrable, therefore the definition makes sense.

Alternatively, any *step function* $\sum_{i=0}^n a_i \chi_{U_i}$, where $a_0 \in \mathbb{R}^+$, $a_1, \ldots, a_n \in \mathbb{R}^+$, $X = U_0 \supseteq U_1 \supseteq \ldots \supseteq U_n$ is a decreasing sequence of opens, and χ_U denotes the indicator function of U ($\chi_U(x) = 1$ if $x \in X$, $\chi_U(x) = 0$ otherwise) is continuous: its integral along ν then equals $\sum_{i=0}^n a_i \nu(U_i)$—for *any* game ν. It is well-known that every bounded continuous function f can be written as the least upper bound of a sequence of step functions $f_K = a + \frac{1}{2^K} \sum_{k=1}^{\lfloor (b-a)2^K \rfloor} \chi_{f^{-1}(a + \frac{k}{2^K}, +\infty)}(x)$, $K \in \mathbb{N}$, where $a = \inf_{x \in X} f(x)$, $b = \sup_{x \in X} f(x)$. Then the integral of f along ν is the least upper bound of the increasing sequence of the integrals of f_K along ν.

The main properties of Choquet integration are as follows. First, the integral is increasing in its function argument: if $f \leq g$ then the integral of f along ν is less than or equal to that of g along ν. If ν is continuous, then integration is also Scott-continuous in its function argument. The integral is also monotonic and Scott-continuous in the game ν. Integration is linear in the game, too, so integrating along $\sum_{i=1}^{n} a_i \nu_i$ is the same as taking the integrals along each ν_i, and computing the obvious linear combination. However, Choquet integration is *not* linear in the function integrated, unless the game ν is a valuation. Still, it is *positively homogeneous*: integrating αf for $\alpha \in \mathbb{R}^+$ yields α times the integral of f. And it is additive on *comonotonic* functions $f, g : X \to \mathbb{R}$ (i.e., there is no pair $x, x' \in X$ such that $f(x) < f(x')$ and $g(x) > g(x')$).

Returning to the example of a simple belief function $\nu = \sum_{i=1}^{n} a_i \mathfrak{u}_{Q_i}$, the properties above imply that the integral of f along ν is $\sum_{i=1}^{n} a_i \min_{x \in Q_i} f(x)$ [5, Proposition 1]. (Note that $f(x)$ indeed attains its minimum over Q_i, which is compact.) Another way to read this is as follows. Imagine P publishes how much money, $f(x)$, she would earn if you picked x. When $\sum_{i=1}^{n} a_i = 1$, it is legitimate to say that the integral of f along ν should be some form of expected income. The formula above states that, when ν is a simple belief function, your expected income is exactly what you would obtain on average by drawing Q_i at random with probability a_i, then letting the malicious adversary C pick some element of Q_i for you—minimizing your earnings $f(x)$. In other words, integrating along a simple belief function computes *average min-payoffs*.

This can be generalized to all continuous, not just simple, belief functions [5, Theorem 4]. More precisely, the space $\mathbf{Cd}_{\leq 1}(X)$ of all continuous belief functions ν on X such that $\nu(X) \leq 1$ is isomorphic to the space $\mathbf{V}_{\leq 1}(\mathcal{Q}(X))$ of continuous valuations ν^* (of total mass at most 1) over the *Smyth powerdomain* $\mathcal{Q}(X)$ of X, provided X is well-filtered and locally compact. $\mathcal{Q}(X)$ is the cpo of non-empty compact saturated subsets of X, ordered by reverse inclusion $\supseteq$, and is a model of demonic non-determinism. (A similar result holds for *normalized* games and valuations ν, i.e., such that $\nu(X) = 1$: $\nu \mapsto \nu^*$ is again an isomorphism from $\mathbf{Cd}_1(X)$ to $\mathbf{V}_1(\mathcal{Q}(X))$.) The construction of ν^* from ν is relatively difficult, however it is noteworthy that when $\nu = \sum_{i=1}^{n} a_i \mathfrak{u}_{Q_i}$ is simple, then ν^* is exactly the simple valuation $\sum_{i=1}^{n} a_i \delta_{Q_i}$, which describes the choice of an element Q_i at random with probability a_i, as intuition would have it.

Similarly, the space $\mathbf{Pb}_{\leq 1}(X)$ of all *continuous plausibilities* ν with $\nu(X) \leq 1$ is isomorphic to $\mathbf{V}_{\leq 1}(\mathcal{H}_u(X))$ when X is stably locally compact, and where $\mathcal{H}_u(X)$ is the topological Hoare powerspace, defined as the set of non-empty closed sets of X, with topology generated by the sub-basic open sets $\Diamond U = \{F \text{ closed} | F \cap U \neq \emptyset\}$, U open in X. (This is the upper topology of $\subseteq$, which in general does not coincide with the Scott topology, unless e.g. X is a continuous cpo.) $\mathcal{H}_u(X)$ is used to model angelic non-determinism. We do not develop this here (see [6]). However, we mention that the corresponding *simple* plausibilities are of the form $\sum_{i=1}^{n} a_i \mathfrak{e}_{F_i}$, where F_i is a non-empty closed subset of X (an element of $\mathcal{H}_u(X)$), and the *example game* $\mathfrak{e}_F$ is defined so that $\mathfrak{e}_F(U) = 1$ if F meets U, $\mathfrak{e}_F(U) = 0$ otherwise: in this case C tries to help you, by finding some element in U that would also be in F, if possible.

Recall that every belief function is convex. One may show that Choquet integration along ν is *super-additive* (the integral of $f + g$ is at least that of f plus that of g) when ν is convex, and *sub-additive* (the integral of $f + g$ is at most that of f plus that of g) when ν is concave. See [4] for the finite case, [6, chapitre 4] for the topological case.

In the sequel, let $\mathbf{J}(X)$, $\bigtriangledown \mathbf{J}(X)$, $\bigtriangleup \mathbf{J}(X)$ be the spaces of plain, convex and concave continuous games respectively ("plain" meaning with no added property).

4 Continuous Previsions

For any space X, let $\langle X \to \mathbb{R}^+\rangle$ be the space of all bounded continuous functions from X to $\mathbb{R}^+$, with the Scott topology. Each continuous game ν on X gives rise to a functional $\alpha_{\mathcal{C}}(\nu)$ from $\langle X \to \mathbb{R}^+\rangle$ to $\mathbb{R}^+$, mapping f to its Choquet integral along ν.

Think of $f(x)$ again as defining how much money you would earn when x is chosen from X, by some computation process. We intentionally leave the notion of computation process undefined. This may be the process of drawing "at random" along a game ν, as in Section 3. In the sequel, we shall explore the view that x is the output of an arbitrary program, defined in some non-deterministic and probabilitic functional language. I.e., any program returns a value x ($\bot$ on non-termination, say), and if so P earns $f(x)$. For purely probabilistic programs (no non-deterministic choice), a prevision F is essentially a function mapping earning functions f to their average value $F(f)$, over all possible executions. Slightly more generally, for any belief function ν, there is a prevision $\alpha_{\mathcal{C}}(\nu)$ that maps each $f \in \langle X \to \mathbb{R}^+\rangle$ to the average min-payoff we get when our final earnings are given by f. Milking out the properties of $\alpha_{\mathcal{C}}(\nu)$, we arrive at:

Definition 1 (Prevision). *A* prevision *is a functional F from $\langle X \to \mathbb{R}^+\rangle$ to $\mathbb{R}^+$ such that F is* positively homogeneous *(for every $\alpha \geq 0$, $F(\alpha f) = \alpha F(f)$), and* monotonic *(if $f \leq g$ [pointwise], then $F(f) \leq F(g)$).*

F is a lower prevision *if moreover F is* super-additive, *i.e., $F(f+g) \geq F(f)+F(g)$. F is an* upper prevision *iff F is* sub-additive*: $F(f+g) \leq F(f)+F(g)$. F is* collinear *iff F is additive on comonotonic pairs, i.e., if whenever f and g are comonotonic, then $F(f+g) = F(f)+F(g)$. A prevision F is* linear *iff $F(f+g) = F(f)+F(g)$ for every $f, g \in \langle X \to \mathbb{R}^+\rangle$.*

Finally, F is continuous *iff it is Scott-continuous: for every directed family $(f_i)_{i \in I}$ of bounded continuous functions with least upper bound f, $F(\sup_{i \in I} f_i) = \sup_{i \in I} F(f_i)$.*

We write $\mathbf{P}(X)$, $\bigtriangledown \mathbf{P}(X)$, $\bigtriangleup \mathbf{P}(X)$ respectively the spaces of all continuous previsions, of continuous lower previsions, of continuous upper previsions equipped with the Scott topology of the pointwise ordering $\leq$. The spaces $\mathbf{P}^{\divideontimes}(X)$, $\bigtriangledown \mathbf{P}^{\divideontimes}(X)$, $\bigtriangleup \mathbf{P}^{\divideontimes}(X)$ will be the subspaces of those that are collinear.

We do not quite follow standard naming conventions. Standardly [28], a lower prevision is just a real-valued functional. *Coherent* lower previsions (taking a more readable definition from [13]) are those F such that $F(f) \geq \sum_{i=1}^n \lambda_i F(f_i) + \lambda_0$ whenever $f \geq \sum_{i=1}^n \lambda_i f_i + \lambda_0$, $\lambda_i > 0$, $\lambda_0 \in \mathbb{R}$. In our case, we reserve the "lower" adjective, so as to have a dual notion of *upper* prevision.

It is clear that any continuous game ν defines a continuous collinear prevision $\alpha_{\mathcal{C}}(\nu)$. Moreover, if ν is convex, then $\alpha_{\mathcal{C}}(\nu)$ is lower, and if ν is concave, then $\alpha_{\mathcal{C}}(\nu)$ is upper. The following isomorphism result, akin to Riesz' Representation Theorem, is known as Schmeidler's Theorem for convex games on discrete topologies. Let $\gamma_{\mathcal{C}}(F)$, for any prevision F, be the capacity ν such that $\nu(U) = F(\chi_U)$ for every open U of X. Order previsions pointwise, then:

Theorem 1. *$\alpha_{\mathcal{C}} \dashv \gamma_{\mathcal{C}}$ is a Galois connection from (plain, convex, concave) games into (plain, lower, upper) collinear previsions, with $\alpha_{\mathcal{C}}$ injective. That is, $\alpha_{\mathcal{C}}$ and $\gamma_{\mathcal{C}}$ are monotonic, $\alpha_{\mathcal{C}}(\gamma_{\mathcal{C}}(F)) \leq F$ for every collinear prevision F, and $\gamma_{\mathcal{C}}(\alpha_{\mathcal{C}}(\nu)) = \nu$ for every game ν.*

Moreover, when restricted to continuous previsions and games, $\alpha_{\mathcal{C}}$ and $\gamma_{\mathcal{C}}$ define an isomorphism between $\mathbf{J}(X)$ and $\mathbf{P}^{\divideontimes}(X)$, between $\bigtriangledown \mathbf{J}(X)$ and $\bigtriangledown \mathbf{P}^{\divideontimes}(X)$, between $\bigtriangleup \mathbf{J}(X)$ and $\bigtriangleup \mathbf{P}^{\divideontimes}(X)$.

Proof. That $\gamma_{\mathcal{C}}(F)$ is a game for any prevision is easy. When F is lower, note that $\chi_{U \cup V}$ and $\chi_{U \cap V}$ are comonotonic, and $\chi_{U \cup V} + \chi_{U \cap V} = \chi_U + \chi_V$. So $\gamma_{\mathcal{C}}(F)(U \cup V) + \gamma_{\mathcal{C}}(F)(U \cap V) = F(\chi_{U \cup V} + \chi_{U \cap V})$ (since F is collinear) $= F(\chi_U + \chi_V) \geq F(\chi_U) + F(\chi_V)$ (since F is super-additive) $= \gamma_{\mathcal{C}}(F)(U) + \gamma_{\mathcal{C}}(F)(V)$. Similarly, $\gamma_{\mathcal{C}}(F)$ is concave if F is upper.

For the converse, we first show that: (A) for any collinear prevision F on X, for any step function f, written $a + \sum_{i=1}^{m} a_i \chi_{U_i}$ with $U_1 \supseteq \ldots \supseteq U_m$, $a \in \mathbb{R}$, $a_1, \ldots, a_m \in \mathbb{R}^+$, then the Choquet integral of f along $\gamma_{\mathcal{C}}(F)$ equals $F(f)$. This is an easy exercise as soon as one realizes that $\sum_{i=0}^{k-1} a_i \chi_{U_i}$ and $a_k \chi_{U_k}$ are comonotonic for every k, $1 \leq k \leq m$. The equality $\gamma_{\mathcal{C}}(\alpha_{\mathcal{C}}(\nu))(U) = \nu(U)$ is obvious, $\alpha_{\mathcal{C}}$ and $\gamma_{\mathcal{C}}$ are clearly monotonic. To show that $\alpha_{\mathcal{C}}(\gamma_{\mathcal{C}}(F)) \leq F$, we must show that the Choquet integral of f along $\gamma_{\mathcal{C}}(F)$ is less than or equal to $F(f)$. Using the step functions f_K, $K \in \mathbb{N}$, by (A) the Choquet integral of f_K is less than or equal to $F(f_K)$. The least upper bound of the Choquet integrals of f_K, $K \in \mathbb{N}$ is that of f, and the least upper bound of $F(f_K)$ is at most $F(f)$. So $\alpha_{\mathcal{C}}(\gamma_{\mathcal{C}}(F))(f) \leq F(f)$. When F is continuous, the least upper bound of $F(f_K)$ is exactly $F(f)$, whence $\alpha_{\mathcal{C}}(\gamma_{\mathcal{C}}(F)) = F$. ☐

One easy, well-known consequence of this is that $\alpha_{\mathcal{C}}$ and $\gamma_{\mathcal{C}}$ define an order isomorphism between the space $\mathbf{V}(X)$ of continuous valuations and that $\mathbf{P}^{\triangle}(X)$ of continuous linear previsions ([7, Theorem 6.2], [22, Satz 4.16]). Intuitively, any continuous game ν gives rise to a continuous collinear prevision $\alpha_{\mathcal{C}}(\nu)$ that computes a generalized form of expectation along ν, and every continuous collinear prevision arises this way.

It is easy to check that $\mathbf{J}$, $\bigtriangledown \mathbf{J}$, $\bigtriangleup \mathbf{J}$, $\mathbf{V}$, $\mathbf{P}^{\divideontimes}$, $\bigtriangledown \mathbf{P}^{\divideontimes}$, $\bigtriangleup \mathbf{P}^{\divideontimes}$, $\mathbf{P}^{\triangle}$ define functors $\boldsymbol{T}$ from $\boldsymbol{Top}$ to $\boldsymbol{Top}$, where $\boldsymbol{Top}$ is the category of topological spaces.

To define a monad structure on $\boldsymbol{T}$, we need a *unit* $\boldsymbol{\eta}_X : X \to \boldsymbol{T}X$, natural in X. This is defined by $\boldsymbol{\eta}_X(x) = \delta_x$. However, there is in general no extension $f^\dagger$ of $f : X \to \boldsymbol{T}Y$. The natural candidate is:

$$f^\dagger(\nu)(V) = \oint^{\mathcal{C}}_{x \in X} f(x)(V) d\nu$$

when $\boldsymbol{T}$ is a game functor ($\mathbf{J}$, $\bigtriangledown \mathbf{J}$, $\bigtriangleup \mathbf{J}$, $\mathbf{V}$), or $f^\dagger(F)(h) = F(\lambda x \in X \cdot f(x)(h))$ when $\boldsymbol{T}$ is a prevision functor ($\mathbf{P}^{\divideontimes}$, $\bigtriangledown \mathbf{P}^{\divideontimes}$, $\bigtriangleup \mathbf{P}^{\divideontimes}$, $\mathbf{P}^{\triangle}$). While this indeed works when $\boldsymbol{T} = \mathbf{V}$ [7, Section 4.2], or when $\boldsymbol{T} = \mathbf{P}^{\triangle}$ using the isomorphism between $\mathbf{V}$ and the latter, it fails for the other functors. To understand why, take $\boldsymbol{T} = \bigtriangledown \mathbf{P}^{\divideontimes}$, and consider $X = \{1, 2\}$, $Y = \{*_{11}, *_{12}, *_{21}, *_{22}\}$ (with their discrete topologies), $F = \alpha_{\mathcal{C}}(\mathfrak{u}_{\{1,2\}})$, i.e., $F(h) = \min(h(1), h(2))$ for every $h : Y \to \mathbb{R}^+$, $f : X \to \boldsymbol{T}Y$ defined by $f(1) = \alpha_{\mathcal{C}}(3/4\delta_{*_{11}} + 1/4\delta_{*_{12}})$ and $f(2) = \alpha_{\mathcal{C}}(1/3\delta_{*_{21}} + 2/3\delta_{*_{22}})$, so that $f(1)(h) = 3/4h(*_{11}) + 1/4h(*_{12})$ and $f(2)(h) = 1/3h(*_{21}) + 2/3h(*_{22})$ for every

$h : Y \to \mathbb{R}^+$. Let h and h' be defined by: $h(*_{11}) = 0.3$, $h(*_{12}) = h(*_{22}) = 0.1$, $h(*_{21}) = 0.7$, $h'(*_{11}) = 0.5$, $h'(*_{12}) = h'(*_{22}) = 0$, $h'(*_{21}) = 0.7$, then $f^\dagger(F)(h) = 0.25$, $f^\dagger(F)(h') = 0.233\ldots$, $f^\dagger(F)(h+h') = 0.533\ldots$, but $f^\dagger(F)(h)+f^\dagger(F)(h') = 0.4833\ldots \neq f^\dagger(F)(h+h')$, although h and h' *are* comonotonic. In other words, $\_^\dagger$ does not preserve collinearity.

In everyday terms, collinear previsions, or more specifically belief functions represent a process where P draws at random first, then C chooses non-deterministically [5]. The example above is a non-deterministic choice (among $\{1, 2\}$) followed by probabilistic choices. In other words, the non-deterministic player C plays first, followed by the probabilistic player P. But it is well-known that you cannot permute non-deterministic and probabilistic choices, and the example above only serves to restate this.

Our cure is simple: drop the collinearity condition. We shall therefore consider monads of continuous (plain, lower, upper) previsions. Let ***Posc*** be the category of posets with Scott-continuous maps, ***Cpo*** its full subcategory of cpos. We consider posets equipped with their Scott topology, whence these two categories are full subcategories of ***Top***. Note that $\mathbf{P}(X)$, $\bigtriangledown\mathbf{P}(X)$, $\bigtriangleup\mathbf{P}(X)$ are only posets, not cpos.

Theorem 2. *Define $\boldsymbol{T}X$ as $\mathbf{P}(X)$, resp. $\bigtriangledown\mathbf{P}(X)$, resp. $\bigtriangleup\mathbf{P}(X)$. Let $\boldsymbol{\eta}_X(x) = \lambda h \in \langle X \to \mathbb{R}^+\rangle \cdot h(x)$, and $f^\dagger(F)(h) = F(\lambda x \in X \cdot f(x)(h))$ for every $f : X \to \boldsymbol{T}Y$. Then $\boldsymbol{T}$ is a monad on* ***Top****, i.e., $(\boldsymbol{T}, \boldsymbol{\eta}, \_^\dagger)$ is a Kleisli triple. On* ***Posc****, $\boldsymbol{T}$ is a strong monad: $\boldsymbol{t}_{X,Y} : X \times \boldsymbol{T}Y \to \boldsymbol{T}(X \times Y)$ defined as $\boldsymbol{t}_{X,Y}(x, F)(h) = F(\lambda y \in Y \cdot h(x, y))$ is a tensorial strength.*

Proof. We must first show that, for every $f : X \to \boldsymbol{T}Y$, $f^\dagger$ is indeed a continuous map from $\boldsymbol{T}X$ to $\boldsymbol{T}Y$. Foremost, we must make sure that for every continuous (plain, lower, upper) prevision F on X, $f^\dagger(F)$ is a continuous (plain, lower, upper) prevision on Y. This is easy, but relatively tedious verification. Now note that the formulae defining $\boldsymbol{\eta}$, $\_^\dagger$, $\boldsymbol{t}$ are exactly the formulae defining the *continuation monad* [17]. It follows that the Kleisli triple axioms also hold in our case.

Contrarily to what might be expected, $\boldsymbol{t}_{X,Y}$ is not defined on all of ***Top***—it may fail to be continuous. On ***Posc***, this is repaired by the fact that a function of two arguments is continuous iff it is continuous in each argument separately (a fact that fails in ***Top***). The tensorial strength equations [17] are checked as for the continuation monad. □

That the formulae for unit, extension, and tensorial strength are the same as for the continuation monad is no accident. Imagine $F \in \boldsymbol{T}X$ is the semantics of a (probabilistic and non-deterministic) program expected to return a result x of type X. As we have already argued, when $F = \alpha_{\mathcal{C}}(P)$, with P a continuous valuation, then $F(h)$ is the *average* payoff, defined as the (Choquet) integral of $h(x)$ along P. When $F = \alpha_{\mathcal{C}}(\nu)$ with ν a continuous belief function, then $F(h)$ is the average min-payoff, where minima are taken over (demonically) non-deterministic choices. When F is not collinear, then more complicated "averaging" processes are involved. In particular, we allow taking means of mins of means of mins… representing plays where P, C, P, C, … take turns. That arbitrarily many turns can be chained in a (not necessarily collinear) prevision will be a consequence of the fact that prevision functors define monads, and in particular have a well-defined multiplication. This is standard in the monadic approach to side-effects [17]: multiplication is the key to defining sequential composition—here, of plays.

More explicitly, take n continuous functions $f_1 : X_0 \to \boldsymbol{T}X_1$, $f_2 : X_1 \to \boldsymbol{T}X_2$, $\ldots$, $f_n : X_{n-1} \to \boldsymbol{T}X_n$. Then, when $\boldsymbol{T}$ is a monad, $f_n^\dagger \circ f_{n-1}^\dagger \circ \ldots \circ f_2^\dagger \circ f_1 : X_0 \to \boldsymbol{T}X_n$ is the sequential composition $f_1; f_2; \ldots; f_n$ of f_1, f_2, $\ldots$, f_{n-1}, f_n in this order: given $x_0 \in X_0$, the process $f_1(x_0)$ computes some element $x_1 \in X_1$ (in our case, by drawing it "at random", say; deterministic computations are of course allowed, too), then $f_2(x_1)$ computes some $x_2 \in X_2$, etc. The monad laws then say that composing with the idle process $\boldsymbol{\eta}_X : X \to \boldsymbol{T}X$ does nothing, and that sequential composition is associative.

While Theorem 2 then establishes a form of soundness (which we shall make more precise below), the goal of the next sections will be to show that the prevision axioms are complete, in the sense that there is no junk: every continuous (lower, upper) prevision is a mix of (demonic, angelic) non-deterministic and probabilistic choices.

One may wonder what the equivalent of *normalized* games ($\nu(X) = 1$) and *sub-normalized* games ($\nu(X) \leq 1$) would be through the correspondence of Theorem 1. Requiring $F(\chi_X)$ to be equal (resp. less than or equal) to 1 is the obvious choice. However, this is not preserved by $\_^\dagger$ when F is not collinear. So we define:

Definition 2. *A prevision F on X is* normalized, *resp.* sub-normalized, *iff for every $f \in \langle X \to \mathbb{R}^+\rangle$, for every $a \in \mathbb{R}^+$, $F(a+f) = a+F(f)$ (resp. $F(a+f) \leq a+F(f)$).*

We let $\mathbf{J}_1(X)$, $\bigtriangledown\, \mathbf{P}_1^{\ast}(X)$, $\bigtriangledown\, \mathbf{P}_1(X)$, $\ldots$, be the subspaces of normalized games/previsions, and $\mathbf{J}_{\leq 1}(X)$, $\bigtriangledown\, \mathbf{P}_{\leq 1}^{\ast}(X)$, $\bigtriangledown\, \mathbf{P}_{\leq 1}(X)$, $\ldots$, those of sub-normalized games/previsions.

Proposition 1. *Theorem 1 again holds for normalized (continuous) games and previsions, and for sub-normalized (continuous) games and previsions.*

Now the spaces of sub-normalized and normalized continuous previsions are cpos. The spaces of sub-normalized continuous previsions are *pointed*, i.e., they have a least element $\perp$, the constant 0 function. If X is itself pointed, then the spaces of normalized continuous previsions are pointed, too, with least element $\alpha_{\mathfrak{C}}(\delta_\perp)$ (a continuous linear prevision). The latter maps $h \in \langle X \to \mathbb{R}^+\rangle$ to $h(\perp)$. Let ***Cpo*** the category of cpos, ***Pcpo*** that of pointed cpos. It follows:

Proposition 2. *Let $\boldsymbol{T}X$ be $\mathbf{P}_{\leq 1}(X)$, $\bigtriangledown\, \mathbf{P}_{\leq 1}(X)$, $\bigtriangleup\, \mathbf{P}_{\leq 1}(X)$, $\mathbf{P}_1(X)$, $\bigtriangledown\, \mathbf{P}_1(X)$, or $\bigtriangleup\, \mathbf{P}_1(X)$. $(\boldsymbol{T}, \boldsymbol{\eta}, \boldsymbol{\mu}, \boldsymbol{t})$ is a strong monad on **Cpo** and on **Pcpo**.*

Theorem 2 allows us to give a semantics to a λ-calculus with both probabilistic and non-deterministic choices. Consider the syntax of terms and types:

$M, N, P ::= x$	variable
$\mid\ c$	constant
$\mid\ MN$	application
$\mid\ \lambda x \cdot M$	abstraction
$\mid\ ()$	empty tuple
$\mid\ (M, N)$	pair
$\mid\ \mathrm{fst}\ M$	first projection
$\mid\ \mathrm{snd}\ M$	second projection
$\mid\ \mathtt{val}\ M$	trivial computation
$\mid\ \mathtt{let\ val}\ x = M\ \mathtt{in}\ N$	let-expression

$\tau ::= \alpha$	base types
$\mid\ \mathtt{u}$	type of $()$
$\mid\ \tau \times \tau$	product
$\mid\ \tau \to \tau$	function types
$\mid\ T\tau$	computation types

The typing rules, as well as the categorical semantics in a let-CCC, are standard [17]. Note that ***Cpo*** and ***Pcpo*** are Cartesian-closed. Together with the strong monads of

Proposition 2, they form let-CCCs. The typing rules for computation types are: if $\Gamma \vdash M : \tau$ then $\Gamma \vdash \mathtt{val}\, M : T\tau$; and if $\Gamma \vdash M : T\tau_1$ and $\Gamma, x : \tau_1 \vdash N : T\tau_2$ then $\Gamma \vdash \mathtt{let\, val}\, x = M\, \mathtt{in}\, N : T\tau_2$.

As should be expected, the semantics has a strong continuation flavor. For each term M of type τ in context $\Gamma = x_1 : \tau_1, \ldots, x_n : \tau_n$, $[\![M]\!]$ is a morphism (a continuous map) from $[\![\Gamma]\!] = [\![\tau_1]\!] \times \ldots \times [\![\tau_n]\!]$ to $[\![\tau]\!]$. The cases for `val` and `let` are given by: $[\![\mathtt{val}\, M]\!]\,(v_1, \ldots, v_n) = \lambda h \in \langle [\![\tau]\!] \to \mathbb{R}^+ \rangle \cdot h([\![M]\!]\,(v_1, \ldots, v_n))$, and $[\![\mathtt{let\, val}\, x = M\, \mathtt{in}\, N]\!]\,(v_1, \ldots, v_n) = \lambda h \in \langle [\![\tau_2]\!] \to \mathbb{R}^+ \rangle \cdot [\![M]\!]\,(v_1, \ldots, v_n)(\lambda v \in [\![\tau_1]\!] \cdot [\![N]\!]\,(v_1, \ldots, v_n, v)(h))$. Let `bool` be a base type, with $[\![\mathtt{bool}]\!] = \mathbb{S}$, where $\mathbb{S} = \{0, 1\}$ is Sierpiński space ($0 < 1$). Constants c may include a least fixpoint operator in ***Pcpo***, the Boolean constants `false`, `true`, a case construct $\mathtt{case} : \mathtt{bool} \times \tau \times \tau \to \tau$ with $[\![\mathtt{case}]\!]\,(0, v_0, v_1) = v_0$ and $[\![\mathtt{case}]\!]\,(1, v_0, v_1) = v_1$. The interpretation of T as a monad of previsions allows us, additionally, to give meaning to a coin-flipping operator $\mathtt{flip} : T\mathtt{bool}$, with $[\![\mathtt{flip}]\!] = \alpha_{\mathcal{C}}(1/2\delta_0 + 1/2\delta_1) = \lambda h \in \langle \mathbb{S} \to \mathbb{R}^+ \rangle \cdot 1/2(h(0) + h(1))$, and a non-deterministic choice operator $\mathtt{amb} : T\mathtt{bool}$. When $\boldsymbol{T}$ is $\bigtriangledown \mathbf{P}_1$, `amb` is the demonic choice (of a Boolean): $[\![\mathtt{amb}]\!] = \alpha_{\mathcal{C}}(\mathfrak{u}_{\{0,1\}}) = \lambda h \in \langle \mathbb{S} \to \mathbb{R}^+ \rangle \cdot \min(h(0), h(1))$ (the chosen Boolean x is the one that minimizes payoff $h(x)$). When $\boldsymbol{T}$ is $\bigtriangleup \mathbf{P}_1$, we get angelic choice: $[\![\mathtt{amb}]\!] = \alpha_{\mathcal{C}}(\mathfrak{e}_{\{0,1\}}) = \lambda h \in \langle \mathbb{S} \to \mathbb{R}^+ \rangle \cdot \max(h(0), h(1))$ (maximize payoff).

One might think that letting $\boldsymbol{T}$ be $\mathbf{P}_1$ would lead to chaotic choice. This certainly accommodates both demonic ($\min$) and angelic choice ($\max$). However, $\mathbf{P}_1$ is a very large space, and seems to contain objects that do not correspond to any mixture of probabilistic and non-deterministic choice. The right notion is suggested by [6, section 7.5].

Definition 3 (Fork). *A* fork *on X is any pair $F = (F^-, F^+)$ where F^- is a lower prevision, F^+ is an upper prevision, and for any $h, h' \in \langle X \to \mathbb{R}^+ \rangle$,*

$$F^-(h + h') \leq F^-(h) + F^+(h') \leq F^+(h + h') \tag{3}$$

F is continuous, *resp.* normalized, sub-normalized, collinear, *whenever both F^- and F^+ are.*

While the above definition was found from purely mathematical arguments, Walley [28, Section 2] defines essentially the same notion in finance. However, we allow any pair (F^-, F^+) satisfying these conditions to be a fork. Walley only observes that whenever F^- is a coherent prevision (in his sense), on a discrete space, then letting $F^+(h) = -F^-(-h)$ yields a fork (F^-, F^+).

One may think of F^- as the pessimistic part of F, which will give us the least expected payoff, while F^+ is the optimistic part, yielding the greatest expected payoff. Taking $h' = 0$ in (3) shows indeed that $F^-(h) \leq F^+(h)$ for each h. Let $\mathbf{F}(X)$ be the space of continuous forks on X, ordered by $\leq \times \leq$. The subspaces $\mathbf{F}_1(X)$ and $\mathbf{F}_{\leq 1}(X)$ of normalized and sub-normalized forks are cpos. The latter is pointed (with least element $(0, 0)$) and the former is as soon as X is (with least element $(\alpha_{\mathcal{C}}(\delta_\bot), \alpha_{\mathcal{C}}(\delta_\bot))$). The semantics is essentially the pairing of two continuation semantics, e.g., $[\![\mathtt{val}\, M]\!]\,(v_1, \ldots, v_n) = (F^-, F^+)$, where $F^- = F^+ = \lambda h \in \langle [\![\tau]\!] \to \mathbb{R}^+ \rangle \cdot h([\![M]\!]\,(v_1, \ldots, v_n))$ (a *linear* prevision); $[\![\mathtt{let\, val}\, x = M\, \mathtt{in}\, N]\!]\,(v_1, \ldots, v_n) = (\lambda h \in \langle [\![\tau_2]\!] \to \mathbb{R}^+ \rangle \cdot F^-(\lambda v \in [\![\tau_1]\!] \cdot F_v^-(h)), \lambda h \in \langle [\![\tau_2]\!] \to \mathbb{R}^+ \rangle \cdot F^+(\lambda v \in [\![\tau_1]\!] \cdot F_v^+(h)))$, where

$(F^-, F^+) = [\![M]\!](v_1, \ldots, v_n)$ and $(F_v^-, F_v^+) = [\![N]\!](v_1, \ldots, v_n, v)$. The constants with the most interesting semantics are `amb`, where $[\![\texttt{amb}]\!] = (\lambda h \in \langle \mathbb{S} \to \mathbb{R}^+ \rangle \cdot \min(h(0), h(1)), \lambda h \in \langle \mathbb{S} \to \mathbb{R}^+ \rangle \cdot \max(h(0), h(1)))$ (i.e., it computes both pessimistic and optimistic outcomes), and `flip`, where $[\![\texttt{flip}]\!] = (F^-, F^+)$ with $F^- = F^+ = \lambda h \in \langle \mathbb{S} \to \mathbb{R}^+ \rangle \cdot 1/2(h(0) + h(1))$. For the rest of the language, we rely on [17] and:

Proposition 3. *Let $\boldsymbol{T}X$ be defined as $\mathbf{F}(X)$, $\mathbf{F}_{\leq 1}(X)$, or $\mathbf{F}_1(X)$. Let $\boldsymbol{\eta}_X(x) = (F^-, F^+)$ with $F^- = F^+ = \lambda h \in \langle X \to \mathbb{R}^+ \rangle \cdot h(x)$, and for every $f : X \to \boldsymbol{T}Y$, let $f^\dagger(F^-, F^+) = (\lambda h \in \langle Y \to \mathbb{R}^+ \rangle \cdot F^-(\lambda x \in X \cdot f^-(x)(h)), \lambda h \in \langle Y \to \mathbb{R}^+ \rangle \cdot F^+(\lambda x \in X \cdot f^+(x)(h)))$, where by convention $f(x) = (f^-(x), f^+(x))$. Then $(\boldsymbol{T}, \boldsymbol{\eta}, \boldsymbol{\mu})$ is a monad on **Top**. Together with $\boldsymbol{t}_{X,Y} : X \times \boldsymbol{T}Y \to \boldsymbol{T}(X \times Y)$ defined by $\boldsymbol{t}_{X,Y}(x, (F^-, F^+)) = (\lambda h \in \langle Y \to \mathbb{R}^+ \rangle \cdot F^-(\lambda y \in Y \cdot h(x, y)), \lambda h \in \langle Y \to \mathbb{R}^+ \rangle \cdot F^+(\lambda y \in Y \cdot h(x, y)))$, it forms a strong monad on **Cpo** and **Pcpo**.*

Proof. That the strong monad laws are satisfied is obvious. The core of the proof is in showing that unit, extension, and tensorial strength are well-defined. We deal with extension. Recall that $f^\dagger(F^-, F^+) = (F'^-, F'^+)$, where $F'^- = \lambda h \in \langle Y \to \mathbb{R}^+ \rangle \cdot F^-(\lambda x \in X \cdot f^-(x)(h))$ and $F'^+ = \lambda h \in \langle Y \to \mathbb{R}^+ \rangle \cdot F^+(\lambda x \in X \cdot f^+(x)(h)))$. Then $F'^-(h + h') = \lambda h \in \langle Y \to \mathbb{R}^+ \rangle \cdot F^-(\lambda x \in X \cdot f^-(x)(h + h')) \leq \lambda h \in \langle Y \to \mathbb{R}^+ \rangle \cdot F^-(\lambda x \in X \cdot f^-(x)(h) + f^+(x)(h'))$ (since $f(x) = (f^-(x), f^+(x)) \in \boldsymbol{T}Y$ and F^- is monotonic) $\leq \lambda h \in \langle Y \to \mathbb{R}^+ \rangle \cdot F^-(\lambda x \in X \cdot f^-(x)(h)) + F^+(\lambda x \in X \cdot f^+(x)(h'))$ (since $(F^-, F^+) \in \boldsymbol{T}X$) $= F'^-(h) + F'^+(h')$. Similarly, $F'^-(h) + F'^+(h') \leq F'^+(h + h')$. ☐

5 Hearts and Skins: Completeness

One of the fundamental theorems of the theory of cooperative games is Shapley's Theorem, which states that every convex game ν has a non-empty core (on finite discrete X)—the core $Core(\nu)$ being the set of measures p such that $\nu \leq p$ and $\nu(X) = p(X)$. A refinement of this is Rosenmuller's Theorem, which states that a game ν is convex iff its core is non-empty and for every function $f : X \to \mathbb{R}^+$, the integral of f along ν is the minimum of all integrals of f along p, $p \in Core(\nu)$. In particular, there is a measure p such that $\nu \leq p$, $\nu(X) = p(X)$, and integrating f along p gives the same result as integrating it along ν [4].

We show that the same results hold in the continuous case in [6, chapitre 10]. Remember that games correspond to collinear previsions. Our purpose here is to show that similar theorems hold on previsions that need not be collinear (see [6, chapitre 11] for a more complete development). The analogue of measures will be linear previsions. We drop the analogue of the $\nu(X) = p(X)$ condition, however we concentrate on normalized games and previsions, because the technical treatment is slightly easier. We call the analogue of cores hearts, and the dual notion skin.

Definition 4 (Heart, Skin). *For any function F from $\langle X \to \mathbb{R}^+ \rangle$ to $\mathbb{R}^+$, its* heart *$Coeur(F)$ is the set of linear functionals G such that $F \leq G$. Its* continuous heart *$CCoeur(F)$ is the subset of those Gs that are continuous. Its* skin *$Peau(F)$ is the set of linear functionals G such that $G \leq F$. Its* continuous skin *$CPeau(F)$ is the subset of those functionals G that are continuous.*

Again, we let $Coeur_1(F)$, $CCoeur_1(F)$, …, be the subsets of the corresponding spaces consisting of normalized previsions only, and similarly $Coeur_{\leq 1}(F)$, …, for those consisting of sub-normalized previsions.

Most of the developments below rest on Roth's Sandwich Theorem ([21], [24, Theorem 3.1]), which states that on every ordered cone C, for every positively homogeneous super-additive function $q : C \to \overline{\mathbb{R}}^+$ and every positively homogeneous sub-additive function $p : C \to \overline{\mathbb{R}}^+$ such that $a \leq b$ implies $q(a) \leq p(b)$ (e.g., when $q \leq p$ and either q or p is monotonic), then there is a monotonic linear function $f : C \to \overline{\mathbb{R}}^+$ such that $q \leq f \leq p$. Here $\overline{\mathbb{R}}^+$ is $\mathbb{R}^+$ plus an extra point at infinity $+\infty$. A *cone* is a set C, together with a binary operation $+$ turning it into a commutative monoid, and a scalar multiplication $\cdot$ from $\mathbb{R}^+ \times C$ to C, such that $1 \cdot a = a$, $0 \cdot a = 0$, $(rs) \cdot a = r \cdot (s \cdot a)$, $r \cdot (a + b) = r \cdot a + r \cdot b$, and $(r + s) \cdot a = r \cdot a + s \cdot a$. An *ordered cone* is equipped in addition with a partial ordering $\leq$ making $+$ and $\cdot$ monotonic. We only use Roth's Theorem on ordered cones of the form $\langle X \to \mathbb{R}^+ \rangle$. Our key result is:

Theorem 3. *Let X be a stably locally compact space, F a continuous lower prevision, and f a bounded continuous function from X to $\mathbb{R}^+$. Then there is a continuous linear functional G from $\langle X \to \mathbb{R}^+ \rangle$ to $\overline{\mathbb{R}}^+$ such that $F \leq G$ and $F(f) = G(f)$.*

Proof. Let F be a lower prevision on X, and $f \in \langle X \to \mathbb{R}^+ \rangle$. Define $\breve{F}_f$ by $\breve{F}_f(g) = \inf_{\substack{\lambda \in \mathbb{R}^+ \\ \lambda f \geq g}} \left[F(\lambda f) - \sup_{\substack{h \in \langle X \to \mathbb{R}^+ \rangle \\ g + h \leq \lambda f}} F(h) \right]$, taking this to be $+\infty$ is there is no $\lambda \in \mathbb{R}^+$ such that $\lambda f \geq g$. One checks that $\breve{F}_f$ is monotonic, positively homogeneous, sub-additive, above F ($\breve{F}_f(g) \geq F(g)$ for all g), touches F at f ($\breve{F}_f(f) = F(f)$). Roth's Sandwich Theorem then gives us a monotonic linear functional G_0 such that $F \leq G_0$ and $F(f) = G_0(f)$. However, G_0 may fail to be continuous. One now observes that $\langle X \to \mathbb{R}^+ \rangle$ is a continuous poset, with a basis B consisting of step functions. By Scott's Formula, the functional G defined by $G(f) = \sup_{g \in B, g \ll f} G_0(g)$ is continuous; in fact, the largest continuous functional below G_0. It follows that $F \leq G$ and $F(f) = G(f)$. The most difficult part of the proof is showing that G is linear. This rests on the fact that $\ll$ is multiplicative i.e., for any $a > 0$, $f \ll g$ iff $a \cdot f \ll a \cdot g$, and additive, i.e., if $h, f, g \in \langle X \to \mathbb{R}^+ \rangle$ are such that $h \ll f + g$, then $h \leq f' + g'$ for some $f', g' \in B$ with $f' \ll f$, $g' \ll g$; and conversely, $f' \ll f$ and $g' \ll g$ imply $f' + g' \ll f + g$. □

Note that G may take the value $+\infty$. We can refine this in the case of normalized previsions (for sub-normalized previsions, see [6, section 11.4]):

Theorem 4. *Let X be a stably locally compact space, F a normalized continuous lower prevision on X, and f a bounded continuous function from X to $\mathbb{R}^+$. Then there is a normalized continuous linear prevision G such that $F \leq G$ and $F(f) = G(f)$.*

Proof. Similar to Theorem 3. However, it may be that $\breve{F}_f$ reaches $+\infty$. Refine this by letting $\breve{\breve{F}}_f(g) = \inf_{\epsilon \in \mathbb{R}^+} \breve{F}_{f+\epsilon}(g)$, and using $\breve{\breve{F}}_f$ instead of $\breve{F}_f$. One checks that, since F is normalized, $\breve{F}_{f+\epsilon}$ is antitone in ϵ. Then $\breve{\breve{F}}_f$ is again monotonic, positively homogeneous, sub-additive (using antitony in ϵ), above F, and touches F at f. Moreover, it is easy to see that $\breve{F}_f(\chi_X) = 1$. We build G_0, then G from $\breve{F}_f$, as in Theorem 3. Additionally, we need X to be compact so as to establish that $G(\chi_X) = 1$. Since G is linear, it follows that G is normalized. $\Box$

One can deal with upper previsions instead, see [6, section 11.5], using a notion we call convex-concave duality to reduce to the above. We then obtain [6, théorème 11.5.22] that, when X is stably compact, F is a normalized continuous upper prevision on X, there is a normalized continuous linear prevision G on X such that $G \leq F$. Moreover, for every $f \in \langle X \to \mathbb{R}^+\rangle$, $F(f) = \sup_{G \in CPeau_1(F)} G(f)$.

Theorem 4 allows us to state a form of Rosenmuller's Theorem:

Theorem 5. *Let X be stably locally compact, F a continuous normalized prevision on X. Then F is lower iff $CCoeur_1(F) \neq \emptyset$ and for every $f \in \langle X \to \mathbb{R}^+\rangle$, $F(f) = \inf_{G \in CCoeur_1(F)} G(f)$. In this case, the inf is attained: $F(f) = \min_{G \in CCoeur_1(F)} G(f)$.*

There is, of course, a dual theorem on upper previsions and their skins [6, théorème 11.7.5]; infs are replaced by sups, which need not be attained.

Our completeness results to come are based on another topology on spaces of previsions: the *weak topology* is the coarsest that makes the function $F \mapsto F(f)$ continuous, for each $f \in \langle X \to \mathbb{R}^+\rangle$. The Scott topology is in general finer. Let $\bigtriangledown \mathbf{P}_{1\ wk}(X)$ be $\bigtriangledown \mathbf{P}_1(X)$ with the weak topology, and similarly for other spaces. Then:

Proposition 4. *Let X be stably compact, F a normalized continuous lower prevision, then $CCoeur_1(F)$ is a non-empty saturated compact convex subset of $\mathbf{P}^{\triangle}_{1\ wk}(X)$.*

Compactness follows from Plotkin's version of the Banach-Alaoglu Theorem [19], or using a technique by Heckmann and Jung [8, Section 3.2], while convexity (i.e., $\alpha F + (1-\alpha)F'$ is in $CCoeur_1(F)$ as soon as F and F' are, $\alpha \in [0,1]$) is clear. That skins are closed is much easier, while non-emptiness is by the dual Rosenmuller Theorem:

Proposition 5. *Let X be a topological space, F a normalized continuous upper prevision, then $CPeau_1(F)$ is a closed convex subset of $\mathbf{P}^{\triangle}_{1\ wk}(X)$. It is non-empty as soon as X is stably compact.*

Finally, call a *lens* of a space X any non-empty intersection $L = Q \cap F$ of a saturated compact Q and a closed subset F. Then:

Proposition 6. *Let X be a stably compact space. The continuous normalized* body *$CCorps_1(F) = CCoeur_1(F^-) \cap CPeau_1(F^+)$ of a continuous normalized fork $F = (F^-, F^+)$ on X is a lens. Moreover, $CCoeur_1(F^-) = \uparrow CCorps_1(F)$ and $CPeau_1(F^+) = \downarrow CCorps_1(F)$.*

Proof. We show that: $(*)$ for each $G \in CCoeur_1(F^-)$, there is a $G' \in CCoeur_1(F^-) \cap CPeau_1(F^+)$ such that $G' \leq G$. Let $F'(h) = \inf_{f,g \in \langle X \to \mathbb{R}^+\rangle / f+g \geq h}(F^+(f) + G(g))$. One checks that $F^- \leq F' \leq G$, that F' is an upper prevision, so by Roth's Sandwich Theorem, there is a linear monotonic functional G_0 such that $F^- \leq G_0 \leq F'$. Since $G_0 \leq F'$, G_0 does not take the value $+\infty$. Build G from G_0 using Scott's Formula, as before. It is easy to see that G is a continuous, normalized, linear prevision. Since $F^- \leq G'$, $G' \in CCoeur_1(F^-)$. Since $G' \leq F' \leq F^+$, $G' \in CPeau_1(F^+)$. Since $F' \leq F' \leq G$, $G' \leq G$ follows.

By $(*)$, $CCoeur_1(F^-) \cap CPeau_1(F^+)$ is non-empty. That $CCoeur_1(F^-) = \uparrow (CCoeur_1(F^-) \cap CPeau_1(F^+))$ is another easy consequence of $(*)$. Next we show $CPeau_1(F^+) = \downarrow (CCoeur_1(F^-) \cap CPeau_1(F^+))$ in a similar way, by defining $F''(h) = \sup_{\substack{f,g \in \langle X \to \mathbb{R}^+ \rangle \\ f+g \leq h}} (F^-(f) + G(g))$, where $G \in CPeau_1(F^-)$, and using F'' to show that there is some $G' \in CCoeur_1(F^-) \cap CPeau_1(F^+)$ such that $G \leq G'$. $\square$

These three propositions state that any normalized continuous lower prevision, resp. upper prevision, resp. fork F gives rise to an element $CCoeur_1(F)$, resp. $CPeau_1(F)$, resp. $CCorps_1(F)$ of the Smyth powerdomain $\mathcal{Q}(\mathbf{P}^{\triangle}_{1\ wk}(X))$ (demonic non-deterministic choice of a probability law—remember that $\mathbf{P}^{\triangle}_1(X) \cong \mathbf{V}_1(X)$), resp. the Hoare powerdomain $\mathcal{H}(\mathbf{P}^{\triangle}_{1\ wk}(X))$ over $\mathbf{P}^{\triangle}_{1\ wk}(X)$ (angelic), resp. the Plotkin powerdomain over $\mathbf{P}^{\triangle}_{1\ wk}(X)$ (chaotic). This is a form of *completeness*: spaces of previsions and of forks contain no junk, and really are no more than mixes of non-deterministic and probabilistic choice. More explicitly, in the demonic case, let F be any normalized continuous lower prevision on stably compact X: then F is the sequential composition $f_0; f_1 = {f_1}^\dagger \circ f_0$ (applied to some dummy $*$), where $f_0(*) = \alpha_{\mathcal{C}}(\mathfrak{u}_{CCoeur_1(F)})$ (non-deterministic choice of some G from the heart of F), and $f_1(G) = G$ (drawing at random along $\gamma_{\mathcal{C}}(G)$). We need Proposition 4 for this to be well-defined. $F = (f_0; f_1)(*)$ follows from the definition of the heart and general properties of the Choquet integral. In the angelic case, any normalized continuous upper prevision F is the sequential composition $f_0; f_1$, where $f_0(*) = \alpha_{\mathcal{C}}(\mathfrak{e}_{\mathcal{C}Peau_1(F)})$, and again $f_1(G) = G$.

In the converse direction, still assuming X stably compact, there is a map $\bigsqcap : \mathcal{Q}(\mathbf{P}^{\triangle}_{1\ wk}(X)) \to \bigtriangledown \mathbf{P}_1(X)$ defined by $\bigsqcap K(f) = \min_{G \in K} G(f)$, and $CCoeur_1 \dashv \bigsqcap$ is a Galois connection consisting of continuous maps [6, théorème 11.7.11], while there is a continuous map $\bigsqcup : \mathcal{H}(\mathbf{P}^{\triangle}_{1\ wk}(X)) \to \bigtriangledown \mathbf{P}_1(X)$ defined by $\bigsqcup C(f) = \sup_{G \in C} G(f)$, so that $\bigsqcup \dashv CPeau_1$ is a Galois insertion.

We conclude by noticing that, when X is a continuous cpo with a least element, $\mathbf{P}^{\triangle}_{1\ wk}(X)$ is homeomorphic to $\mathbf{V}_1(X)$ with the weak topology, and the latter coincides then with the Scott topology [8]. Apart from spurious details (e.g., we bound our valuations by 1 instead of $+\infty$), there is therefore a strong connection with the models of Mislove [16] and Tix [23, 24]. This is explored, under a different light, in [9].

References

1. Samson Abramsky and Achim Jung. Domain theory. In S. Abramsky, D. M. Gabbay, and T. S. E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 3, pages 1–168. Oxford University Press, 1994.
2. Neil Ghani and Tarmo Uustalu. Coproducts of ideal monads. *Theoretical Informatics and Applications*, 38(4):321–342, 2004.
3. Gerhard Gierz, Karl Heinrich Hofmann, Klaus Keimel, Jimmie D. Lawson, Michael Mislove, and Dana S. Scott. Continuous lattices and domains. In *Encyclopedia of Mathematics and its Applications*, volume 93. Cambridge University Press, 2003.
4. Itzhak Gilboa and David Schmeidler. Additive representation of non-additive measures and the Choquet integral. Discussion Papers 985, Northwestern University, Center for Mathematical Studies in Economics and Management Science, 1992.

5. Jean Goubault-Larrecq. Continuous capacities on continuous state spaces. In *ICALP'2007*. Springer-Verlag LNCS, 2007.
6. Jean Goubault-Larrecq. Une introduction aux capacités, aux jeux et aux prévisions. `http://www.lsv.ens-cachan.fr/~goubault/ProNobis/pp.pdf`, June 2007. Version 1.7.
7. Claire Jones. *Probabilistic Non-Determinism*. PhD thesis, University of Edinburgh, 1990. Technical Report ECS-LFCS-90-105.
8. Achim Jung. Stably compact spaces and the probabilistic powerspace construction. In J. Desharnais and P. Panangaden, editors, *Domain-theoretic Methods in Probabilistic Processes*, volume 87 of *Electronic Lecture Notes in Computer Science*. Elsevier, 2004. 15pp.
9. Klaus Keimel and Gordon Plotkin. Predicate transformers for convex powerdomains. *Math. Struct. Comp. Sci.*, 2007. Submitted. 42 pages.
10. G. Max Kelly. A unified treatment of transfinite constructions for free algebras, free monoids, colimits, associated sheaves and so on. *Bull. Austr. Math. Soc.*, 22:1–83, 1980.
11. Christoph Lüth. *Categorical Term Rewriting: Monads and Modularity*. PhD thesis, University of Edinburgh, 1997.
12. Sebastian Maaß. Coherent lower previsions as exact functionals and their (sigma-)core. In G. de Cooman, T. Fine, and T. Seidenfeld, editors, *Proc. 2nd Intl. Symp. Imprecise Probabilities and their Applications (ISIPTA'01)*, pages 230–236, 2001.
13. Sebastian Maass. Continuous linear representation of coherent lower previsions. In J.-M. Bernard, T. Seidenfeld, and M. Zaffalon, editors, *Proc. 3rd Intl. Symp. on Imprecise Probabilities and Their Applications (ISIPTA'03)*, pages 371–381. Carleton Sci. Proc. in Informatics 18, 2003.
14. Ernest G. Manes. *Algebraic Theories*, volume 26 of *Graduate Texts in Mathematics*. Springer-Verlag, 1976.
15. Michael Mislove. Topology, domain theory and theoretical computer science. *Topology and Its Applications*, 89:3–59, 1998.
16. Michael Mislove. Nondeterminism and probabilistic choice: Obeying the law. In *Proc. 11th Conf. Concur. Theory (CONCUR'00)*, pages 350–364. Springer Verlag LNCS 1877, 2000.
17. Eugenio Moggi. Notions of computation and monads. *Inf. & Comp.*, 93:55–92, 1991.
18. Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, 1994.
19. Gordon Plotkin. A domain-theoretic Banach-Alaoglu theorem. *Math. Struct. Comp. Sci.*, 16:299–311, 2006.
20. Norman Ramsey and Avi Pfeffer. Stochastic lambda calculus and monads of probability distributions. In *Proc. 29th Ann. ACM SIGPLAN-SIGACT Symp. Principles of Programming Languages (POPL'02)*, pages 154–165, 2002.
21. Walter Roth. Hahn-Banach type theorems for locally convex cones. *Journal of the Australian Mathematical Society*, 68(1):104–125, 2000.
22. Regina Tix. *Stetige Bewertungen auf topologischen Räumen*. Diplomarbeit, TH Darmstadt, June 1995.
23. Regina Tix. *Continuous D-Cones: Convexity and Powerdomain Constructions*. PhD thesis, Technische Universität Darmstadt, 1999.
24. Regina Tix, Klaus Keimel, and Gordon Plotkin. Semantic domains for combining probability and non-determinism. *Electronic Notes in Theoretical Computer Science*, 129:1–104, 2005.
25. Daniele Varacca. The powerdomain of indexed valuations. In *Proc. 17th Ann. Symp. Logic in Computer Science (LICS'02)*, pages 299–308. IEEE Computer Society Press, 2002.
26. Daniele Varacca and Glynn Winskel. Distributing probability over nondeterminism. *Math. Struct. Comp. Sci.*, 2005. Accepted. 26 pages.
27. Philip Wadler. Comprehending monads. *Math. Struct. Comp. Sci.*, 2:461–493, 1992.
28. Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.