# Random Variables over Domains

Michael Mislove
Tulane University

Work supported by NSF and ONR
ICALP 2005

# Nondeterminism vs. Probabilistic Choice

- *Nondeterminism:* Represents the environment making choices
  - Like riders selecting floors on an elevator
- *Probabilistic choice:* Represents random events affecting the system
  - Like random stops the elevator makes

# Standard Models for Nondeterminism and Probabilistic Choice

- $S$ - finite set of states
  - *Nondeterministic choice:* $(\mathcal{P}(S), \cup)$
  - *Probabilistic choice:* $(\mathbb{V}(S), \{_r+ \mid 0 \leq r \leq 1\})$

$$\mathbb{V}(S) = \{\textstyle\sum_{i=1}^{n} r_i \delta_{x_i} \mid 0 \leq r_i;\ \sum_i r_i \leq 1;\ x_i \in S\}$$

$$\textstyle\sum_{i=1}^{m} r_i \delta_{x_i} {}_r{+} \sum_{j=1}^{n} s_j \delta_{y_j} = \sum_{i=1}^{m} r \cdot r_i \delta_{x_i} + \sum_{j=1}^{n} (1-r) \cdot s_j \delta_{y_j}$$

$$\textstyle\sum_{i=1}^{m} r_i \delta_{x_i} \sqsubseteq \sum_{j=1}^{n} s_j \delta_{y_j} \quad \text{iff} \quad \sum_{x_i \in X} r_i \leq \sum_{y_j \in X} s_j \quad \forall X \in \mathcal{P}(S)$$

# Nondeterminism & Probability

- *Each defines an algebraic theory*
- *Nondeterminism: theory of semilattices*

$$x \sqcap (y \sqcap z) = (x \sqcap y) \sqcap z; \quad x \sqcap y = y \sqcap x; \quad x \sqcap x = x$$

- *Probabilistic choice: theory of probabilistic algebras*

$$x \,{}_r\!+ y = y \,{}_{1-r}\!+ x \qquad x \,{}_1\!+ y = x \qquad x \,{}_r\!+ x = x$$

$$(x \,{}_r\!+ y) \,{}_s\!+ z \quad = \quad x \,{}_{r\cdot s}\!+ (y \,{}_{\frac{s(1-r)}{1-r\cdot s}}\!+ z) \ \text{ if } r < 1$$

# Combining Theories

- $\mathcal{P}(\mathbb{V}(S))$ - *Sets of valuations*
- $X \;{}_r{+}\; Y = \{x \;{}_r{+}\; y \mid x \in X, y \in Y\}$
- Nondeterministic choice and probabilistic choice are entangled:
  $(\{\delta_x\} \cup \{\delta_y\}) \;{}_{\frac{1}{2}}{+}\; (\{\delta_x\} \cup \{\delta_y\}) = \{\delta_x, \delta_y, \delta_x \;{}_{\frac{1}{2}}{+}\; \delta_y\}$
- $X \sqcap Y = X \cup Y \cup (\cup_r \; X \;{}_r{+}\; Y)$
- **Theorem** *(Tix 1999/ M- 2000)*
  The power set induces an endofunctor on probabilistic algebras.

# Two Theorems

- **Theorem 1** *(Beck)*
  If $\langle S, \eta_S, \mu_S \rangle, \langle T, \eta_T, \mu_T \rangle \colon \mathsf{A} \to \mathsf{A}$ are monads, then the following are equivalent:
  - There is a distributive law $d \colon ST \dot{\to} TS$
  - $S$ lifts to a monad of $T$-algebras
- **Theorem 2** *(Plotkin & Varacca)*
  There is no distributive law of the power set over the probabilistic power domain, or *vice versa*.

# Alternative approach

- *Weaken (eliminate) one of the laws:*

$$x\ {}_r\!+ y = y\ {}_{1-r}\!+ x \qquad x\ {}_1\!+ y = x \qquad \text{~~}x\ {}_r\!+ x = x\text{~~}$$

~~$x_r + x = x$~~

$$(x\ {}_r\!+ y)\ {}_s\!+ z \quad = \quad x\ {}_{r\cdot s}\!+ (y\ {}_{\frac{s(1-r)}{1-r\cdot s}}\!+ z) \ \text{ if } r < 1$$

- New structures - *(finite) indexed valuations:*

$$IV(X) = [\dot{\bigcup}_{n>0}\ (\mathbb{R}_+ \times X)^n/ \equiv)]\ \cup \{\underline{0}\}$$

$$(r_i, x_i) \equiv (s_i, y_i) \text{ iff } (\exists \phi \in S(n))$$

$$(r_{\phi^{-1}(i)}, x_{\phi^{-1}(i)}) = (s_i, y_i)\ (i = 1, \ldots, n)$$

# Understanding Indexed Valuations

- $[r_i, x_i]_m \simeq \sum_{i=1}^{m} r_i \delta_{x_i};\qquad [1, x] \not\simeq [(\frac{1}{2}, x), (\frac{1}{2}, x)]$
- $r \cdot [r_i, x_i]_m \mapsto [r \cdot r_i, x_i]_m : \mathbb{R}_+ \times IV(X) \to IV(X)$
- $$\begin{aligned} [r_i, s_i]_m \oplus [s_j, y_j]_n &= [t_k, z_k]_{m+n} \\ (t_k, z_k) &= \begin{cases} (r_i, x_i) & \text{if } k \leq m \\ (s_j, y_j) & \text{if } m < k \leq n \end{cases} \end{aligned}$$
- $[r_i, x_i]_m \,{}_r\!+ [s_j, y_j]_n = r \cdot [r_i, x_i]_m \oplus (1 - r) \cdot [s_j, y_j]_n$

# Universal Properties

- *Indexed valuations are real quasi-cones:*

$$A + (B + C) = (A + B) + C \qquad A + B = B + A$$

$$r(A + B) = rA + rB \qquad r(sA) = (rs)A$$

$$0A = \underline{0}, \quad 1A = A \qquad \underline{0} + A = A$$

$$\xcancel{(r + s)A = rA + sA} \qquad r, s \in \mathbb{R}_+ \quad A, B, C \in IV(X)$$

- **Theorem** *(Varacca)*
  $IV$ : Set $\rightarrow$ Set defines a monad of real quasi-cones that enjoys a distributive law over $\mathcal{P}$. So, $(\mathcal{P} \circ IV)(X)$ is a real quasi-cone that also is a semilattice.

# Justifying Indexed Valuations

$$[r_i, x_i]_m \simeq \sum_{i=1}^{m} r_i \delta_{x_i}; \qquad [1, x] \not\simeq [(\frac{1}{2}, x), (\frac{1}{2}, x)]$$

$f\colon (P, \mu) \to (X, \Omega)$ a *random variable.*

$f\colon (P, \mu) \to (X, \Omega)$ induces $(f \cdot \mu)(U) = \mu(f^{-1}(U))$

– too coarse

$Flat\colon IV(X) \to \mathbb{P}(X)$ by $Flat([r_i, x_i]_m) = \sum_i r_i \delta_{x_i}$

– morphism of real quasi-cones.

# Generalizing to Domains

*Domain: Partial order in which directed sets have suprema*

- $A \subseteq D$ directed *if each finite subset has an upper bound in* $A$
- *Continuous:* $x \ll y$ *iff* $y \leq \sup A \Rightarrow x \leq a \in A$

  $y = \sup \{ x \mid x \ll y \}$ *- directed*

Example: $([0,1], \leq)$ $x \ll y$ *iff* $x = 0$ *or* $x < y$

# Categories of Domains

$f : D \to E$ *continuous* if $f$ preserves the order
and $f$ preserves sups of directed sets

Dom – domains and continuous functions
– *not* cartesian closed

BCD – *bounded complete domains* and continuous functions
$\cap$ – *is* cartesian closed
RB – retracts of *bifinite domains* and continuous functions
$\cap$ – *is* cartesian closed
FS – *FS-domains* and continuous functions
– *maximal* cartesian closed

# Constructing Bag Domains

$E \simeq D \times E = \dot{\bigcup}_n D^n$ – domain of lists over $D$

- leaves RB, FS invariant
- free domain monoid over $D$

$D^n / \equiv_{S(n)}$ – domain of $n$-bags over $D$

- leaves RB, FS invariant

$E_C = \dot{\bigcup}_n (D^n / \equiv_{S(n)})$ – bag domain over $D$

- leaves RB, FS invariant
- free commutative domain monoid over $D$

# Applying the Construction

$$IV(D) = \dot{\cup}_n \; ((\overline{\mathbb{R}_+} \times D)^n / \equiv_{S(n)}) \;\cup\; \{\underline{0}\}$$

- leaves RB, FS invariant
- free commutative domain monoid over $\overline{\mathbb{R}_+} \times D$ with $\bot = 0$

$$\mathbb{RV}(D) = \{[r_i, x_i]_m \in IV(D) \mid \textstyle\sum_i r_i \leq 1\} \cup \{\underline{0}\}$$

- discrete random variables over $D$

**Theorem:** $\mathbb{RV}$: RB $\to$ RB is a continuous endofunctor; the same is true for FS. $\mathit{Flat}\colon \mathbb{RV}(D) \to \mathbb{V}(D)$ is an epimorphism.

# Summary and Future Work

- Daniele Varacca first defined indexed valuations
  - Used *abstract bases*
  - No categorical results
- Our work first to introduce random variables
- Categorical results also new
- Bag domain results also new
- Possible further work:
  - Generalize to nondiscrete random variables
  - Applications to quantum computing - *entropy* & *majorization*