Programming Research Group

ENTROPIC GEOMETRY FROM LOGIC

Bob Coecke

PRG-RR-02-12

Oxford University Computing Laboratory
Wolfson Building, Parks Road, Oxford OX1 3QD

Abstract

1. A finitary probability space (= all probability measures on a fixed finite support) can be faithfully represented by a partial order equipped with a measure of content (e.g. Shannon entropy).

2. This partial order can be obtained via a purely order-theoretic systematic procedure starting from an algebra of properties. This procedure applies to any poset envisioned as an algebra of properties.

1 Introduction

For almost a century now, the dominant formal conception of uncertainty has been one in terms of probability measures on some support. However, already in the late twenties a conception of probability as "the logic of partial knowledge" has been put forward by F. P. Ramsey [6]. Later D. S. Scott relied on a more general notion of partial knowledge to propose the mathematical structure of a domain [7]. It is the intention of this paper to show that a probability space cannot only be conceived, not only be represented, but can also be order-theoretically constructed in terms of partial knowledge.

This work is part of the research program started in [3], a paper jointly written with K. Martin, where we proposed a partial order on classical and quantum probability measures. It is exactly this partial order that arises when introducing partial knowledge on a boolean logic.

The deep connection between domains and measures of content as revealed by K. Martin in [5] incarnates in the case of ordered probability spaces in having Shannon entropy as a measure of content [3]. A first unification of qualitative theories of information ($\sim$ domain theory) and quantitative ones ($\sim$ Shannon) is thus established. This, together with the fact the construction proposed in this paper has striking geometric features warrents the name "entropic geometry" for the resulting mathematical shape.[1]

We now recall some results from [3]. Let Δ^n be all probability distributions on $\{1,\ldots,n\}$, that is, either a list $x=(x_1,\ldots,x_n)\in[0,1]^n$ or a map $x:\{1,\ldots,n\}\to[0,1]::i\mapsto x_i$, with $\sum_{i=1}^{i=n}x_i=1$. The *spectrum* of x is the set $\mathrm{spec}(x):=\{x_i\mid 1\leq i\leq n\}$. Denote the collection of all permutations $\sigma:\{1,\ldots,n\}\to\{1,\ldots,n\}$ as $S(n)$. For a poset D we call $e\in D$ *maximal* iff its upper set is a singleton, that is $\uparrow e=\{e\}$. We denote the set of maximal elements by $\mathrm{Max}(D)$. The *bottom* $\perp$ (if it exists) satisfies $\uparrow\perp=D$.

Definition 1.1 [3] *For $x,y\in\Delta^n$, we have $x\sqsubseteq y$ iff there exists $\sigma\in S(n)$ such that $x\cdot\sigma$ and $y\cdot\sigma$ are monotone decreasing and if we have*

$$\forall i\in\{1,\ldots,n-1\}:(x\cdot\sigma)_i(y\cdot\sigma)_{i+1}\leq(x\cdot\sigma)_{i+1}(y\cdot\sigma)_i\,. \tag{1}$$

[1] This name was established in exchanges with K. Martin.

Theorem 1.2 [3] *Let $n \geq 2$. Then, $(\Delta^n, \sqsubseteq)$ is a partially ordered set with* $\mathrm{Max}(\Delta^n) = \{e \in \Delta^n \mid \mathrm{spec}(x) = \{0,1\}\}$ *and* $\perp = (1/n, \ldots, 1/n)$. *Moreover, it is a dcpo and admits the notions of partiality and approximation. Thus, Δ^n is entitled to be called a domain. Finally, Shannon entropy*

$$\mu(x) = -\sum_{i=1}^{n} x_i \log x_i$$

is a measurement on Δ^n in the sense of [5].

The significance of $x \sqsubseteq y$ is: "*State* y is more *informative* than *state* x". In epistemic terms this becomes: "*Observer* y has more *knowledge about the system* than *observer* x". Defining *Bayesian projections* for $x \in \Delta^{n+1}$ as

$$p_i(x) = \frac{1}{1 - x_i}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}) \in \Delta^n$$

we have the following property for $(\Delta^{n+1}, \sqsubseteq)$ in terms of $(\Delta^n, \sqsubseteq)$:

$$x \sqsubseteq y \Leftrightarrow (\forall i)(x_i, y_i < 1 \Rightarrow p_i(x) \sqsubseteq p_i(y)) . \tag{2}$$

This interprets as follows. The pure states $\{e_i\}_i$ are to be seen as the actual states the system can be in, while general mixed states x and y should be conceived as being epistemic. Eq.(2) expresses that whenever a state x stands for less knowledge about the system than state y, then, after Bayesian update with respect to the new knowledge that the actual state of the system is not e_i, the state $p_i(x)$ still stands for less knowledge than $p_i(y)$ due to the initial advantage in knowledge of y as compared to x. (For a detailed exposition see [3] §2.1 and §4.4.) Remarkably, eq.(2) envisioned as an inductive step provides a definition equivalent to Definition 1.1 when a base case $n = 2$ is postulated as:

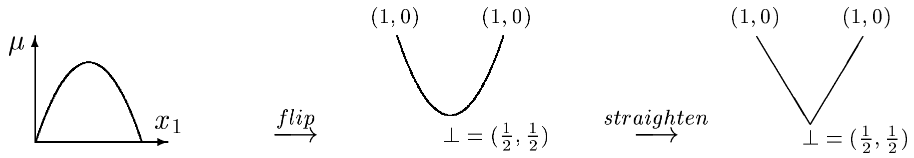
Definition 1.3 [3] *For $x, y \in \Delta^2$ we set*

$$(x_1, x_2) \sqsubseteq (y_1, y_2) \Leftrightarrow (y_1 \leq x_1 \leq 1/2) \text{ or } (1/2 \leq x_1 \leq y_1) .$$

Theorem 1.4 [3] *The order of Definition 1.3 is the only partial order on Δ^2 which has* $\perp = (1/2, 1/2)$ *and satisfies the mixing law:*

$$x \sqsubseteq y \text{ and } p \in [0,1] \;\Rightarrow\; x \sqsubseteq (1-p)x + py \sqsubseteq y .$$

The canonicity of this choice for the order on Δ^2 is also indicated by the correspondence between the shape of the Shannon entropy curve (left) and the graph of the order (right):



Conclusively, there exists an order on Δ^n with uniqueness properties that arises from envisioning probability distributions as informative objects. In this paper we address the following two questions:

1. How much structure does this order actually capture? The answer is provided in Theorem 4.1 and Corollary 4.2.

2. Is a probability space *a priori* required to produce these partial orders? The answer is provided in Section 5.

2 Symmetries and degeneration

Decreasing monotone distributions in Δ^n, that is, for all $i \in \{1, \ldots, n-1\}$ we have $x_i \geq x_{i+1}$, are denoted by Λ^n. For $x \in \Delta^n$, whenever $\sigma \in S(n)$ is such that $x \cdot \sigma \in \Lambda^n$ the monotone map $x^\Lambda := x \cdot \sigma : \{1, \ldots, n\} \to [0,1]^n$ does not depend on the particular choice of σ. It then follows that some $\sigma \in S(n)$ monotonizes $x \in \Delta^n$ iff it makes the following diagram commute:

$$\begin{tikzcd} \{1,\ldots,n\} \arrow[r, "x"] & {[0,1]} \\ \{1,\ldots,n\} \arrow[u, "\sigma"] \arrow[ur, "x^\Lambda"'] & \end{tikzcd} \tag{3}$$

The conditions eq.(1) can now be restated without explicit reference to σ.

Proposition 2.1 *For $x, y \in \Delta^n$, we have $x \sqsubseteq y$ iff*

1. *There exists at least one $\sigma \in S(n)$ such that $x \cdot \sigma$, $y \cdot \sigma \in \Lambda^n$*

2. *For all $i \in \{1, \ldots, n-1\}$ we have $x_i^\Lambda \cdot y_{i+1}^\Lambda \leq x_{i+1}^\Lambda \cdot y_i^\Lambda$.*

In the case that $x_{i+1}^\Lambda \neq 0$ and $y_{i+1}^\Lambda \neq 0$ the inequalities of Proposition 2.1 take the intuitively more transparent form $x_i^\Lambda / x_{i+1}^\Lambda \leq y_i^\Lambda / y_{i+1}^\Lambda$.

Let $x \in \Delta^n$. Let n^x be the cardinality of $\mathrm{spec}(x)$. Let $x^{\mathrm{spec}} \in \Lambda^{n^x}$ be the decreasingly ordered spectrum of x. Denote the multiplicity of value x_j^{spec} in the list x^Λ by n_j^x, or, n_j when it is clear from the context to which state this number applies. Then, set $K_1^{(x)} := \{1, \ldots, n_1\}$ and set

$$\forall j \in \{1, \ldots, n^x\} : \bar{n}_j := \sum_{i=1}^{i=j} n_i \qquad \forall j \in \{2, \ldots, n^x\} : K_j^{(x)} := \{\bar{n}_{j-1}^{(x)} + 1, \ldots, \bar{n}_j^{(x)}\}$$

that is $i \in K_j \;\Leftrightarrow\; x_i^\Lambda = x_j^{\rm spec}$. The diagram in eq.(3) then splits up in

$$
\begin{tikzcd}
\{1,\ldots,n\} \arrow[r, "x"] & {[0,1]} \\
K_1 \arrow[u, "\sigma"] \arrow[ur, "x^{\rm spec}(1)"'] &
\end{tikzcd}
\qquad \cdots \qquad
\begin{tikzcd}
\{1,\ldots,n\} \arrow[r, "x"] & {[0,1]} \\
K_{n^x} \arrow[u, "\sigma"] \arrow[ur, "x^{\rm spec}(n)"'] &
\end{tikzcd}
$$

where $x^{\rm spec}(1),\ldots,x^{\rm spec}(n)$ are constant maps. Requiring commutation then imposes an ordered partition $(\sigma[K_1],\ldots,\sigma[K_{n^x}])$ on $\{1,\ldots,n\}$.

For $i,j \in \{1,\ldots,n\}$ set $i \sim j$ whenever $x_i = x_j$. The corresponding equivalence classes then admit a total ordering $I_1^{(x)} \succ \ldots \succ I_{n^x}^{(x)}$ which is such that $I_k \succ I_l$ whenever for $i \in I_k$ and $j \in I_l$ we have $x_i > x_j$. Thus

$$i \in I_j \;\Leftrightarrow\; x_i = x_j^{\rm spec}\,. \tag{4}$$

The cardinality of I_j is that same as that of K_j, namely n_j.

Lemma 2.2 *For $x \in \Delta^n$ and $\sigma \in S(n)$ we have $x \cdot \sigma \in \Lambda^n$ iff*

$$\forall j \in \{1,\ldots,n^x\} : \sigma[K_j] = I_j\,.$$

Proof. Since by eq.(3) we have $x\cdot\sigma \in \Lambda^n \Leftrightarrow \forall i \in \{1,\ldots,n\} : x_i^\Lambda = (x\cdot\sigma)(i)$ the equivalence follows from $\sigma(i) \in \sigma[K_j] \Leftrightarrow i \in K_j \Leftrightarrow x_j^{\rm spec} = x_i^\Lambda$ and $\sigma(i) \in I_j \Leftrightarrow x_j^{\rm spec} = x_{\sigma(i)} = (x\cdot\sigma)(i)$. $\square$

Proposition 2.3 *Each $x \in \Delta^n$ is fully characterized by the pair*

- *the ordered partition $\mathcal{I}^x := (I_1,\ldots,I_{n^x})$ on $\{1,\ldots,n\}$,*
- *the $[0,1]$-valued n^x-element set ${\rm spec}(x)$.*

Conversely, each such pair defines a state $x \in \Delta^n$ provided that

$$\sum_{j=1}^{j=n^x} n_j \cdot x_j^{\rm spec} = 1\,.$$

Proof. Direction $\Rightarrow$ of eq.(4) fixes x given ${\rm spec}(x)$ and $(I_1,\ldots,I_{n^x})$. The converse follows by construction. $\square$

The degeneration of the spectrum of $x \in \Delta^n$ which is now encoded in the ordered partition $\mathcal{I}^x$ is of crucial importance due to the following.

Lemma 2.4 (Degeneration) [3] *If $x \sqsubseteq y$ in Δ^n, then*

$$x_i = 0 \quad \Rightarrow \quad y_i = 0 \tag{5}$$
$$y_i = y_j > 0 \quad \Rightarrow \quad x_i = x_j \tag{6}$$

This yields a hierarchy in Δ^n with respect to increase of the order:

zero-values/degeneration
non-degenerated non-zero values
degenerated non-zero values

Setting

$$\begin{cases} n_0^{(x)} := n^x & 0 \notin \mathrm{spec}(x) \\ n_0^{(x)} := n^x - 1 \ , \ \bar{n}_0 := \sum_{i=1}^{i=n_0} n_i \ , \ I_0 := I_{n^x} \ , \ K_0 := K_{n^x} & 0 \in \mathrm{spec}(x) \,. \end{cases}$$

We can express the Degeneration Lemma in terms of $\mathcal{I}^x$.

Lemma 2.5 (Degeneration') *If $x \sqsubseteq y$ in Δ^n, then*

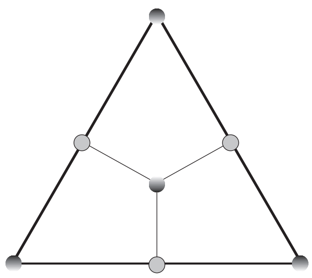
$$I_0^x \subseteq I_0^y \quad \& \quad \forall i \in \{1, \ldots, n_0^y\}, \exists j \in \{1, \ldots, n_0^x\} : I_i^y \subseteq I_j^x \,.$$

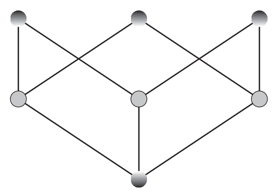
Proof. If I_i^y is a singleton then the claim follows from eq.(5). If in addition it has multiple elements then it follows from eq.(5) and eq.(6). □

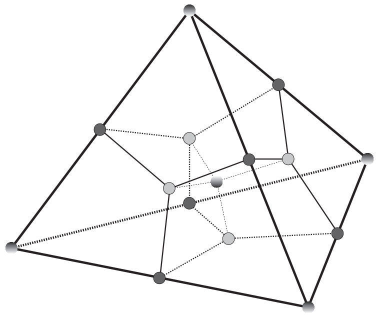
3 Coordinates

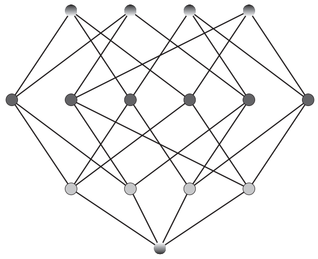
Definition 3.1 (Coordinates) *Let* $\mathrm{Coord}(\Delta^{\mathrm{n}})$ *be all $x \in \Delta^n$ with an at most binary spectrum. Let the degenerated coordinates* $\mathrm{Ir}_\perp(\Delta^{\mathrm{n}})$ *be the set of all* $x \in \mathrm{Coord}(\Delta^{\mathrm{n}})$ *with* $0 \in \mathrm{spec}(x)$*. For* $x \in \mathrm{Ir}_\perp(\Delta^{\mathrm{n}})$ *let the x-axis be the set of all* $y \in \mathrm{Coord}(\Delta^{\mathrm{n}})$ *with* $I_1^y = I_1^x$ *(and thus also* $I_2^y = I_2^x$*).*

As shown in [3] §4.3, $\mathrm{Ir}_\perp(\Delta^{\mathrm{n}})$ constitutes a subposet of Δ^n which, when top and bottom are added to it, is isomorphic to the powerset $\mathcal{P}(\{1, \ldots, n\})$. The illustrations below expose $\mathrm{Ir}_\perp(\Delta^{\mathrm{n}}) \cup \{\perp\}$ in the "triangle" Δ^3 and the "tetrahedron" Δ^4. The figures on the right are their Hasse diagrams.

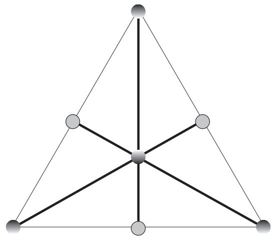


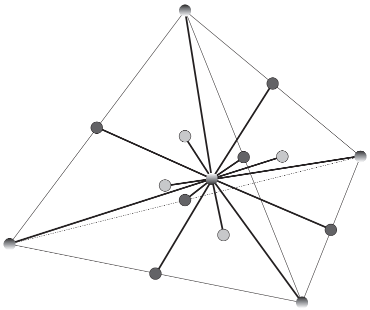






The segments represent increase of the order and coincide on the left and the right, the increase being respectively radially and upwardly. The coordinate axis of Δ^3 and Δ^4 look as follows.





Proposition 3.2 *Coordinates and coodinate axis are order-theoretical:*

- $\mathrm{Ir}_\perp(\Delta^{\mathrm{n}}) \cup \{\perp\}$ *are the infima of sets in* $\mathcal{P}(\mathrm{Max}(\Delta^n)) \setminus \{\emptyset\}$. (ref. [3])
- *If* $x \in \mathrm{Coord}(\Delta^{\mathrm{n}}) \setminus \mathrm{Ir}_\perp(\Delta^{\mathrm{n}})$ *then* $\downarrow x$ *is a chain. Conversely, if* $\downarrow x$ *is a chain then* $x \in \mathrm{Coord}(\Delta^{\mathrm{n}})$.
- *A coordinate axis is a maximal chain of non-degenerated coordinates completed with its supremum.*

Proof. Maximal elements and bottom are order-theoretic by definition and so are all $x \in \mathrm{Ir}_\perp(\Delta^n)$ since by [3] §4.3 we have $x = \bigwedge(\uparrow x \cap \mathrm{Max}(\Delta^n))$.

For $x \in \mathrm{Coord}(\Delta^{\mathrm{n}}) \setminus \mathrm{Ir}_\perp(\Delta^{\mathrm{n}})$ we have $x = \perp$ or $\mathcal{I}^x = (I_1^x, I_2^x)$. Let $x \neq \perp$. If $y \sqsubseteq x$ by Lemma 2.5 we have $I_1^x \subseteq I^y$ and $I_2^x \subseteq J^y$ for some $I^y, J^y \in \mathcal{I}^y$. Thus, $\mathcal{I}^y = \mathcal{I}^x$ or $I_1^y = \{1, \ldots, n\}$. If $y, z \in \downarrow x$ with $y \neq \perp \neq z$ then $\mathcal{I}^y = \mathcal{I}^z = \mathcal{I}^x$ and either $y^+ \cdot z^- \leq z^+ \cdot y^-$ or $z^+ \cdot y^- \leq y^+ \cdot z^-$ so y and z compare. The cases $x = \perp$, $y = \perp$ and $z = \perp$ are trivial so $\downarrow x$ is a chain.

Let $x \notin \mathrm{Coord}(\Delta^{\mathrm{n}})$. Then $\{I_1^x, I_2^x, I_3^x\} \subseteq \mathcal{I}^x$. But $y, z \in \Delta^n$ defined by

$$\begin{array}{cc} \mathcal{I}^y = \{I_1^x\ , \{1, \ldots, n\} \setminus I_1^x\} & \qquad \mathcal{I}^z = \{I_1^x \cup I_2^x\ , \{1, \ldots, n\} \setminus (I_1^x \cup I_2^x)\} \\ y_1^{\mathrm{spec}} \cdot x_2^{\mathrm{spec}} = x_1^{\mathrm{spec}} \cdot y_2^{\mathrm{spec}} & \qquad z_1^{\mathrm{spec}} \cdot x_3^{\mathrm{spec}} = x_2^{\mathrm{spec}} \cdot z_2^{\mathrm{spec}} \end{array}$$

(cfr. Proposition 2.3) don't compare although $y, z \sqsubseteq x$ so $\downarrow x$ is not a chain.

From the above we also know that for $x \in \mathrm{Coord}(\Delta^{\mathrm{n}}) \setminus \mathrm{Ir}_\perp(\Delta^{\mathrm{n}})$ and $y \sqsubseteq x$ we have $y \in \mathrm{Coord}(\Delta^{\mathrm{n}})$ and in particular that y belongs to the same axis as x. Thus for $y, z \in x$-axis with $z \neq x$ we have that $y \sqsubseteq w \sqsubseteq z$ forces $w \in x$-axis. Thus x-axis$\setminus\{x\}$ is a maximal chain in $\mathrm{Coord}(\Delta^{\mathrm{n}})\setminus\mathrm{Ir}_\perp(\Delta^{\mathrm{n}})$. By [3] Proposition 2.16 we then have $x = \bigsqcup(x\text{-axis}\setminus\{x\})$. $\square$

To $x \in \Delta^n \setminus \{\perp\}$ we attribute $\mathcal{C}^x = \{c(1), \ldots, c(n^x - 1)\} \subset \mathrm{Coord}(\Delta^{\mathrm{n}})$ as "its coordinates", where, using Proposition 2.3, each $c(j)$ is defined by

$$\mathcal{I}^{c(j)} = \left\{ \bigcup_{i=1}^{i=j} I_i^x \ , \ \bigcup_{i=j+1}^{i=n^x} I_i^x \right\} \qquad c(j)_1^{\mathrm{spec}} \cdot x_{j+1}^{\mathrm{spec}} = x_j^{\mathrm{spec}} \cdot c(j)_2^{\mathrm{spec}} .$$

Further we set $\mathcal{C}^\perp = \emptyset$. If $0 \in \mathrm{spec}(x)$ we set $c_0 := c(n^x - 1) \in \mathrm{Ir}_\perp(\Delta^{\mathrm{n}})$.

Theorem 3.3 (Decomposition in coordinates) *The elements of Δ^n and their coordinates are in bijective order-theoretic correspondence. Explicitly, for all $x \in \Delta^n$ we have*

$$x = \bigsqcup \mathcal{C}^x \qquad \text{and} \qquad \mathcal{C}^x = \mathrm{Max}(\mathrm{Coord}(\Delta^{\mathrm{n}}) \cap {\downarrow} x) \setminus \{\perp\} .$$

Proof. We exclude the trivial case $x = \perp$. Note that by counting we obtain

$$K_1^{c(j)} = \bigcup_{i=1}^{i=j} K_i^x \quad K_2^{c(j)} = \bigcup_{i=j+1}^{i=n^x} K_i^x \quad K_1^{c_0} = \bigcup_{i=1}^{i=n_0^x} K_i^x \quad K_2^{c_0} = K_0^x .$$

Let $x \cdot \sigma_x \in \Lambda^n$. By Lemma 2.2 we have $\forall i \in \{1, \ldots, n^x\}$ that $\sigma[K_i] = I_i$ and as such we have for all $j \in \{1, \ldots, n_0^x - 1\}$

$$\sigma_x \left[K_1^{c(j)} \right] = \sigma_x \left[\bigcup_{i=1}^{i=j} K_i^x \right] = \bigcup_{i=1}^{i=j} \sigma_x \left[K_i^x \right] = \bigcup_{i=1}^{i=j} I_i^x = I_1^{c(j)} .$$

Analogously, $\sigma_x[K_2^{c(j)}] = I_2^{c(j)}$, $\sigma_x \left[K_1^{c_0}\right] = I_1^{c_0}$ and $\sigma_x \left[K_2^{c_0}\right] = I_2^{c_0}$. Thus, again by Lemma 2.2, for all $c(j) \in \mathcal{C}^x$ we have $c(j) \cdot \sigma_x \in \Lambda^n$ so x and $c(j)$ admit joint monotonization. Again, let $j \in \{1, \ldots, n_0^x - 1\}$. We have:

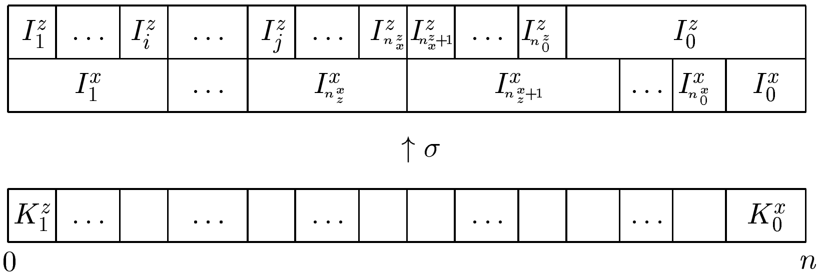
1. $c(j)^\Lambda_{\bar{n}_j^x} / c(j)^\Lambda_{\bar{n}_j^x + 1} = x^\Lambda_{\bar{n}_j^x} / x^\Lambda_{\bar{n}_j^x + 1}$;

2. $c(j)^\Lambda_i / c(j)^\Lambda_{i+1} = 1 \leq x^\Lambda_i / x^\Lambda_{i+1}$ for $i \in \{1, \ldots, \bar{n}_0^x - 1\} \setminus \{\bar{n}_j^x\}$;

3. $c(j)^\Lambda_i \cdot x^\Lambda_{i+1} = c(j)^\Lambda_i \cdot 0 \leq c(j)^\Lambda_{i+1} \cdot x^\Lambda_i$ for $i \in \{\bar{n}_0^x, \ldots, n - 1\}$.

Thus $c(j) \sqsubseteq x$ by Proposition 2.1. Analogously, in the case that $0 \in \mathrm{spec}(x)$ we have $c_0 \sqsubseteq x$. Thus, x is an upper bound for $\mathcal{C}^x$.

Let $z \in \Delta^n$ be such that $\forall c \in \mathcal{C}^x : c \sqsubseteq z$ and $\sigma_x, \sigma_z \in S(n)$ such that $x \cdot \sigma_x \in \Lambda^n$ and $z \cdot \sigma_z \in \Lambda^n$. First we construct $\sigma \in S(n)$ that monotonizes both x and z. Set $n_z^x := \sup(\{0\} \cup \{j \in \{1, \ldots, n^x\} \mid K_j^x \cap K_0^z = \emptyset\})$.



Assume $n_z^x \neq 0$ (if not, skip this paragraph). We have for $i \in I_1^{c(1)} = I_1^x$ and for $k \in I_2^{c(1)} = \{1, \ldots, n\} \setminus I_1^x$ that $c(1)_i > c(1)_k \neq 0$. Since $c(1) \sqsubseteq z$ we have by Lemma 2.4 that $\forall i \in I_1^x, \forall k \in \{1, \ldots, n\} \setminus I_1^x : z_i > z_k$ so $\sigma_z[K_1^x] = I_1^x$. By induction on $j \in \{1, \ldots, n_z^x\}$, since $c(j) \sqsubseteq z$ we have

$$\forall l \in \bigcup_{l=1}^{l=j-1} I_l^x \,, \forall i \in I_j^x \,, \forall k \in \{1, \ldots, n\} \setminus \bigcup_{l=1}^{l=j} I_l^x : z_l > z_i > z_k$$

so $\sigma_z[K_j^x] = I_j^x$. Let n_x^z be such that $\bigcup_{j=1}^{j=n_x^z} K_j^z = \bigcup_{j=1}^{j=n_z^x} K_j^x$. Setting

$$\forall j \in \{1, \ldots, n_x^z\} : \sigma[K_j^z] := \sigma_z[K_j^z] = I_j^z$$

we also obtain $\forall j \in \{1, \ldots, n_z^x\} : \sigma[K_j^x] = I_j^x$.

Next we set

- $\forall j \in \{n_x^z + 1, \ldots, n_0^z\} : \sigma[K_j^z] := \sigma_z[K_j^z] = I_j^z$
- $\sigma[K_{n_z^x+1}^x \cap K_0^z] := \sigma_x[K_{n_z^x+1}^x] \cap \sigma_z[K_0^z] = I_{n_z^x+1}^x \cap I_0^z$
- $\forall K_j^x \subseteq K_0^z : \sigma[K_j^x] := \sigma_x[K_j^x] = I_j^x$

Since $c(n_z^x + 1) \sqsubseteq z$ we obtain along the same lines as above that $\sigma[K_{n_z^x+1}^x] = I_{n_z^x+1}^x$ and $\sigma[K_0^z] = I_0^z$. Conclusively, σ monotonizes both x and z. We now verify the inequalities of Proposition 2.1 in order to prove that $x \sqsubseteq z$.

1. $x_{\bar{n}_j^x}^\Lambda / x_{\bar{n}_j^x+1}^\Lambda = c(j)_{\bar{n}_j^x}^\Lambda / c(j)_{\bar{n}_j^x+1}^\Lambda \leq z_{\bar{n}_j^x}^\Lambda / z_{\bar{n}_j^x+1}^\Lambda$ for $j \in \{1, \ldots, n_0^x - 1\}$;
2. $x_i^\Lambda / x_{i+1}^\Lambda = 1 \leq z_i^\Lambda / z_{i+1}^\Lambda$ for $i \in \{1, \ldots, \bar{n}_0^z - 1\} \setminus \{\bar{n}_j^x \,|\, j \in \{1, \ldots, n_0^x - 1\}\}$;
3. $x_i^\Lambda \cdot z_{i+1}^\Lambda = x_i^\Lambda \cdot 0 \leq x_{i+1}^\Lambda \cdot z_i^\Lambda$ for $i \in \{\bar{n}_0^z, \ldots, n-1\}$.

Conversely, $\mathcal{C}^x = \mathrm{Max}(\mathrm{Coord}(\Delta^n) \cap \downarrow x) \setminus \{\perp\}$ follows by Lemma 2.5 and the fact that $c(j)_1^{\mathrm{spec}} \cdot x_{j+1}^{\mathrm{spec}} = x_j^{\mathrm{spec}} \cdot c(j)_2^{\mathrm{spec}}$ maximizes those coordinates below x that are on the same axis. $\square$

One easily verifies that this decomposition is irreducible, that is, $\mathcal{C}^x$ is the infimum for inclusion of all finite $\mathcal{C} \subseteq \mathrm{Coord}(\Delta^n)$ with $x = \bigsqcup \mathcal{C}$. We also provide a characterization of sets of coordinates that arise in that way.

Proposition 3.4 $\{c(1), \ldots, c(m)\}$ *are the coordinates of some* $x \in \Delta^n$ *iff*

1. $m \leq n-1$
2. $x^1 \sqsupset \ldots \sqsupset x^m$ *where* $\forall j \in \{1, \ldots, m\} : c(j) \in x^j$*-axis*$\setminus\{\perp\}$
3. $c(j) = x^j \Rightarrow j = m$

Proof. For each $c(j)$ we obtain x^j such that $c(j) \in x^j$-axis by setting $\mathcal{I}^{x^j} = \mathcal{I}^{c(j)}$ and $\mathrm{spec}(x^j) = \{0,1\}$. (2.) is then easily verified. (1.) and (3.) are also obvious. Conversely, defining $\mathcal{I}^x$ by intersecting the sets $\mathcal{I}^{c(j)}$ for all $j \in \{1, \ldots, m\}$ and imposing $c(j)_1^{\mathrm{spec}} \cdot x_{j+1}^{\mathrm{spec}} = x_j^{\mathrm{spec}} \cdot c(j)_2^{\mathrm{spec}}$ we construct that $x \in \Delta^n$ which satisfies $\mathcal{C}^x = \{c(1), \ldots, c(m)\}$. $\square$

4 Isomorphisms

Theorem 4.1 (Isomorphisms) *Order-isomorphisms of* Δ^n *are in bijective correspondence with pairs consisting of*

- $\sigma \in S(n)$, ($\sim$ permuting the elements in $\mathrm{Max}(\Delta^n)$)
- $2^n - 2$ *order-isomorphisms of the unit interval.* ($\sim$ gauging each axis)

Proof. Let $h : \Delta^n \to \Delta^n$ be an order-isomorphism. We have $h(\perp) = \perp$. Next, since a maximal element has to be mapped on a maximal element this induces a permutation $\sigma \in S(n)$ via $\sigma(e_i) = h(e_i)$. This permutation extends to one on all $x \in \mathrm{Ir}_\perp(\Delta^n)$ since they are of the form $\bigwedge(\uparrow x \cap \max(\Delta^n))$ which on its turn extends by Proposition 3.2 to all coordinate axis (as a whole). For each coordinate axis set

$$f_x : x\text{-axis} \to x\text{-axis} :: y \mapsto h(y \cdot \sigma^{-1})$$

Since h is an order-isomorphism, so is f_x. Finally, the action on each $x \in \Delta^n$ is fully characterized via $x = \bigsqcup \mathcal{C}^x$.

Conversely, let $\sigma \in S(n)$ and let $\{f_x : x\text{-axis} \to x\text{-axis} \mid \mathrm{x} \in \mathrm{Ir}_\perp(\Delta^\mathrm{n})\}$ the $2^n - 2$ order-isomorphisms of the unit interval. We can define an order isomorphism of Δ^n by setting

$$h : \Delta^n \to \Delta^n :: y \mapsto \bigsqcup \{f_x(c(j)) \cdot \sigma \mid c(j) \in \mathcal{C}^y, c(j) \in x\text{-axis}\}\,.$$

Existence of the suprema follows from Proposition 3.4, bijectivity from Theorem 3.3 and monotonicity from $\mathcal{C}^x = \mathrm{Max}(\mathrm{Coord}(\Delta^\mathrm{n}) \cap \downarrow x) \setminus \{\perp\}$. Indeed, when $x \sqsubseteq y$ then this forces each $c(j) \in \mathcal{C}^x$ to have an upper bound in $\mathcal{C}^y$ since then $\downarrow x \subseteq \downarrow y$. Applying this argument to h^{-1} yields strictness. $\square$

Corollary 4.2 *The identity is the only order-isomorphism of* Δ^n *which preserves both* $\mathrm{Max}(\Delta^n)$ *and Shannon entropy (or any other map that is stricktly increasing on coordinate axis).*

Proof. By Theorem 4.1 it suffices to verify that Shannon entropy is strictly increasing on each coordinate axis. Its preservation forces all maps $\{f_x : x\text{-axis} \to x\text{-axis} \mid x \in \mathrm{Ir}_\perp(\Delta^n)\}$ to be identities. $\square$

So when Δ^n goes equipped with Shannon entropy it faithfully represents the set of all probability measures on $\{1, \ldots, n\}$. We can thus envision a probability space as an order-theoretic structure with a measure on it. This a priori order theoretic perspective is even more justified by the fact that we can construct the ordered probability space Δ^n from the algebra $\mathrm{P}(\{1, \ldots, n\})$ in purely order-theoretic manner.

5 Probability from algebra

We will reconstruct Δ^n from the algebra $A := \mathrm{P}(\{1, \ldots, n\})$ via a systematic procedure that allows extension to other algebraic logics, e.g. the lattice of closed subspaces of a Hilbert space [1, 4]. First we provide the underlying intuition, then the explicit mathematical operations applicable to any poset and finally the geometric and physical pictures that go with the construction.

Intuition. Let A be the algebra of properties of a system with n atomic properties $\{e_1, \ldots, e_n\}$. We will extend A in order to increase expressibility by allowing statements on truth of elements of the algebra to be "partial". Let Γ be a bounded chain with $\top$ as top and $\perp$ as bottom. Consider pairs in $A \times \Gamma$, where $(a, \top)$ stands for "certainty on truth of a", that is, ordinary truth of a, and $(a, \perp)$ stands for "no knowledge on truth of a", that is, truth of the trivial property. For $\gamma \in \Gamma \setminus \{\perp, \top\}$ the expressions (a, γ) allow true "partial knowledge on truth of a".

Next, we extend further by considering conjunctions of such expressions. In particular we consider those of the form $\vec{a} \cdot \vec{\gamma} := ((a_1, \gamma_1), \ldots, (a_{n-1}, \gamma_{n-1}))$ for $a_1 \sqsupset \ldots \sqsupset a_{n-1}$ a maximal chain in the poset obtained by removing the trivial and the absurd property from A. The pointwise ordering

$$\vec{a} \cdot \vec{\gamma} \sqsubseteq \vec{a} \cdot \vec{\varphi} \;\Leftrightarrow\; \forall i \in \{1, \ldots, n-1\} : \gamma_i \sqsubseteq \varphi_i$$

then arises naturally as an implicative relation. However, since strong statements are above weak ones with respect to the order of Γ, we will reverse the order of A such that the respective interpretations of the order in Γ and A become compatible. Thus, the atomic properties of A are now maximal.

The extremal elements of Γ require some additional considerations. For all $a \in A$ the expressions $(a, \perp)$ are all equivalent to the trivial property. Thus they are "void" so they can be omitted. Given both $(a, \top)$ and $a \sqsupset b$ then this forces $\gamma = \top$ in (b, γ). Moreover, in that case $(b, \top)$ is redundant in the presence of $(a, \top)$ and can also be omitted.

We translate this now into mathematics.

Formal procedure. Let A be a bounded poset. Let Γ be a bounded chain.

1. Denote by $A^*_{0,1}$ the poset obtained by removing top and bottom from A and by reversing the order.

2. Let $\mathrm{MChain}(A^*_{0,1})$ be all maximal chains $\vec{a} = \{a_1 \sqsupset \ldots \sqsupset a_{n-1}\}$ in $A^*_{0,1}$. In benefit of transparency of the argument we assume that all these chains have equal finite length. (This is however not a necessity!)

3. Denote by $\mathrm{Cl}_\top(\Gamma^{n-1})$ the set of all Γ-valued tuples $\vec{\gamma} = (\gamma_1, \ldots, \gamma_{n-1})$ subjected to the closure[2]
$$\forall i < j \in \{1, \ldots, n-1\} : \gamma_i = \top \;\Rightarrow\; \gamma_j = \top\,. \tag{7}$$

4. Set $[A^*_{0,1}, \Gamma] := \{\vec{a} \cdot \vec{\gamma} \mid \vec{a} \in \mathrm{MChain}(A^*_{0,1})\,, \vec{\gamma} \in \mathrm{Cl}_\top(\Gamma^{n-1})\}\,.$

5. Introduce the pointwisely induced relation
$$\vec{a} \cdot \vec{\gamma} \sqsubseteq \vec{b} \cdot \vec{\varphi} \iff \vec{a} = \vec{b} \text{ and } \forall i \in \{1, \ldots, n-1\} : \gamma_i \sqsubseteq \varphi_i\,.$$

6. Set
$$I(\vec{\gamma}) = \{i \in \{1, \ldots, n-1\} \mid \gamma_i \notin \{\perp, \top\}\}$$
and
$$\iota(\vec{\gamma}) = \inf\{i \in \{1, \ldots, n-1\} \mid \gamma_i = \top\}\,.$$
Let $\overline{[A^*_{0,1}, \Gamma]}$ be the set of equivalence classes in $[A^*_{0,1}, \Gamma]$ obtained for the equivalence relation
$$\vec{a} \cdot \vec{\gamma} = \vec{b} \cdot \vec{\varphi} \iff \vec{\gamma} = \vec{\varphi} \text{ and } (i \in I(\vec{\gamma}) \cup \{\iota(\vec{\gamma})\} \Rightarrow a_i = b_i)\,.$$

7. Finally, $\overline{[A^*_{0,1}, \Gamma]}$ inherits the relation $\sqsubseteq$ on $[A^*_{0,1}, \Gamma]$, explicitly,
$$\vec{a} \cdot \vec{\gamma} \sqsubseteq \vec{a} \cdot \vec{\varphi} \implies [\vec{a} \cdot \vec{\gamma}] \sqsubseteq [\vec{a} \cdot \vec{\varphi}]\,.$$

Proposition 5.1 $\left(\overline{[A^*_{0,1}, \Gamma]}, \sqsubseteq\right)$ *is a poset with a bottom.*

Proof. We have to prove anti-symmetry and transitivity of $\sqsubseteq$ on $\overline{[A^*_{0,1}, \Gamma]}$.

Anti-symmetry. Let $\vec{a} \cdot \vec{\gamma} \sqsubseteq \vec{a} \cdot \vec{\varphi}$ and $\vec{b} \cdot \vec{\gamma} \sqsupseteq \vec{b} \cdot \vec{\varphi}$ with $[\vec{a} \cdot \vec{\gamma}] = [\vec{b} \cdot \vec{\gamma}]$ and $[\vec{a} \cdot \vec{\varphi}] = [\vec{b} \cdot \vec{\varphi}]$. We must then for all $i \in \{1, \ldots, n-1\}$ both have $\gamma_i \sqsubseteq \varphi_i$ and $\varphi_i \sqsubseteq \gamma_i$ from which $\vec{a} \cdot \vec{\gamma} = \vec{a} \cdot \vec{\varphi}$ and thus $[\vec{a} \cdot \vec{\gamma}] = [\vec{a} \cdot \vec{\varphi}]$ follows.

Transitivity. Let $\vec{a} \cdot \vec{\gamma}^- \sqsubseteq \vec{a} \cdot \vec{\gamma}$ and $\vec{b} \cdot \vec{\gamma} \sqsubseteq \vec{b} \cdot \vec{\gamma}^+$ with $[\vec{a} \cdot \vec{\gamma}] = [\vec{b} \cdot \vec{\gamma}]$. We have to prove that $[\vec{a} \cdot \vec{\gamma}^-] \sqsubseteq [\vec{b} \cdot \vec{\gamma}^+]$. We define $\vec{c} \in \mathrm{MChain}(A^*_{0,1})$ as follows. For $i \in I(\vec{\gamma}) : c_i := a_i = b_i$, for $i \in \{\iota(\vec{\gamma}), \ldots, n-1\} : c_i := a_i$ and in all other cases, that is $\gamma_i = \perp$, we set $c_i := b_i$. Since $\gamma_i^- \sqsubseteq \gamma_i$ implies $\gamma_i = \perp \Rightarrow \gamma_i^- = \perp$ and $\gamma_i \sqsubseteq \gamma_i^+$ implies $\gamma_i = \top \Rightarrow \gamma_i^+ = \top$ it respectively follows that $[\vec{c} \cdot \vec{\gamma}^-] = [\vec{a} \cdot \vec{\gamma}^-]$ and $[\vec{c} \cdot \vec{\gamma}^+] = [\vec{a} \cdot \vec{\gamma}^+]$. Thus, since $\vec{c} \cdot \vec{\gamma}^- \sqsubseteq \vec{c} \cdot \vec{\gamma}^+$ due to $\gamma_i^- \sqsubseteq \gamma_i \sqsubseteq \gamma_i^+$ for all $i \in \{1, \ldots, n-1\}$ we obtain $[\vec{a} \cdot \vec{\gamma}^-] \sqsubseteq [\vec{b} \cdot \vec{\gamma}^+]$.

Finally, choosing $\vec{a}$ arbitrary in $\mathrm{MChain}(A^*_{0,1})$ and setting $\vec{\gamma} = (\perp, \ldots, \perp)$, we obtain $[\vec{a} \cdot \vec{\gamma}]$ as the bottom of $\overline{[A^*_{0,1}, \Gamma]}$. □

[2] $\mathrm{Cl}_\top$ acts as a closure operator on the complete lattice Γ^{n-1} by imposing eq.(7), and thus, $\mathrm{Cl}_\top(\Gamma^{n-1})$ is itself a complete lattice. For all $n \geq 2$ we actually have $\Lambda^n \cong \mathrm{Cl}_\top([0,1]^{n-1})$ so monotone states constitute complete lattices. One can also show that Δ^n admits all non-empty meets and that any subset of Δ^n which is bounded from above has a supremum [2].

Theorem 5.2 (Construction of classical states)

$$\overline{\left[\mathrm{P}(\{1,\ldots,n\})^*_{0,1}\,,[0,1]\right]} \cong \Delta^n$$

Proof. Assume $\xi : [0,1] \to [1,\infty]$ to be a fixed order isomorphism. Let $\vec{a}\cdot\vec{\gamma} \in [\mathrm{P}(\{1,\ldots,n\})^*_{0,1}\,,[0,1]]$. We can define a set $\mathcal{C}^{\vec{a}\cdot\vec{\gamma}}$ of coordinates as follows. For each $a_i \in \vec{a}$ such that $i \in I(\vec{\gamma}) \cup \{\iota(\vec{\gamma})\}$ define $c(i) \in \mathrm{Coord}(\Delta^n)$ such that $\mathcal{I}^{c(i)} = (I^i, \{1,\ldots,n\} \setminus I^i)$ where I^i is implicitly defined by $a_i = \bigvee\{e_j \mid j \in I^i\}$, and by setting $c^i_1/c^i_2 = \xi(\gamma_i)$ whenever $\gamma_i \neq 1$ and $c^i_2 = 0$ otherwise. The set $\mathcal{C}^{\vec{a}\cdot\vec{\gamma}} = \{c^i \mid i \in I(\vec{\gamma}) \cup \{\iota(\vec{\gamma})\}\}$ satisfies the conditions in Proposition 3.4 and as such $\mathcal{C}^{\vec{a}\cdot\vec{\gamma}} = \mathcal{C}^x$ for $x = \bigsqcup \mathcal{C}^{\vec{a}\cdot\vec{\gamma}}$. For $\vec{a}\cdot\vec{\gamma}, \vec{b}\cdot\vec{\varphi} \in [\mathrm{P}(\{1,\ldots,n\})^*_{0,1}\,,[0,1]]$ we have $\mathcal{C}^{\vec{a}\cdot\vec{\gamma}} = \mathcal{C}^{\vec{b}\cdot\vec{\varphi}}$ iff $\vec{a}\cdot\vec{\gamma} \sim \vec{b}\cdot\vec{\varphi}$ in the above defined equivalence relation on $[\mathrm{P}(\{1,\ldots,n\})^*_{0,1}\,,[0,1]]$. Due to uniqueness of the decomposition in coordinates (Theorem 3.3) we obtain an injective correspondence between $\overline{[\mathrm{P}(\{1,\ldots,n\})^*_{0,1}\,,[0,1]]}$ and Δ^n and by Proposition 3.4 it follows that it is also surjective.

We now show that this correspondence also preserves the order. It follows from the definition of $\sqsubseteq$ that for $[\vec{a}\cdot\vec{\gamma}], [\vec{b}\cdot\vec{\varphi}] \in \overline{[\mathrm{P}(\{1,\ldots,n\})^*_{0,1}\,,[0,1]]}$ we have $[\vec{a}\cdot\vec{\gamma}] \sqsubseteq [\vec{b}\cdot\vec{\varphi}]$ iff there exists $\vec{c} \in \mathrm{MChain}(\mathrm{P}(\{1,\ldots,n\})^*_{0,1})$ such that $\vec{c}\cdot\vec{\gamma} \in [\vec{a}\cdot\vec{\gamma}]$ and $\vec{c}\cdot\vec{\varphi} \in [\vec{b}\cdot\vec{\varphi}]$ and such that $\vec{c}\cdot\vec{\varphi} \sqsubseteq \vec{b}\cdot\vec{\varphi}$. Moreover,

1. $\mathrm{MChain}(\mathrm{P}(\{1,\ldots,n\})^*_{0,1})$ and $S(n)$ are in bijective correspondence via

$$\forall i \in \{1,\ldots,n-1\} : a_i = \bigvee\{e_j \mid j \in \sigma[\{1,\ldots,i\}]\}$$

   and in particular, existence of $\vec{c} \in \mathrm{MChain}(\mathrm{P}(\{1,\ldots,n\})^*_{0,1})$ with $\vec{c}\cdot\vec{\gamma} \in [\vec{a}\cdot\vec{\gamma}]$ and $\vec{c}\cdot\vec{\varphi} \in [\vec{b}\cdot\vec{\varphi}]$ coincides with existence of $\sigma \in S(n)$ which monotonizes both $x = \bigsqcup \mathcal{C}^{\vec{a}\cdot\vec{\gamma}}$ and $y = \bigsqcup \mathcal{C}^{\vec{b}\cdot\vec{\varphi}}$.

2. Due to $c^i_1/c^i_2 = \xi(\gamma_i)$ for $\gamma_i \neq 1$ and $c^i_2 = 0$ for $\gamma_i = 1$, the pointwisely defined order for $\vec{\gamma}$ and $\vec{\varphi}$ induces eq.(1) for $x = \bigsqcup \mathcal{C}^{\vec{a}\cdot\vec{\gamma}}$ and $y = \bigsqcup \mathcal{C}^{\vec{b}\cdot\vec{\varphi}}$.

Explicit verification of the above completes the proof. □

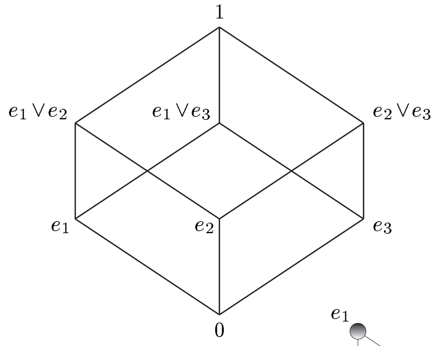
This theorem also extends to the quantum case. Explicitly, the **construction of quantum states** is given by

$$\overline{\left[\mathbb{L}(n)^*_{0,1}\,,[0,1]\right]} \cong \Omega^n$$

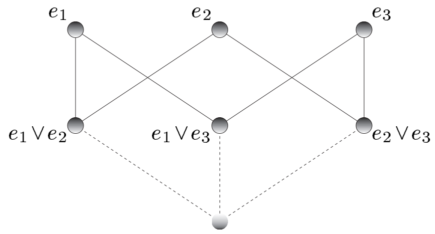
where $\mathbb{L}(n)$ is the lattice of closed subspaces of a n-dimensional Hilbert space and Ω^n are the quantum probability measures on this Hilbert space ordered according to [3] §3. We however omit the lengthy proof here.

The geometric picture. We illustrate the above for the case of $n = 3$.

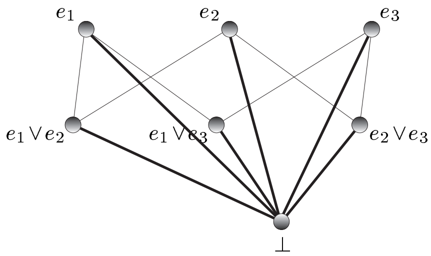
$A =$



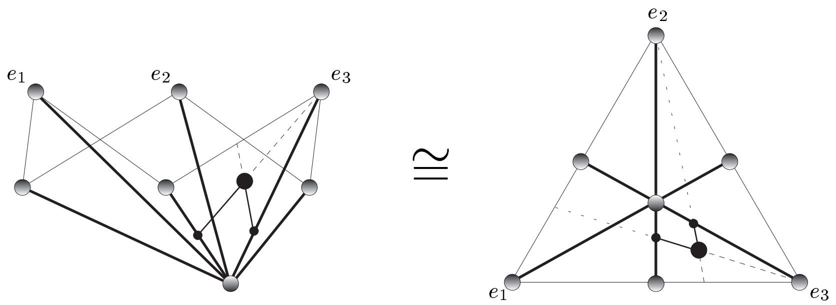
$A^*_{0,1} =$



Pairing elements of $A^*_{0,1}$ with those of Γ creates increasing "lines" which all rise from a common source, namely the "void" statement (denoted as $\perp$).



Finally, the formation of lists for all chains in $\mathrm{MChain}(A^*_{0,1})$ fills the regions enclosed by the corresponding lines resulting in a triangle.



Note how the formation of lists of pairs (= conjunctive) corresponds with the generation of points as joins of coordinates ($\sim$ reversed order).

The physical picture. Entropic geometry is not merely a geometry of lines but one of directed lines. The triangle or the tetrahedron are not merely convex geometric objects but combine both (static $\sim$ *space-like*) form and (dynamic $\sim$ *time-like*) directedness. As an example, the center of the triangle is a special point from which directed lines emerge, which stand for the increase of certainty, or, if one prefers, the decrease of entropy, or, if one prefers, the flow of time (either forward or backward). In such a dynamic perspective where the lines Γ obtain the connotation of "flow", the bounds $\perp$ and $\top$ obtain the connotation of *initiation* and *termination*. The fact that the 4-tuple $(A, \Gamma, \perp, \top)$ generates an entropic geometry by the above presented systematic formal procedure can thus be interpreted as

$$\text{Entropic Geometry} = \text{Logic} + \text{Flow} + \text{Initiation} + \text{Termination}\,.$$

6 Acknowledgements

I thank Keye Martin for sharing ideas on presentation of these results. I also thank Samson Abramsky and Prakash Panangaden for logistic, Dusko Pavlovic for recreational and Rhada Jagadeesan for gastronomic support, and also for their constructive feedback on [3].

References

[1] G. Birkhoff and J. von Neumann (1936) "The Logic of Quantum Mechanics", *Annals of Mathematics*, **37**, 823–843, 1936.

[2] B. Coecke (2002) "Lattices of Probabilities", Unpublished notes.

[3] B. Coecke and K. Martin (2002) *A Partial Order on Classical and Quantum States*, Research Report PRG-RR-02-07, Oxford University Computing Laboratory.

[4] B. Coecke, D. J. Moore, and A. Wilce, Eds. (2002) *Current Research in Operational Quantum Logic: Algebras, Categories, Languages*, Kluwer.

[5] K. Martin (2000) *A Foundation for Computation*, Ph.D. Thesis, Department of Mathematics, Tulane University.

[6] F. P. Ramsey (1926) *Truth and Probability*, Reprinted in: R. B. Braithwaite, editor, *The Foundations of Mathematics and Other Logical Essays*, Ch. 7, Routledge, 1931.

[7] D. S. Scott (1970) *Outline of a Mathematical Theory of Computation*, Tech. Monograph PRG-2, Oxford University Computing Laboratory.