PROBABILITY AS ORDER

c[B]o[B]ecke — Oxford

lattices@dimacs

July 2003

- **A partial order on probability measures**
- **Partial order = Logic + Partiality**
- **Probability = Partial order + Entropy**
- **Quantum states from quantum logic**

References:

- B. Coecke and K. Martin. A partial order on classical and quantum states. PRG-RR-02-07 Oxford University Computing Laboratory.
  `web.comlab.ox.ac.uk/oucl/publications/tr/rr-02-07.html`
- B. Coecke. Entropic geometry from logic. ENTCS **69** (MFPS'03).
  `arXiv:quant-ph/0212065`

Computing applications:

- S. Abramsky. Qualitative and quantitative theories of information. 2003. Research proposal.
- K. Martin. Epistemic motion in quantum searching. 2003.
  `web.comlab.ox.ac.uk/oucl/people/keye.martin.html`

Probability space:

$$\Delta^n \; := \left\{ (x_1, \ldots, x_n) \in [0,1]^n \;\middle|\; \sum_{i=1}^{i=n} x_i = 1 \right\}$$

E.g.: the **line** Δ^2, the **triangle** Δ^3, the **tetrahedron** Δ^4.

Bayesian/knowledge/epistemic updating:

	1		i		$n{+}1$	1		i		$n{+}1$
knowledge x ::	x_1		x_i		x_{n+1}	$\frac{x_1}{1-x_i}$		0		$\frac{x_{n+1}}{1-x_i}$
knowledge y ::	y_1		y_i		y_{n+1}	$\frac{y_1}{1-y_i}$		0		$\frac{y_{n+1}}{1-y_i}$
public knowledge ::								$\mathbf{0}$		

Then, for $x \in \Delta^{n+1}$ with $x_i < 1$:

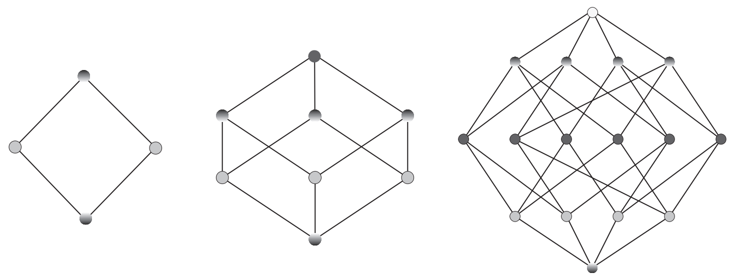
$$p_i(x) = \frac{1}{1-x_i}(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_{n+1}) \in \Delta^{n+1}\,.$$

or

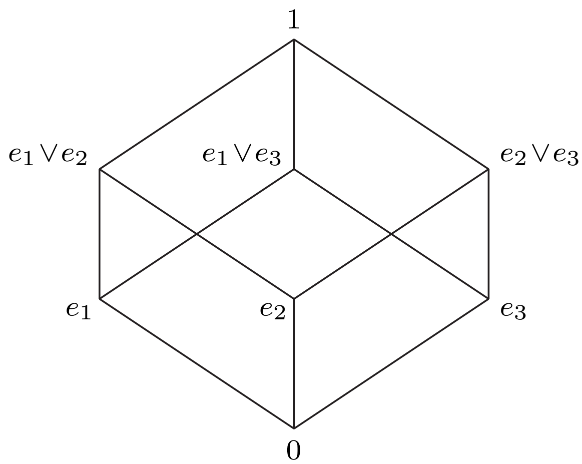
$$p_i(x) = \frac{1}{1-x_i}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}) \in \Delta^{n}\,.$$

LOGIC (AS ALGEBRA):

i. Boolean algebra $\simeq$ classical logic



$\simeq$ powerset $\simeq$ generated as disjuncts of atomic propositions:



Update $\simeq$ eliminate atoms in disjunctions:

$$p_i(a := e_i \vee e_j \vee e_k) = e_j \vee e_k$$

This update is monotone. Moreover, trivially:

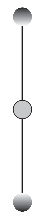
$$a \leq b \iff \forall i : \, p_i(a) \leq p_i(b) \, .$$

This update coincides with Bayesian update when setting:

$$\begin{cases} x_i^a = 0 & e_i \not\leq a \\ x_i^a = k & e_i \leq a \end{cases}$$

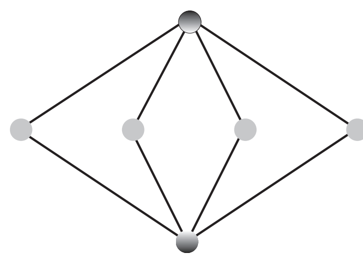
with $k^{-1} = |\{e_i \leq a\}|$.

ii. Heyting algebra $\simeq$ intuitionistic logic



$\simeq$ implicative by definition in terms of $(\wedge\,,\rightarrow)$-fragment.

iii. Quantum property lattice $\simeq$ quantum "logic"



"This is not a logic" since it has no $(\wedge\,,\rightarrow)$-fragment. But it has physical significance as the subsets of the state space that guaranty definite outcomes in measurements ($\simeq$ **properties**).

Partial order:

> $\simeq$ inclusion with respect to obtainable outcomes in a measurement when the system has that property
>
> $\simeq$ knowledge order on what one would obtain when performing a measurement.

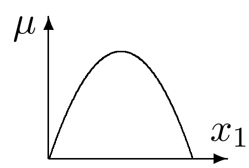
PARTIALTY, APPROXIMATION AND ENTROPY:

i. Probability expresses partiality of knowledge. E.g. finding a ball in a box; distributions in statistical physics.

ii. Entropy expresses partiality of knowledge for probability **quantitatively**:

$$\mu : \Delta^n \to [0,1] :: x \mapsto -\sum_{i=1}^{n} x_i \log x_i$$

Shannon entropy on Δ^2:



- Physical processes: Entropy increases by thermodynamics' 2nd law. Entropy is the heart of irreversibility.
- Computational processes: Aim at decreasing entropy.

Entropy is a number $\simeq$ total order $\simeq$ everything compares.

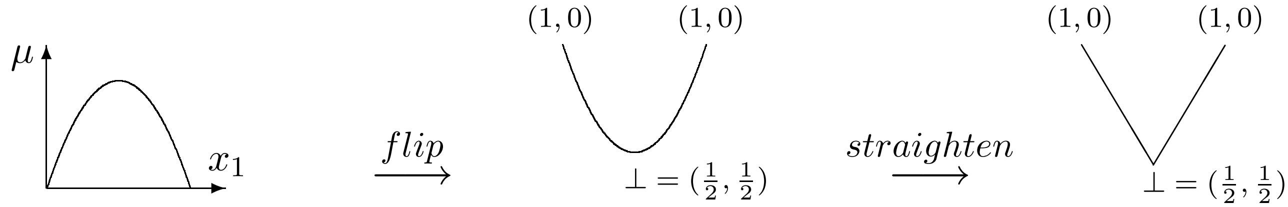
iii. Domains express partiality of knowledge **qualitatively**:

$$x \sqsubseteq y \iff x \text{ has only part of } y\text{'s knowledge}\,.$$

Total elements are approximated by partial ones as limits $\bigsqcup A$.

Question: Does probabilistic partial knowledge have a more refined structure which enriches entropy?

A PARTIAL ORDER ON DISTRIBUTIONS:



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

So there exists a canonical partial order on Δ^2 which goes "hand in hand" with Shannon entropy. Thus, **we combined quantitative and qualitative theories of information**.

Question. Does this case extends to higher dimensions?

Answer. Yes. And the solution $(\Delta^n, \sqsubseteq)$ is non-trivial.

Theorem. Let $n \geq 2$. Then, $(\Delta^n, \sqsubseteq)$ is a dcpo with

$$\mathrm{Max}(\Delta^n) = \{e \in \Delta^n \mid \mathrm{spec}(x) = \{0, 1\}\} \;\&\; \perp = (1/n, \ldots, 1/n)\,,$$

which admits the notions *partiality* and *approximation*. Shannon entropy is a measure of content on $(\Delta^n, \sqsubseteq)$ in a very strong sense ($\simeq$ in the sense of Keye Martin's PhD thesis).

Theorem. The above theorem extends to a partial order on quantum mixed states; so does everything that follows below.

General intuition on $x \sqsubseteq y$:

- "*State* y is more *informative* than *state* x".

or, epistemically:

- "*Observer* y has more *knowledge* than *observer* x".

E.g. $(1/3\,,\,1/3\,,\,1/3) \sqsubseteq (1/2\,,\,1/2\,,\,0) \sqsubseteq (1\,,\,0\,,\,0)$.

Definition. Given the $n = 2$ case introduce $\sqsubseteq$ by:

$$x \sqsubseteq y \iff \forall i : p_i(x) \sqsubseteq p_i(y)\,,$$

that is, in words:

1. If x has less knowledge than y, then, after update w.r.t. common knowledge, $p_i(x)$ has less knowledge than $p_i(y)$.
2. This operational behavior w.r.t. update defines $\sqsubseteq$.

Definition.' Let $n \geq 2$. For $x, y \in \Delta^n$, we have $x \sqsubseteq y$ iff

$$\exists \sigma \in S(n) \text{ such that } x \cdot \sigma, y \cdot \sigma \in \Lambda^n$$

and $\forall i \in \{1, \ldots, n-1\}$:

$$(x \cdot \sigma)_i (y \cdot \sigma)_{i+1} \leq (x \cdot \sigma)_{i+1} (y \cdot \sigma)_i\,.$$

where

- $\Lambda^n :=$ decreasing distributions in Δ^n, i.e., $x_i \geq x_{i+1}$
- $S(n) := \{\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\} \mid \text{permutation}\}$

There exists a partial order on Δ^n which canonically arises from envisioning probability distributions as informative objects, and which goes hand in hand with Shannon entropy.

LOGIC FROM PROBABILITY

Set:

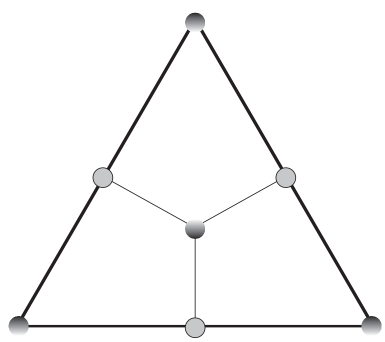
$$\mathrm{Ir}(D) := \left\{ x \in D \;\middle|\; \bigwedge (\uparrow x \cap \mathrm{Max}(D)) = x \right\}$$

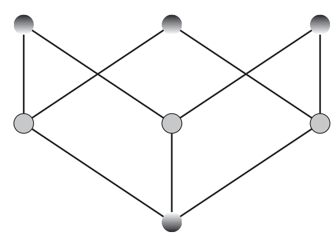
Then:

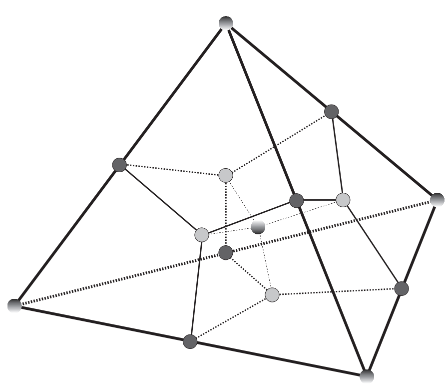
$$\mathrm{Ir}(\Delta^n) \cup \{0\} \;\cong\; \mathrm{P}(\{1, \ldots, n\})$$

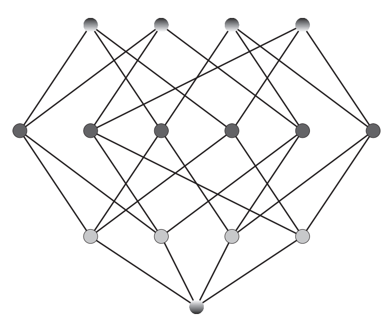
$$\mathrm{Ir}(\Omega^n) \cup \{0\} \;\cong\; \mathbb{L}^n$$

where $\mathbb{L}^n$ is the lattice of subspaces of a n-dimensional Hilbert space, i.e., the Birkhoff – von Neumann quantum logic.









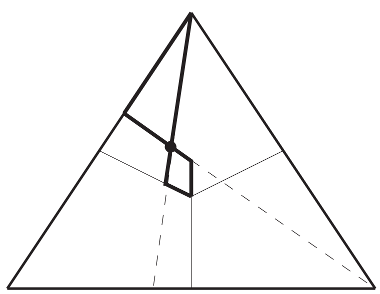
Where did logic come in?

Lemma 1. *If $x \sqsubseteq y$ in Δ^n then*

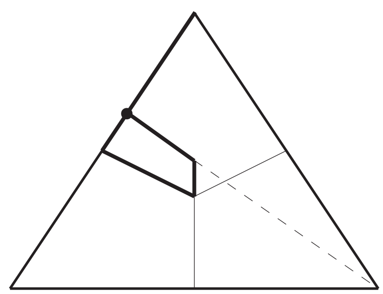
$$x_i = 0 \;\Rightarrow\; y_i = 0 \qquad \& \qquad y_i = y_j > 0 \;\Rightarrow\; x_i = x_j$$

zero-values/degeneration		001
non-degenerated non-zero values	e.g.	$0\frac{1}{3}\frac{2}{3}$
degenerated non-zero values		$\frac{1}{3}\frac{1}{3}\frac{1}{3}$

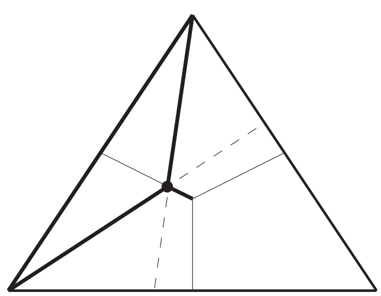
Non-degenerated:

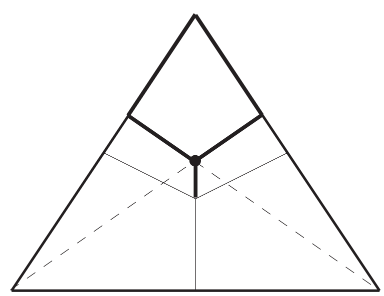


Zero-values:

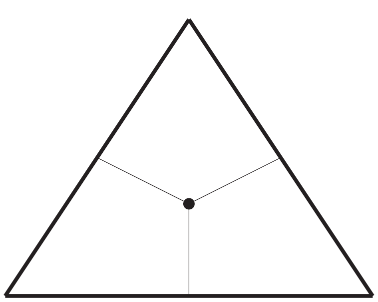


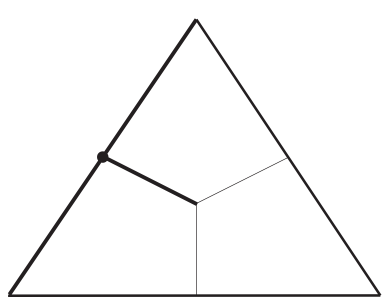
Non-zero degenerated:

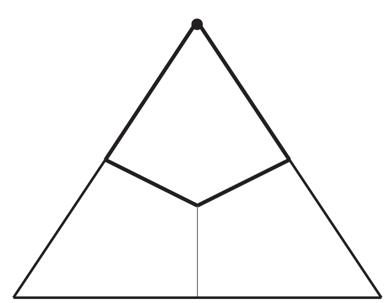




Extremal degeneration:

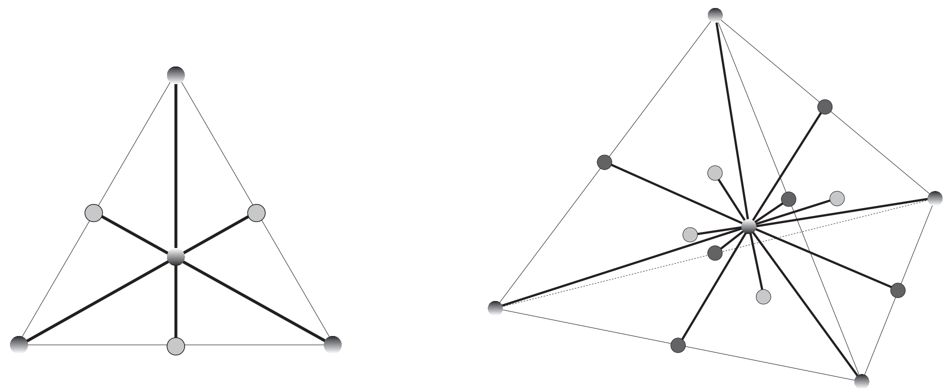






PROBABILITY FROM LOGIC

Coordinates:=



Theorem. Coordinates and coordinate-axis are order-theoretic.

Theorem. States $x \in \Delta^n$ and irreducible coordinate tuples $\mathcal{C}^x$ are in bijective order-theoretic correspondence:

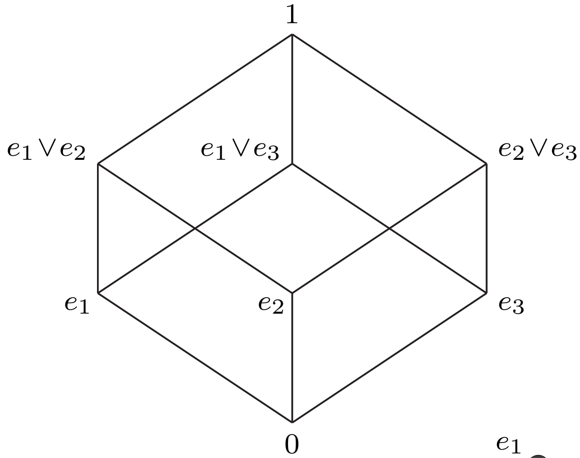
$$x = \bigsqcup \mathcal{C}^x \qquad \text{and} \qquad \mathcal{C}^x = \mathrm{Max}(\mathrm{Coord}(\Delta^{\mathrm{n}}) \cap \downarrow x) \setminus \{\perp\}\,.$$

where $\{c(1), \ldots, c(m)\}$ is a coordinate-tuple iff

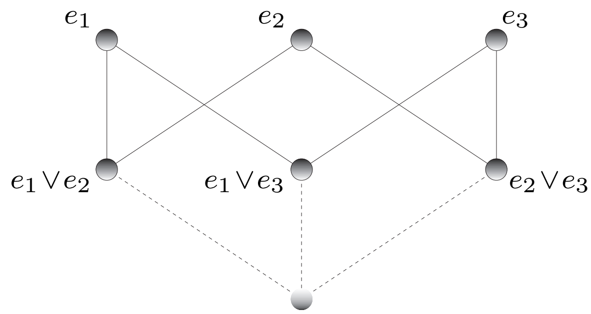
1. $m \leq n - 1\,;$
2. $x^1 \sqsupset \ldots \sqsupset x^m$ for $c(j) \in x^j$-axis $\setminus \{\perp\}\,;$
3. $c(j) = x^j \Rightarrow j = m\,.$

Theorem. The partially ordered sets $(\Delta^n, \sqsubseteq)$ and $(\Omega^n, \sqsubseteq)$ can be constructed in purely order theoretic fashion respectively from $\mathrm{P}(\{1, \ldots, n\})$ and $\mathbb{L}^n$.

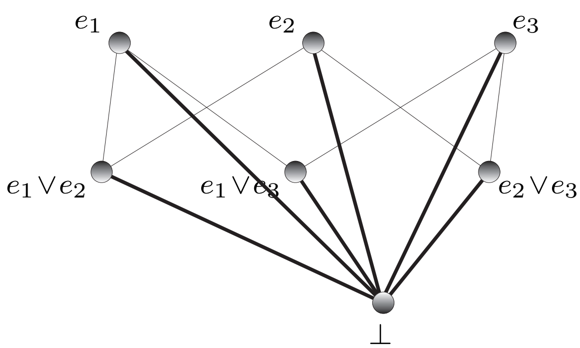
$A =$



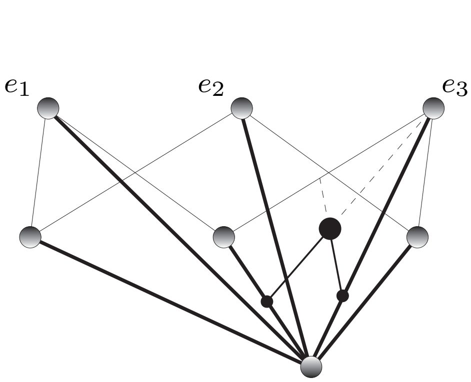
$A^*_{0,1} =$



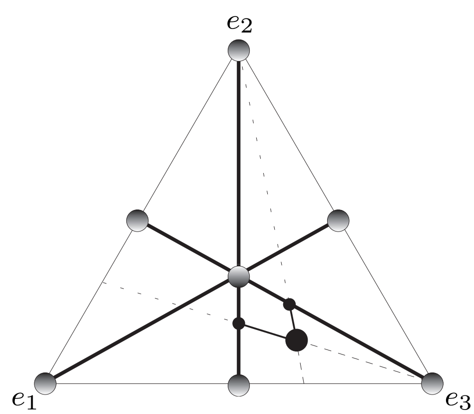
Introduce partial knowledge on $A^*_{0,1}$ as "lines":



Introduce spans (= conjunctive lists of partial statements):



$\cong$



Formally, $A :=$ bounded poset & $\Gamma :=$ bounded chain.

1. $A^*_{0,1} :=$ remove top and bottom and reverse order;
2. $\mathrm{Cl}_\top(\Gamma^{n-1}) :=$ all Γ-valued tuples $\vec{\gamma} = (\gamma_1, \ldots, \gamma_{n-1})$ s.t.
$$\forall i < j \in \{1, \ldots, n-1\} : \gamma_i = \top \ \Rightarrow\ \gamma_j = \top\,;$$
3. Set $[A^*_{0,1}, \Gamma] := \{\vec{a}\cdot\vec{\gamma} \mid \vec{a} \in \mathrm{MaxChain}(A^*_{0,1})\,, \vec{\gamma} \in \mathrm{Cl}_\top(\Gamma^{n-1})\}\,;$
4. $\vec{a}\cdot\vec{\gamma} \sqsubseteq \vec{b}\cdot\vec{\varphi} \iff \vec{a} = \vec{b} \wedge \forall i \in \{1, \ldots, n-1\} : \gamma_i \sqsubseteq \varphi_i\,;$
5. $\overline{[A^*_{0,1}, \Gamma]} :=$ equivalence classes in $[A^*_{0,1}, \Gamma]$ for
$$\vec{a}\cdot\vec{\gamma} \sim \vec{b}\cdot\vec{\varphi} \iff \vec{\gamma} = \vec{\varphi} \wedge (i \in I(\vec{\gamma}) \cup \{\iota(\vec{\gamma})\} \Rightarrow a_i = b_i)\,,$$
$$I(\vec{\gamma}) := \{i \mid \gamma_i \notin \{\bot, \top\}\} \ \ \&\ \ \iota(\vec{\gamma}) := \inf\{i \mid \gamma_i = \top\}\,;$$
$\overline{[A^*_{0,1}, \Gamma]}$ inherits the relation $\sqsubseteq$ on $[A^*_{0,1}, \Gamma]$.

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

Next, set $\Gamma \simeq [0,1]$ as a partial order.

Theorem 1. (Construction of classical states)
$$\left(\overline{\left[\mathrm{P}(\{1, \ldots, n\})^*_{0,1}\,, \Gamma\right]}\,, \sqsubseteq\right) \cong (\Delta^n, \sqsubseteq)\,.$$

Theorem 2. (Construction of quantum states)
$$\left(\overline{\left[(\mathbb{L}^n)^*_{0,1}\,, \Gamma\right]}\,, \sqsubseteq\right) \cong (\Omega^n, \sqsubseteq)\,.$$

Proposition 2. (Construction of monotone states)
$$\left(\overline{\left[D^*_{0,1}\,, \Gamma\right]}\,, \sqsubseteq\right) \cong (\Lambda^n, \sqsubseteq)$$
for D a $(n+1)$-chain.

REPRESENTING PROBABILITY

Theorem. Order-isomorphisms of $(\Delta^n, \sqsubseteq)$ are in bijective correspondence with pairs consisting of

- $\sigma \in S(n)$ $\sim$labeling the maximal elements
- $2^n - 2$ order-isomorphisms of $[0,1]$ $\sim$gauging each axis

Theorem. The identity is the only order-isomorphism of $(\Delta^n, \sqsubseteq)$ which preserves both $\mathrm{Max}(\Delta^n)$ and Shannon entropy.

$$\begin{tikzcd}
[0,1] \arrow[r, "id"] & [0,1] \\
D^n \arrow[u, "\mu"] \arrow[r, "h"] & (\Delta^n, \sqsubseteq) \arrow[u, "\mu"']
\end{tikzcd}$$

Digest:

$$\mathsf{Probability} = \mathsf{Qualitative\ Entropy} + \mathsf{Quantitative\ Entropy}$$

$$\mathsf{Qualitative\ Entropy} := \mathsf{Logic} + \mathsf{Partiality\ of\ Knowledge}$$

Note:

$$\mathsf{Qualitative\ Entropy} \simeq \mathsf{Qualitative\ Probability}$$

THE QUANTUM MEASUREMENT PROCESS

Each quantum measurement M formally represents as a spectral decomposition in **projectors**:

$$\mathrm{M} := 1 \cdot \mathrm{P}_1 + 2 \cdot \mathrm{P}_2 + 3 \cdot \mathrm{P}_3 + \ldots + i \cdot \mathrm{P}_i + \ldots$$

$$(\mathrm{P}_i : \mathcal{H}^n \to \mathcal{H}^n \quad \text{s.t.} \quad \mathrm{P}_i \circ \mathrm{P}_i = \mathrm{P}_i = \mathrm{P}_i^\dagger \quad \& \ i \neq j \Rightarrow \mathrm{P}_i \perp \mathrm{P}_j)$$

The actual **measurement process** is:

1. **A value** i **is delivered to the observer**;
2. **The state of the system undergoes change** P_i.

Thus two transitions take place:

$$\begin{cases} \textsf{observer's knowledge :} & \perp \ \mapsto \ i \\ \textsf{state of system :} & \psi \ \mapsto \ \mathrm{P}_i(\psi) \end{cases}$$

(for each i with some probability)

The **action** $\mathrm{P}_i : \mathcal{H}^n \to \mathcal{H}^n$ is captured in different structures:

- **The (non-distributive) lattice of properties** $\mathbb{L}^n$

$$A \in \mathbb{L}^n \iff \exists \mathrm{M}, \exists i : A = \{\psi \in \mathcal{H}^n \mid \mathrm{P}_i(\psi) = \psi\}$$

  Those (pure) states which “guarantee behavior”.

- **The manual of tests** $\mathcal{M}^n$

$$T \in \mathcal{M}^n \iff \exists \mathrm{M} : T = \{\mathrm{P}_i(\psi) \mid \psi \in \mathcal{H}^n, i \in \sigma(\mathrm{M})\}$$

  Those (pure) states which are “jointly reachable”.

Mixed state := represents behavioral probabilities under measurement, that is, for each measurement M it provides

$$\mathrm{Prob}(\psi, \mathrm{M}, i) \in [0, 1]\,.$$

Both $\mathbb{L}^n$ and $\mathcal{M}^n$ suffice to characterize all mixed states:

- All measures on $\omega : \mathbb{L}^n \to [0, 1]$ which are additive on mutually orthogonal elements (Gleason 1957);
- All measures on $\omega : \bigcup \mathcal{M}^n \to [0, 1]$ which sum to one on each test (Foulis and Randall 1970);

In both cases we exactly get the density matrices Ω^n.

Surveys: B. Coecke, D. Moore and A. Wilce (Eds.): *Current research in operational quantum logic: algebras, categries, languages.* Kluwer. 2000. Introduction + references at `arXiv:quant-ph/0008019`

In a Boolean algebra orthogonality is captured by order:

$$a \perp b \Leftrightarrow a \wedge b = 0\,.$$

$\mathbb{L}^n$ admits many orthocomplementations, i.e., antitone involutions $\,' : L \to L$ s.t. $a \wedge a' = 0 \;\&\; a \vee a' = 1$. Orthogonality:

$$a \perp b \Leftrightarrow a \leq b'\,.$$

Mixed quantum states are measures $\omega : \mathbb{L}^n \to [0,1]$ s.t.

$$a \perp b \;\Rightarrow\; \omega(a \vee b) = \omega(a) + \omega(b)\,.$$

By Gleason's theorem these are in bijective correspondence with the density matrices Ω^n. We can envision a constructor $\underline{\mathrm{Val}}[-]$, acting on all posets with orthogonality $(D, \perp)$:

$$\underline{\mathrm{Val}}\,[(\mathrm{P}(\{1,\ldots,n\}), (-)^c)] \cong (\Delta^n, \sqsubseteq)$$

$$\underline{\mathrm{Val}}\,[(\mathbb{L}^n, (-)')] \cong (\Omega^n, \sqsubseteq)\,,$$

The $\underline{\mathrm{EntGeom}}$ construction enables to produce an isomorphic copy of $(\Omega^n, \sqsubseteq)$ without specification of orthogonality:

$$\underline{\mathrm{EntGeom}}\,[(n+1) - \mathrm{chain}] \cong (\Lambda^n, \sqsubseteq)\,,$$

$$\underline{\mathrm{EntGeom}}\,[\mathrm{P}(\{1,\ldots,n\})] \cong (\Delta^n, \sqsubseteq)\,,$$

$$\underline{\mathrm{EntGeom}}\,[\mathbb{L}^n] \cong (\Omega^n, \sqsubseteq)\,.$$

Moreover:

$$\underline{\mathrm{Ir}} \circ \underline{\mathrm{EntGeom}} = \underline{\mathrm{id}}\,.$$