New Foundations for Information Theory: The transition to quantum information theory

David Ellerman

University of California-Riverside

Abstract

"For in the general we must note, That whatever is capable of a competent Difference, perceptible to any Sense, may be a sufficient Means whereby to express the Cogitations." (John Wilkins 1641)
****

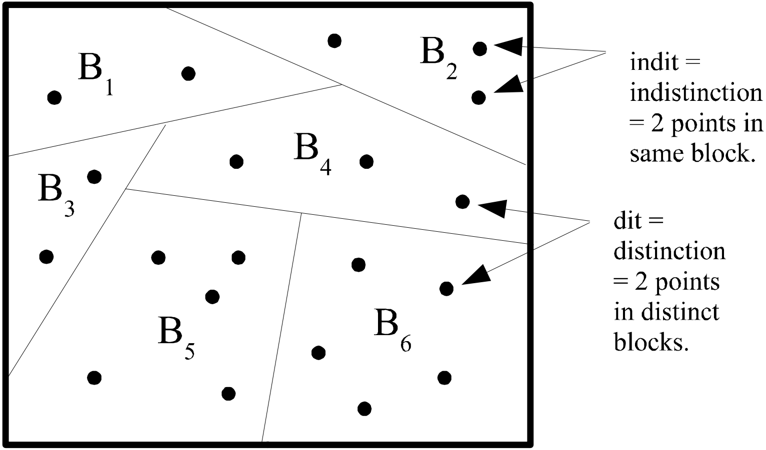
"So information really is a very useful abstraction. It is the notion of distinguishability abstracted away from what we are distinguishing, or from the carrier of information. ...And we ought to develop a theory of information which generalizes the theory of distinguishability to include these quantum properties... ." (Charles H. Bennett 2003)

Duality of Subsets and Partitions: I

f: X→Y	Subsets	Partitions
General case	Image f(X) = Subset of codomain Y	Inverse-image $f^{-1}(Y)$ = Partition on domain X
Definition of a function	Binary relation $f \subseteq X \times Y$ that transmits elements	+Binary relation $f \subseteq X \times Y$ that reflects distinctions
Basic duality	Elements of Subset	Distinctions of Partition

Duality of Subsets and Partitions &
Duality of Elements and Distinctions ("Its" & "Dits")

Duality of Subsets and Partitions: II



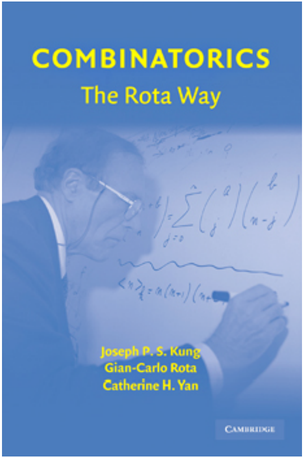
Partition $\pi = \{B_1, ..., B_6\}$ on set $U = \{u_1, ..., u_n\}$.

Duality of Subsets and Partitions: III

Table 1	Subset Logic	Partition Logic
Logic of...	Subsets $S \subseteq U$	Partitions π on U
'Elements' (its or dits)	*Elements* u of a subset S	*Distinctions* (u,u') of a partition π
All 'elements'	Universe set U (all elements)	Discrete partition **1** (all dits)
No 'elements'	Empty set $\emptyset$ (no elements)	Indiscrete partition **0** (no dits)
Partial order on...	$S \subseteq T$ = Inclusion of elements	Refinement $\sigma \preceq \pi$ of partitions = inclusion of ditsets: $\text{dit}(\sigma) \subseteq \text{dit}(\pi)$
Formula variables	Subsets of U	Partitions on U
Logical operations $\cup, \cap, \Rightarrow, \ldots$	Operations on subsets	Operations on partitions
Propositional interp. of $\Phi(\pi,\sigma,\ldots)$	Subset $\Phi(\pi,\sigma,\ldots)$ contains an element u.	Partition $\Phi(\pi,\sigma,\ldots)$ makes a distinction (u,u').
Valid formula $\Phi(\pi,\sigma,\ldots)$	$\Phi(\pi,\sigma,\ldots) = U$ for any subsets $\pi,\sigma,\ldots$ of any U ($\|U\| \geq 1$), i.e., contains all elements u.	$\Phi(\pi,\sigma,\ldots) = \mathbf{1}$ for any partitions $\pi,\sigma,\ldots$ on any U ($\|U\| \geq 2$), i.e., makes all distinctions (u,u').

Dual Logics: Boolean subset logic of subsets and partition logic

Duality of Subsets and Partitions: IV



"Lattice of partitions plays for information the role that the Boolean algebra of subsets plays for size or probability."

$$\frac{\text{Information}}{\text{Partitions}} \approx \frac{\text{Probability}}{\text{Subsets}}$$

Gian-Carlo Rota

Duality of Subsets and Partitions: V

- Rota: "Probability is a measure on the Boolean algebra of events" that gives quantitatively the "intuitive idea of the size of a set", so we may ask by "analogy" for some measure to capture a property for a partition like "what size is to a set." Rota goes on to ask:

  How shall we be led to such a property? We have already an inkling of what it should be: it should be a measure of information provided by a random variable. Is there a candidate for the measure of the amount of information? (Rota's Fubini Lecture)

- Elements : Subsets :: Dits : Partitions, so

$$\#\text{elements (“size of subset”)} \approx \#\text{dits (“size of partition”)}.$$

The logical theory of information: I

- New foundations of information theory starts with sets, not probabilities.

  Information theory must precede probability theory, and not be based on it. By the very essence of this discipline, the foundations of information theory have a finite combinatorial character. [Kolmogorov, A. N. 1983]

- The notion of information-as-distinctions thus starts with the *set of distinctions*, the *information set*, of a partition $\pi = \{B, B', ...\}$ on a finite set U where that set of distinctions (dits) is:

$$\operatorname{dit}(\pi) = \{(u, u') : \exists B, B' \in \pi, B \neq B', u \in B, u' \in B'\}.$$

The logical theory of information: II

- The ditset of a partition is the complement in $U \times U$ of the equivalence relation associated with the partition π.
- Given any probability measure $p : U \rightarrow [0,1]$ on $U = \{u_1, ..., u_n\}$ which defines $p_i = p(u_i)$ for $i = 1, ..., n$, the *product measure* $p \times p : U \times U \rightarrow [0,1]$ has for any $S \subseteq U \times U$ the value of:

$$p \times p(S) = \sum_{(u_i,u_j)\in S} p(u_i)\, p(u_j) = \sum_{(u_i,u_j)\in S} p_i p_j.$$

- The *logical entropy* of π is the product measure of its ditset:

$$h(\pi) = p \times p(\mathrm{dit}(\pi)) = \sum_{(u_i,u_j)\in \mathrm{dit}(\pi)} p_i p_j = 1 - \sum_{B\in\pi} p(B)^2.$$

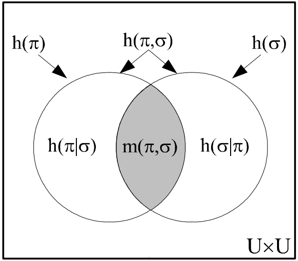
The logical theory of information: III

Table 2	Logical Probability Theory	Logical Information Theory
'Outcomes'	Elements $u \in U$ finite	Distinctions $(u,u') \in U \times U$ finite
'Events'	Subsets $S \subseteq U$	Ditsets $\mathrm{dit}(\pi) \subseteq U \times U$
Equiprobable outcomes	$\Pr(S) = \|S\|/\|U\|$ = logical probability of event S	$h(\pi) = \|\mathrm{dit}(\pi)\|/\|U \times U\|$ = logical entropy of partition π
Point probabilities	$\Pr(S) = \Sigma\{p_j : u_j \in S\} = p(S)$ = logical prob. of event S	$h(\pi) = \Sigma\{p_j p_k : (u_j,u_k) \in \mathrm{dit}(\pi)\}$ = logical entropy of π
Interpretation	$\Pr(S)$ = one draw probability of getting an element from S	$h(\pi)$ = two draw probability (w/replacement) of getting a distinction of π

Compound logical entropies: I

- Given partitions $\pi = \{B_1, ..., B_I\}, \sigma = \{C_1, ..., C_J\}$ on U, the *information set* or *ditset for their join* is: $\operatorname{dit}(\pi \vee \sigma) = \operatorname{dit}(\pi) \cup \operatorname{dit}(\sigma) \subseteq U \times U$.
- Given probabilities $p = \{p_1, ..., p_n\}$, the *joint logical entropy* is: $h(\pi, \sigma) = h(\pi \vee \sigma) = p \times p(\operatorname{dit}(\pi) \cup \operatorname{dit}(\sigma)) = 1 - \sum_{i,j} p(B_i \cap C_j)^2$.
- The infoset for the *conditional logical entropy* $h(\pi|\sigma)$ is the difference of ditsets, and thus: $h(\pi|\sigma) = p \times p(\operatorname{dit}(\pi) - \operatorname{dit}(\sigma))$.
- The infoset for the *logical mutual information* $m(\pi, \sigma)$ is the intersection of ditsets, so: $m(\pi, \sigma) = p \times p(\operatorname{dit}(\pi) \cap \operatorname{dit}(\sigma))$.

Compound logical entropies: II



Logical entropies Venn diagram

- *Information algebra* $\mathcal{I}(\pi,\sigma)$ = Boolean subalgebra of $\wp(U \times U)$ generated by ditsets and their complements.

Deriving the Shannon entropies from the logical entropies: I

- All the Shannon entropies can be derived from the logical definitions by a uniform transformation–since they are just two different ways to measure distinctions.
- Canonical case of n equiprobable elements, $p_i = \frac{1}{n}$, the logical entropy of $\mathbf{1}_U = \{B_1, ..., B_n\}$ where $B_i = \{u_i\}$ with $p = \{\frac{1}{n}, ..., \frac{1}{n}\}$ is:

$$h(p(B_i)) = \frac{|U \times U - \Delta|}{|U \times U|} = \frac{n^2 - n}{n^2} = 1 - \frac{1}{n} = 1 - p(B_i).$$

- General case $\pi = \{B_1, ..., B_m\}$ is average of this dit-count $1 - p(B_i)$:

Deriving the Shannon entropies from the logical entropies: II

$$h(\pi) = \sum_i p(B_i)(1 - p(B_i)).$$

- Canonical case of 2^n equiprobable elements and discrete partition so $p(B_i) = \frac{1}{2^n}$, the minimum number of binary partitions ("yes-or-no questions") or "bits" it takes to uniquely determine or *encode* each distinct element or block is n, so the Shannon-Hartley entropy is:

$$H(p(B_i)) = n = \log_2(2^n) = \log_2\left(\frac{1}{1/2^n}\right) = \log_2\left(\frac{1}{p(B_i)}\right).$$

- General case is average of this bit-count $\log_2\left(\frac{1}{p(B_i)}\right)$:

Deriving the Shannon entropies from the logical entropies: III

$$H(\pi) = \sum_i p(B_i) \log_2 \left(\frac{1}{p(B_i)} \right).$$

- *Dit-Bit Transform*: express any logical entropy concept (joint, conditional, or mutual) as average of dit-counts $1 - p(B_i)$, and then substitute the bit-count $\log\left(\frac{1}{p(B_i)}\right)$ to obtain the corresponding formula as defined by Shannon.

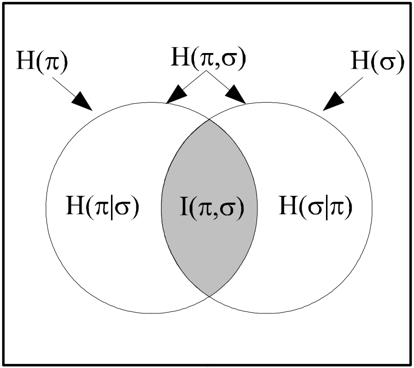
Deriving the Shannon entropies from the logical entropies: IV

Table 3	The Dit-Bit Transform: $1-p(B_i) \rightsquigarrow \log(1/p(B_i))$
Entropy	$h(\pi) = \Sigma_i\, p(B_i)(1-p(B_i))$ $H(\pi) = \Sigma_i\, p(B_i)(\log(1/p(B_i)))$
Joint Entropy	$h(\pi,\sigma) = \Sigma_{i,j} p(B_i \cap C_j)(1-p(B_i \cap C_j))$ $H(\pi,\sigma) = \Sigma_{i,j} p(B_i \cap C_j)\log(1/p(B_i \cap C_j))$
Conditional Entropy	$h(\pi\vert\sigma) = \Sigma_{i,j} p(B_i \cap C_j)(1-p(B_i \cap C_j)) - \Sigma_j p(C_j)(1-p(C_j))$ $H(\pi\vert\sigma) = \Sigma_{i,j} p(B_i \cap C_j)\log(1/p(B_i \cap C_j)) - \Sigma_j p(C_j)\log(1/p(C_j))$
Mutual Information	$m(\pi,\sigma) = \Sigma_{i,j} p(B_i \cap C_j)[(1-p(B_i)) + (1-p(C_j)) - (1-p(B_i \cap C_j))]$ $I(\pi,\sigma) = \Sigma_{i,j} p(B_i \cap C_j)[(\log(1/p(B_i)) + (\log(1/p(C_j)) - (\log(1/p(B_i \cap C_j))]$

Deriving the Shannon entropies from the logical entropies: V

- The dit-bit transform is linear in the sense of preserving plus and minus, so the Shannon formulas satisfy the same Venn diagram formulas in spite of not being a measure (in the sense of measure theory):

Deriving the Shannon entropies from the logical entropies: VI



Venn diagram ‘mnemonic’ for Shannon entropies

Logical entropy via density matrices: I

- All this will carry over to quantum logical entropy using density matrices.
- Let $U = \{u_1, ..., u_n\}$ be the sample space with the point probabilities $p = (p_1, ..., p_n)$. An event $S \subseteq U$ has the probability $p(S) = \sum_{u_j \in S} p_j$.
- For any event S with $p(S) > 0$, let $|S\rangle = \frac{1}{\sqrt{p(S)}} (\chi_S(u_1)\sqrt{p_1}, ..., \chi_S(u_n)\sqrt{p_n})^t$ which is a normalized column vector in $\mathbb{R}^n$ where $\chi_S : U \to \{0,1\}$ is the characteristic function for S, and let $\langle S|$ be the corresponding row vector. Since $|S\rangle$ is normalized, $\langle S|S\rangle = 1$.
- Then the *density matrix* representing the event S is the $n \times n$ symmetric real matrix:

Logical entropy via density matrices: II

$$\rho(S) = |S\rangle\langle S| = \begin{cases} \frac{1}{p(S)}\sqrt{p_j p_k} \text{ for } u_j, u_k \in S \\ 0 \text{ otherwise} \end{cases}.$$

- Then $\rho(S)^2 = |S\rangle\langle S|S\rangle\langle S| = \rho(S)$ so borrowing language from QM, $|S\rangle$ is said to be a *pure* state or event.
- Given any partition $\pi = \{B_1, ..., B_I\}$ on U, its density matrix is the average of the block density matrices:

$$\rho(\pi) = \sum_i p(B_i)\,\rho(B_i).$$

- Then $\rho(\pi)$ represents the *mixed* state, experiment, or lottery where the event B_i occurs with probability $p(B_i)$.

Logical entropy via density matrices: III

Example

For the throw of a fair die, $U = \{u_1, u_3, u_5, u_2, u_4, u_6\}$ (where u_j represents the number j coming up), the density matrix $\rho(\mathbf{0}_U)$ is the "pure state" 6×6 matrix with each entry being $\frac{1}{6}$.

$$\rho(\mathbf{0}_U) = \begin{array}{cc} \begin{bmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix} & \begin{array}{c} u_1 \\ u_3 \\ u_5 \\ u_2 \\ u_4 \\ u_6 \end{array} \end{array}.$$

Logical entropy via density matrices: IV

- Nonzero off-diagonal entries represents indistinctions or indits of partition $\mathbf{0}_U$, or in quantum terms as "coherences" where all 6 "eigenstates" cohere together in a pure "superposition" state. All pure states have logical entropy of zero, i.e., $h(\mathbf{0}_U) = 0$ (i.e., no dits).

Example (continued)

Now classify or "measure" the elements by parity (odd or even) partition (observable) $\pi = \{B_{odd}, B_{even}\} = \{\{u_1, u_3, u_5\}, \{u_2, u_4, u_6\}\}$. Mathematically, this is done by the Lüders mixture operation where P_{odd} and P_{even} are the projections to the odd or even components:

Logical entropy via density matrices: V

$$P_{odd}\rho\left(\mathbf{0}_U\right)P_{odd} + P_{even}\rho\left(\mathbf{0}_U\right)P_{even} = \sum_{i=1}^{m} p\left(B_i\right)\rho\left(B_i\right) = \rho\left(\pi\right).$$

$$\rho\left(\mathbf{0}_U\right) = \begin{bmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix} \rightsquigarrow$$

$$\begin{bmatrix} 1/6 & 1/6 & 1/6 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 0 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 0 & 1/6 & 1/6 & 1/6 \end{bmatrix} = \rho\left(\pi\right)$$

Logical entropy via density matrices: VI

Theorem (Basic)

The increase in logical entropy due to a Lüders mixture operation is the sum of amplitudes squared of the non-zero off-diagonal entries of the beginning density matrix that are zeroed in the final density matrix.

Proof.

Since for any density matrix ρ, $\operatorname{tr}\left[\rho^2\right] = \sum_{i,j} \left|\rho_{ij}\right|^2$,
$h\left(\rho\left(\pi\right)\right) - h\left(\rho\left(\mathbf{0}_U\right)\right) = \left(1 - \operatorname{tr}\left[\rho\left(\pi\right)^2\right]\right) - \left(1 - \operatorname{tr}\left[\rho\left(\mathbf{0}_U\right)^2\right]\right) =$
$\operatorname{tr}\left[\rho\left(\mathbf{0}_U\right)^2\right] - \operatorname{tr}\left[\rho\left(\pi\right)^2\right] = \sum_{i,j} \left|\rho_{ij}\left(\mathbf{0}_U\right)\right|^2 - \sum_{i,j} \left|\rho_{ij}\left(\pi\right)\right|^2.$ □

Logical entropy via density matrices: VII

Example (continued)

In comparison with the matrix $\rho(\mathbf{0}_U)$ of all entries $\frac{1}{6}$, the entries that got zeroed in $\rho(\mathbf{0}_U) \rightsquigarrow \rho(\pi)$ correspond to the distinctions created in the transition $\mathbf{0}_U = \{U\} \rightsquigarrow \pi = \{\{u_1, u_3, u_5\}, \{u_2, u_4, u_6\}\}$. Increase in logical entropy = $h(\pi) - h(\mathbf{0}_U) = 2 \times 9 \times \left(\frac{1}{6}\right)^2 = \frac{18}{36} = \frac{1}{2}$. Usual calculations: $h(\pi) = 1 - 2 \times \left(\frac{1}{2}\right)^2 = \frac{1}{2}$ and $h(\mathbf{0}_U) = 1 - 1^2 = 0$.

Logical entropy via density matrices: VIII

- Since a projective measurement's effect on a density matrix *is* the Lüders mixture operation, that means that the effects of the measurement is the above-described "making distinctions" by decohering or zeroing certain coherence terms in the density matrix, and the sum of the absolute squares of the coherences that were decohered is the change in the logical entropy.

Generalization to Quantum Info. Theory: I

[Information] is the notion of distinguishability abstracted away from what we are distinguishing, or from the carrier of information. ...And we ought to develop a theory of information which generalizes the theory of distinguishability to include these quantum properties... . [Bennett, 2003]

- *Qubit* = pair of states definitely distinguishable by any observable, e.g., distinction of self-adjoint operator $\sum_k kP_{[u_k]}$.
- In general, a qubit (or qudit?) needs to be r*elativized to an observable*–classically entropy is the entropy *of a partition*.
- A *qubit of* F is a pair $(u_k, u_{k'})$ in the eigenbasis definitely distinguishable by F, i.e., $\phi(u_k) \neq \phi(u_{k'})$, distinct eigenvalues.

Generalization to Quantum Info. Theory: II

Table 4a (w/o probs.)	'Classical' Logical Info. Theory	Quantum Logical Info. Theory
Universe	$U = \{u_1,\ldots,u_n\}$	Orthonormal basis $\{u_i\}$ Hilbert space V
Attribute/Observable	Real-valued 'random' variables $f, g: U \to \mathbb{R}$	Commuting self-adjoint operators F, G $\{u_i\}$ O.N. basis of simult. eigenvectors
Values	Image values $\{\phi_i\}_{i\in I}$ of f Image values $\{\gamma_j\}_{j\in J}$ of g	Eigenvalues $\{\phi_i\}_{i\in I}$ of F Eigenvalues $\{\gamma_j\}_{j\in J}$ of G
Partitions / Direct-sum decompositions	Inverse-image $\pi = \{f^{-1}(\phi_i)\}_{i\in I}$ Inverse-image $\sigma = \{g^{-1}(\gamma_j)\}_{j\in J}$	Eigenspace Direct-sum Decomp. F Eigenspace Direct-sum Decomp. G
Distinctions	Dits of π: $(u_k,u_{k'})\in U^2$, $f(u_k) \neq f(u_{k'})$ Dits of σ: $(u_k,u_{k'})\in U^2$, $g(u_k) \neq g(u_{k'})$	Qubits of F: $u_k\otimes u_{k'}\in V\otimes V$, $\phi(u_k) \neq \phi(u_{k'})$ Qubits of G: $u_k\otimes u_{k'}\in V\otimes V$, $\gamma(u_k) \neq \gamma(u_{k'})$
Information sets/spaces	$\mathrm{dit}(\pi) \subseteq U\times U$ $\mathrm{dit}(\sigma) \subseteq U\times U$	[qubit(F)] = subspace gen. by qubits of F [qubit(G)] = subspace gen. by qubits of G
Joint = Conditional = Mutual =	$\mathrm{dit}(\pi) \cup \mathrm{dit}(\sigma) \subseteq U\times U$ $\mathrm{dit}(\pi) - \mathrm{dit}(\sigma) \subseteq U\times U$ $\mathrm{dit}(\pi) \cap \mathrm{dit}(\sigma) \subseteq U\times U$	$[\mathrm{qubit}(F) \cup \mathrm{qubit}(G)] \subseteq V\otimes V$ $[\mathrm{qubit}(F) - \mathrm{qubit}(G)] \subseteq V\otimes V$ $[\mathrm{qubit}(F) \cap \mathrm{qubit}(G)] \subseteq V\otimes V$

Generalization to Quantum Info. Theory: III

Table 4b (w/ probs.)	'Classical' Logical Info. Theory	Quantum Logical Info. Theory
Probability dist.	Pure state density matrix, e.g., $\rho(\mathbf{0}_U)$	Pure state density matrix $\rho(\psi)$
Product prob. dist.	$p \times p$ on $U \times U$	$\rho(\psi) \otimes \rho(\psi)$ on $V \otimes V$
Logical entropies	$h(\mathbf{0}_U) = 1 - \mathrm{tr}[\rho(\mathbf{0}_U)^2] = 0$ $h(\pi) = p \times p(\mathrm{dit}(\pi))$ $h(\pi,\sigma) = p \times p(\mathrm{dit}(\pi) \cup \mathrm{dit}(\sigma))$ $h(\pi \vert \sigma) = p \times p(\mathrm{dit}(\pi) - \mathrm{dit}(\sigma))$ $m(\pi,\sigma) = p \times p(\mathrm{dit}(\pi) \cap \mathrm{dit}(\sigma))$	$h(\rho(\psi)) = 1 - \mathrm{tr}[\rho(\psi)^2] = 0$ $h(F{:}\psi) = \mathrm{tr}[P_{[\mathrm{qubit}(F)]}\rho(\psi) \otimes \rho(\psi)]$ $h(F,G{:}\psi) = \mathrm{tr}[P_{[\mathrm{qubit}(F) \cup \mathrm{qubit}(G)]}\rho(\psi) \otimes \rho(\psi)]$ $h(F \vert G{:}\psi) = \mathrm{tr}[P_{[\mathrm{qubit}(F) - \mathrm{qubit}(G)]}\rho(\psi) \otimes \rho(\psi)]$ $m(F,G{:}\psi) = \mathrm{tr}[P_{[\mathrm{qubit}(F) \cap \mathrm{qubit}(G)]}\rho(\psi) \otimes \rho(\psi)]$
Venn diagram from being prob. measure	$h(\pi,\sigma) = h(\pi \vert \sigma) + h(\sigma \vert \pi) + m(\pi,\sigma)$ $h(\pi) = h(\pi \vert \sigma) + m(\pi,\sigma)$	$h(F,G) = h(F \vert G) + h(G \vert F) + m(F,G)$ $h(F) = h(F \vert G) + m(F,G)$
Interpretation	$h(\pi)$ = two-draw prob. of getting a dit of π, i.e., different f values.	$h(F{:}\psi)$ = prob. in two indep. F meas. of ψ in getting different eigenvalues.
Lüders Mixture	$\rho(\pi) = \Sigma_i P_{B_i}\rho(\mathbf{0}_U)P_{B_i}$ and $h(\pi) = p \times p(\mathrm{dit}(\pi)) = 1 - \mathrm{tr}[\rho(\pi)^2]$	$\rho'(\psi) = \Sigma_i P_{\phi_i}\rho(\psi)P_{\phi_i}$ and $h(F{:}\psi) = 1 - \mathrm{tr}[\rho'(\psi)^2]$
Thm. on L-entropy and measurement.	$h(\pi)$ = sum of squares of terms zeroed in measurement operation: $\rho(\mathbf{0}_U) \rightarrow \rho(\pi)$.	$h(F{:}\psi)$ = sum of absol. squares of terms zeroed in the measurement operation: $\rho(\psi) \rightarrow \rho'(\psi)$.

Basic Theorem about measurement

- Nonzero off-diagonal terms in density matrix $\rho(\psi)$ are called "coherences"–like indistinctions in classical case.
- Measurement creates distinctions, i.e., turn coherences into 'decoherences'–classically, turn indistinctions into distinctions.
- **Basic Theorem**: Measure of distinctions created in measuring pure state ψ by F = sum of absolute squares of off-diagonal terms zeroed (i.e., coherences that were decohered) in measurement = logical entropy increase, e.g., $h(F:\psi) = h(\rho'(\psi)) - h(\rho(\psi))$ = probability that two independent measurements of ψ will yield a qubit of F.