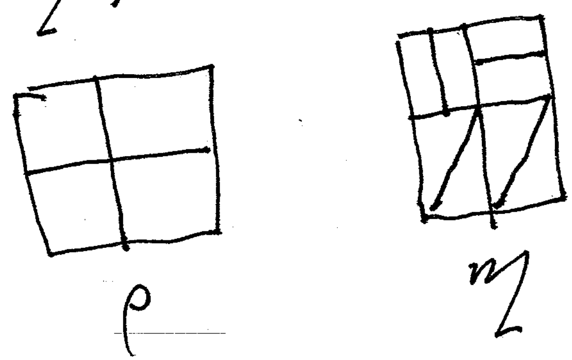
Entropy, Cont.

· To bring the dynamics into play we need more information about partitions and finite algebras

<u>DEF</u> If ρ, η are two finite partitions of $(X, \mathcal{B}, \mu)$

$\rho \leq \eta$ means that each element of ρ is a union of elements of η, so η is a refinement of ρ



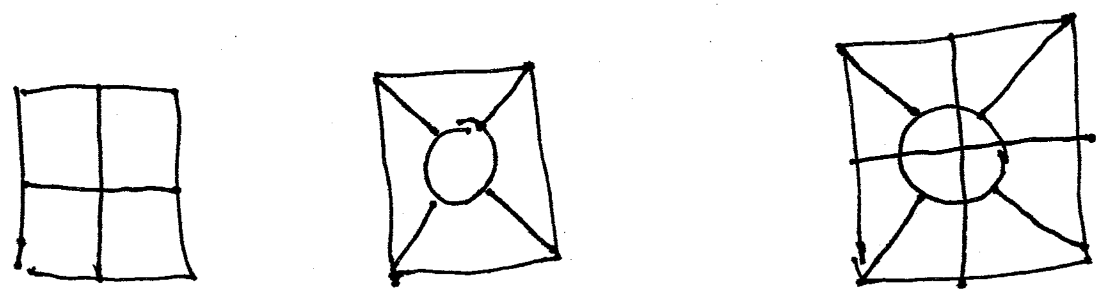
<u>Remark</u> (1) $\rho \leq \eta \iff \mathcal{A}(\rho) \subseteq \mathcal{A}(\eta)$

(2) $\mathcal{A} \subseteq \mathcal{C} \iff \rho(\mathcal{A}) \leq \rho(\mathcal{C})$

<u>DEF</u>: If $\rho = \{A_1, \dots, A_n\}$, $\eta = \{C_1, \dots, C_k\}$

their join is

$$\rho \vee \eta = \{A_i \wedge C_j : 1 \le i \le n,\ 1 \le j \le k\}$$



ρ $\qquad$ η $\qquad$ $\rho \vee \eta$

- For finite subalgebras, $\mathcal{A} \vee \mathcal{C}$ is all sets of the form $A \wedge C$, $A \in \mathcal{A}$, $C \in \mathcal{C}$ and is the smallest subalgebra of $\mathcal{B}$ containing $\mathcal{A}$ and $\mathcal{C}$
- $\rho(\mathcal{A} \vee \mathcal{C}) = \rho(\mathcal{A}) \vee \rho(\mathcal{C})$, $\mathcal{A}(\rho \vee \eta) = \mathcal{A}(\rho) \vee \mathcal{A}(\eta)$.

(3)

DEF: If $\rho = \{A_1, \dots, A_k\}$ and $f : (X, \mathcal{B}, \mu) \circlearrowleft$ is a mpt $\Rightarrow \forall n \geq 0$, $f^{-n}(\rho) = \{f^{-n}(A_1), \dots, f^{-n}(A_k)\}$ is also a finite partition. For finite subalgebras, $\mathcal{A}$,

$$f^{-n}(\mathcal{A}) = \{f^{-n}(A) : A \in \mathcal{A}\}$$

FACTS Since f^{-n} preserves set theoretic operations

$$f^{-n}(\rho \vee \eta) = f^{-n}\rho \vee f^{-n}\eta$$

$$\rho \leq \eta \Rightarrow f^{-n}\rho \leq f^{-n}\eta$$

- Recall $H(\rho) = -\sum_{i=1}^{k} \mu(A_i) \log \mu(A_i)$

Since f is an mpt, $H(f^{-n}\rho) = -\sum_{i=1}^{k} \mu(f^{-n}A_i) \log \mu(f^{-n}(A_i))$

$= -\sum_{i=1}^{k} \mu(A_i) \log \mu(A_i) = H(\rho)$ which reflects the fact

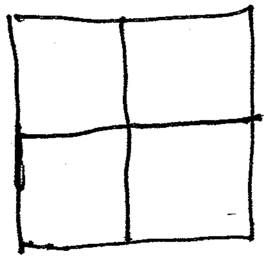
that f preserves probabilities.

- The most important partition for the entropy is

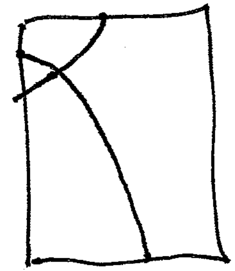
$$\rho \vee f^{-1}\rho \vee f^{-2}\rho \vee \dots \vee f^{-(n-1)}\rho$$

$$= \bigvee_{i=0}^{n-1} f^{-i}\rho$$

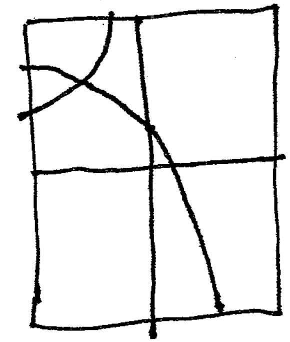
with a similar definition for finite algebras.



ρ



$f^{-1}\rho$



$\rho \vee f^{-1}\rho$

what does $x \in A_1 \cap f^{-1}A_2$ mean?

$x \in A_1$ and $f(x) \in A_2$

so, if the A_i are events, x starts in state A_1
then after time 1 is in state A_2.

So in the coding of last time, x has itinerary
that starts . 1 2

Some folks write

$$\rho^n = \bigvee_{i=0}^{n-1} f^{-i}\rho$$ as a short hand with f implicit

Thus the partition ρ^n tells us what happens for n steps of the experiment.

<u>Example</u>: In Σ_2^+, let $\rho = \{{}_0[0], {}_0[1]\}$ and $f = \sigma$. Then $\sigma^{-1}(\rho) = \{{}_1[0], {}_1[1]\}$. Thus

$$\rho^2 = \rho \vee \sigma^{-1}(\rho) = \{{}_0[0] \wedge {}_1[0], {}_0[0] \wedge {}_1[1], {}_0[1] \wedge {}_1[0], {}_0[1] \wedge {}_1[1]\} = \{{}_0[00], {}_0[01], {}_0[10], {}_0[11]\}$$

cylinder sets. This encodes the partition into length 2 two tosses of a coin in succession.

<u>DEF</u> Given a partition ρ and a mpt $f:(X,\mathcal{B},\rho)\circlearrowleft$ the entropy of the partition under f is

$$h(f,\rho) = \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} f^{-i}\rho\right)$$

where we have to prove the limit exists.
with a similar definition for finite algebras

<u>Example</u> Let $\rho = \{{}_0[0], {}_0[1]\}$ in Σ_2^+ with the Bernoulli measure $(1/2, 1/2)$

$$H(\rho) = -\left[\tfrac{1}{2}\log 1/2 + 1/2 \log 1/2\right] = -\log\left(\left(\tfrac{1}{\sqrt{2}}\right)^2\right) = \log 2$$

$$H(\rho^2) = -\sum_{i=1}^{4} 1/4 \log 1/4 = -\log\left(\left(\tfrac{1}{4}\right)^{1/4}\right)^4 = \log 4 = 2\log 2$$

$$H(\rho^n) = \log(2^n) = n \log 2$$

$$\text{so } \frac{1}{n} H(\rho^n) = \log 2 = h(T,\rho).$$

- NOTE: $h(f,\rho)$ is the average information gained per time step

- Finally, the <u>entropy</u> of f (under same hypotheses) is $h(f) = \sup\{h(f,\rho) : \rho \text{ is a finite partition of } X\}$

- $h(f) \geq 0$ but $h(f) = \infty$ is possible
- This is the step that makes h hard to compute.
- Much work has gone into finding <u>generators</u> ρ which have the property that $h(f,\rho) = h(f)$
- Often the entropy is written $h_\mu(f)$ to make clear what the measure is
- This entropy is sometimes called the metric or measure theoretic entropy to distinguish it from the topological entropy.
- In the example, in fact $h(T) = \log 2$

To prove the existence of the limit

$$\lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} f^{-i}\rho\right)$$

we need two lemmas. The first is standard analysis

<u>DEF</u>: A sequence $\{a_n\}$ is called subadditive $(n=1,2,\dots)$ if $a_{n+p} \le a_n + a_p$ $\forall\, n,p \ge 1$

<u>Lemma 1</u> If $\{a_n\}$ is subadditive then

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf\left\{\frac{a_n}{n}\right\}$$

<u>Remark</u>: The limit could be $-\infty$ but if $a_n \ge 0$ then the limit is also.

Proof of Lemma 1:

$\boxed{a_{kp} \le k a_p}$

Fix a $p > 0$. Any n can be written as $n = kp + i$ with $0 \le i < p$. Then

$$\frac{a_n}{n} = \frac{a_{kp+i}}{i+kp} \le \frac{a_{kp} + a_i}{kp} \le \frac{k a_p + a_i}{kp} = \frac{a_p}{p} + \frac{a_i}{kp}$$

Now $n \to \infty$ as $k \to \infty$ and so $\limsup \frac{a_n}{n} \le \frac{a_p}{p}$, $\forall p > 0$.

inf

But $\inf \frac{a_p}{p} \le \liminf \frac{a_n}{n}$ from the definitions

and so $\limsup \frac{a_n}{n} \le \inf \frac{a_p}{p} \le \liminf \frac{a_n}{n}$. Since always $\liminf \le \limsup$, we have $\limsup \frac{a_n}{n} = \liminf \frac{a_n}{n}$

$= \lim \frac{a_n}{n} = \inf \frac{a_p}{p}$ ∎

Lemma 2: If ρ and η are finite partitions then

$$H(\rho \vee \eta) \le H(\rho) + H(\eta).$$

Prop: $\lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} f^{-i}\rho\right)$ exists and is ≥ 0.

Proof: let $a_n = H\left(\bigvee_{i=0}^{n-1} f^{-i}\rho\right)$, we show $\{a_n\}$ is subadditive

For $p>0$, $a_{n+p} = H\left(\bigvee_{i=0}^{n+p-1} f^{-i}\rho\right) = H\left(\bigvee_{i=0}^{n-1} f^{-i}\rho \vee \bigvee_{i=n}^{n+p-1} f^{-i}\rho\right)$

$$\le H\left(\bigvee_{i=0}^{n-1} f^{-i}\rho\right) + H\left(\bigvee_{i=n}^{n+p-1} f^{-i}\rho\right)$$

$$= a_n + H\left(\bigvee_{i=0}^{p-1} f^{-i}\rho\right) = a_n + a_p$$

using $f^{-n}\bigvee_{i=0}^{p-1} f^{-i}\rho = \bigvee_{i=n}^{n+p} f^{-i}\rho$ and $H(f^{-n}\rho) = H(\rho)$ $\blacksquare$