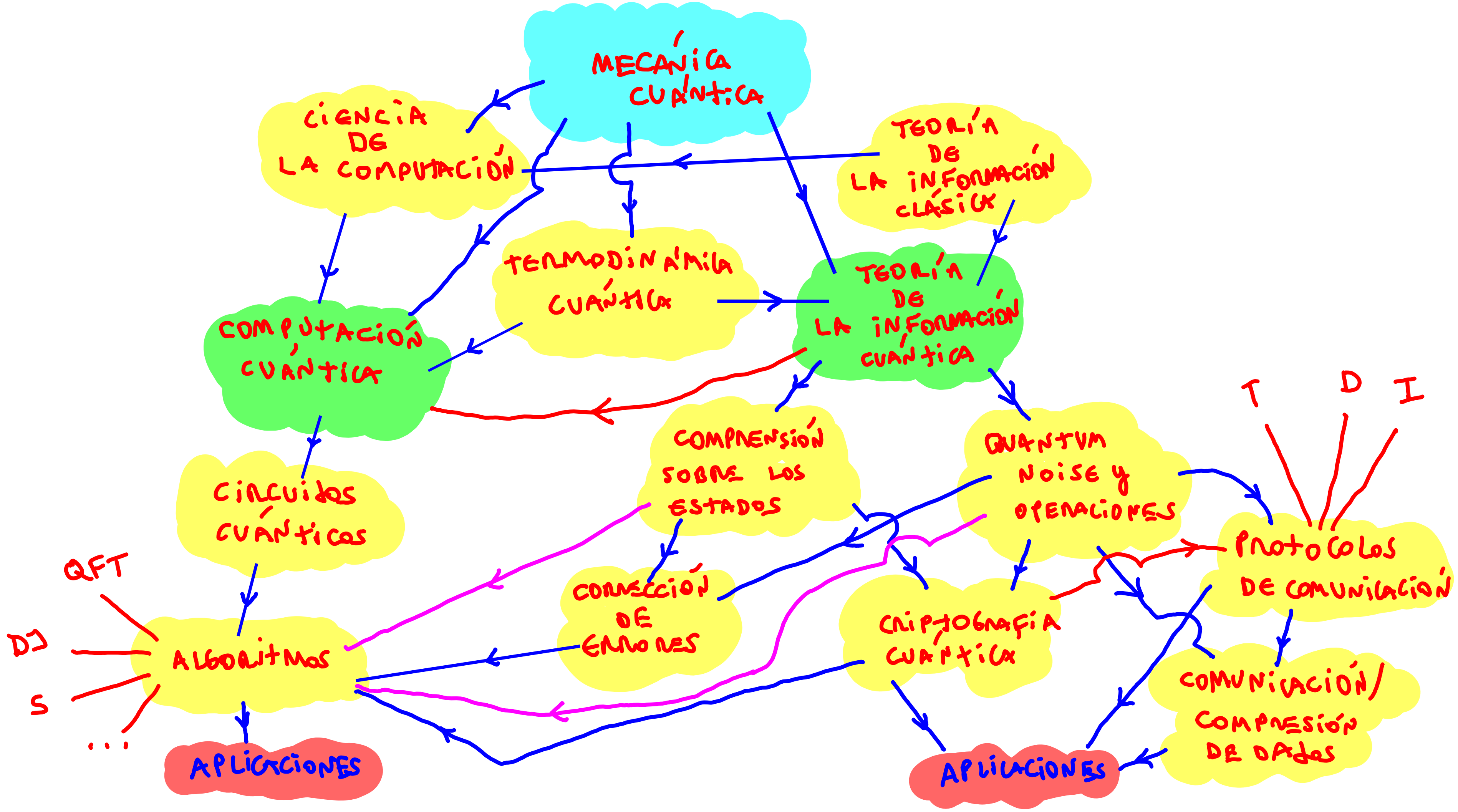
INFORMACIÓN CUÁNTICA

CLASE 2

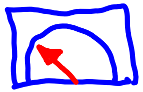
PROF. DR. ANDRÉ TANUS CESÁRIO (ANDRETANUS@GMAIL.COM)
PROF. DR. TEÓFILO VARGAS



MECÁNICA CUÁNTICA
CIENCIA DE LA COMPUTACIÓN
TEORÍA DE LA INFORMACIÓN CLÁSICA
TERMODINÁMICA CUÁNTICA
COMPUTACIÓN CUÁNTICA
TEORÍA DE LA INFORMACIÓN CUÁNTICA
CIRCUITOS CUÁNTICOS
COMPRENSIÓN SOBRE LOS ESTADOS
QUANTUM NOISE Y OPERACIONES
T
D
I
QFT
DJ
S
...
ALGORITMOS
CORRECCIÓN DE ERRORES
CRIPTOGRAFÍA CUÁNTICA
PROTOCOLOS DE COMUNICACIÓN
COMUNICACIÓN/ COMPRESIÓN DE DATOS
APLICACIONES
APLICACIONES

2. TEORÍA CLÁSICA DE LA INFORMACIÓN

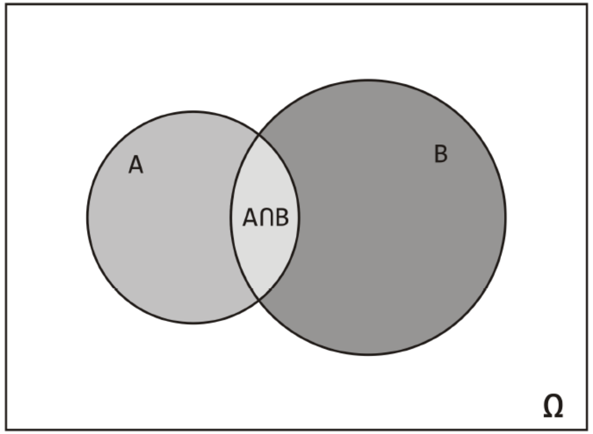
REGRESANDO

	INFORMACIÓN	CODIFICACIÓN	COMUNICACIÓN	DECODIFICACIÓN/MEDICIÓN
CLÁSICO		m bits 100110101... ↓ $f(m)$	 GRABACIÓN $f'(m)$ → DECOHERENCIA → ↓ FIDELIDAD →  TRANSMISIÓN $f''(m)$ CANAL CLÁSICO →)) 	 $g(f'(m))$ ↓ ? $m'' \approx m$
CUÁNTICO		100110101... CODIFICADO EN UN SISTEMA FÍSICO ρ_m $\mathcal{E}(\rho_m)$	- GRABACIÓN: LÍMITE DE LANDAUER $\mathcal{E}'(\rho_m)$ - TRANSMISIÓN: CAPACIDAD DE UN CANAL CUÁNTICO $I(\rho_m)$ ⇒ DECOHERENCIA Y RUIDO	MEDICIÓN  $\rho_m' \approx \rho_m$

2. TEORÍA CLÁSICA DE LA INFORMACIÓN

2.1 REPASO DE LA TEORÍA DE LA PROBABILIDAD

TEOREMA DE BAYES

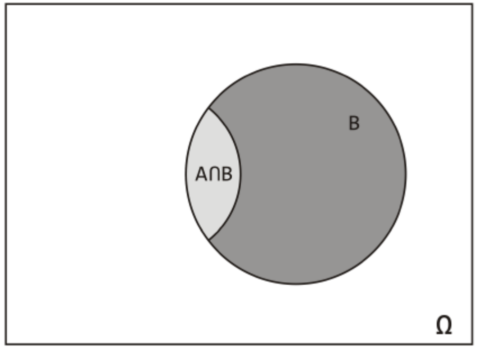


- EL UNIVERSO Ω : TODOS LOS RESULTADOS POSIBLES.
- $P(\Omega) = \dfrac{|\Omega|}{|\Omega|} = 1$, ADÓNDE $|\cdot|$ SIGNIFICA ÁREA (UNA MEDIDA)
- $A \cap B$ ES LA PARTE ADÓNDE A Y B OCURREN
- $P(B) = \dfrac{|B|}{|\Omega|}$ Y $P(A) = \dfrac{|A|}{|\Omega|}$

- $P(A \cap B) = \dfrac{|A \cap B|}{|\Omega|}$

- PERO CUAL LA PROBABILIDAD DE A OCURRIR, YA QUE B YA HA OCURRIDO

$$P(A|B) = \frac{|A \cap B|}{|B|} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A|B) \cdot P(B)$$



LA MISMA IDEA

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \Rightarrow P(A \cap B) = P(B|A) \cdot P(A)$$

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

CUANTIFICACIÓN DE LA INFORMACIÓN

- ¿CÓMO SE CUANTIFICA LA INFORMACIÓN?
- SHANNON (1948) DESARROLLÓ LA 1ª TEORÍA EXITOSA DE LA INFORMACIÓN (P/ BELL LAB.)
- SU TEORÍA SE USA P/ BIOLOGÍA, SOCIOLOGÍA, ECONOMÍA, ETC.

• SHANNON MODELÓ INFORMACIÓN COMO EVENTOS QUE OCURREN CON CIERTAS PROBABILIDADES.

• CON ESTO, ÉL POSTULÓ LOS SIGUIENTES REFINAMIENTOS QUE CUALQUIER MEDIDA DE INFORMACIÓN DEBE POSEER:

1. LA CANTIDAD DE INFORMACIÓN DE UN EVENTO X DEBE DEPENDER SOLO DE SU PROBABILIDAD P(X). CUANTO MAS SORPRENDIDOS ESTAMOS POR UN EVENTO QUE OCURRE, MAS INFORMACIÓN LLEVA ESTE EVENTO.

EJEMPLO: CAÍDA DE NIEVE EN EL ÁRTICO VERSUS EN LIMA.

2. LA FUNCIÓN $I(p)$ DEBE SER CONTÍNUA EN p (IGNORANDO TRANSICIONES DE FASE Y ALGUNAS ANOMALÍAS). $I(p+dp) \approx dI(p)$. TIENE POCO SENTIDO QUE SI LA PROBABILIDAD DE UN EVENTO NOS DEJE MUCHO MÁS SORPRENDIDOS CUANDO EL EVENTO OCURRA.

3. ADITIVIDAD: $I(P(X,Y)) = I(P(X)) + I(P(Y))$ SI LOS EVENTOS X Y Y SON INDEPENDIENTES, ES DECIR, SI $P(X,Y) = P(X)P(Y)$

SHANNON MOSTRÓ QUE LAS 3 CONDICIONES IMPLICAN QUE

$$I(P(X)) = \log_2\left(\frac{1}{P(X)}\right) = -\log_2(P(X))$$

- LA ENTROPÍA DE SHANNON APARECE COMO LA INFORMACIÓN PROMEDIA REQUERIDA PARA ESPECIFICAR X, CUANDO EL RECEPTOR RECIBE P(X).

- MIDE DE MANERA EQUIVALENTE, LA CANTIDAD DE INCERTIDUMBRE REPRESENTADA POR P(X).

- EN NUESTRO CONTEXTO, EQUIVALE A LA CANTIDAD MÍNIMA EN BITS QUE DEBEN TRANSMITIRSE PARA DETERMINARNOS LA VARIABLE X, PERO ESTE CONCEPTO DEBERÍA SER CONECTADO CON LOS TEOREMAS DE COMPRESIÓN Y CODIFICACIÓN DE DATOS.

- VALOR ESPERADO O VALOR PROMEDIO DE UNA FUNCIÓN g DE UNA VARIABLE ALEATORIA X QUE OCURRE CON PROBABILIDAD P(X)

$$E_{P(X)}(g(X)) = \sum_{x} P(x)\, g(x)$$

ENTROPÍA DE SHANNON

$$H(x) = E_{p(x)}(I(p(x))) = \sum_x p(x) \log_2\left(\frac{1}{p(x)}\right) = -\sum_x p(x) \cdot \log(p(x))$$

$$\boxed{H(x) = -\sum_{x=1}^{n} p(x) \cdot \log_2(p(x))}$$

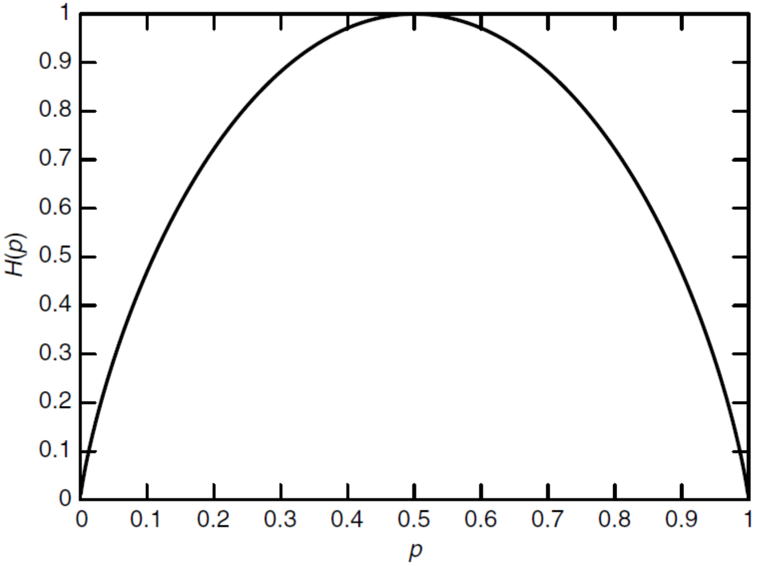
- CUANTIFICA CUANTA INFORMACIÓN ES NECESARIA, EN PROMEDIO, PARA CODIFICAR n BITS DE INFORMACIÓN.

- VAMOS SUPONER QUE TENEMOS UN ALFABETO BINARIO $X = \{0, 1\}$ CON $P(0) = (1-p)$ y $P(1) = p$

- LA ENTROPÍA DE SHANNON DE $P(x)$ ES :

$$H(X) = -\sum_{x=0}^{1} P(x) \cdot \log_2(P(x)) = -P(0)\log_2(P(0)) - P(1)\log_2(P(1))$$

$$H(X) = -(1-p) \cdot \log_2(1-p) - p \cdot \log_2(p)$$



- SUPONGAMOS QUE ALICE JUEGA UNA MONEDA SESGADA n VECES y QUIERE ENVIAR A BOB EL RESULTADO DEL JUEGO.

- DE EXPERIMENTOS PREVIOS SABEN QUE

p() $= 0{,}75 \longrightarrow P(0) = 0{,}75$

y

p() $= 0{,}25 \longrightarrow P(1) = 0{,}25$

- SI ALICE JUEGA $n = 1$ MILLÓN DE VECES, $n_0 \approx 750000$ y $n_1 \approx 250000$

- A MEDIDA QUE $n \to \infty$, ELLOS PUEDEN ESTAR AUN MAS SEGUROS QUE VAN A RECIBIR UNA SECUENCIA CON NUMERO $n_0 = 0{,}75 \cdot n$ y $n_1 = 0{,}25\, n$.

- ESTA SECUENCIA SE LLAMA SECUENCIA TÍPICA

- AHORA VAMOS SUPONER QUE VAMOS ENVÍAR UNA CANTIDAD n GRANDE.

- LAS SECUENCIAS TÍPICAS VAN A TENER, APROXIMADAMENTE

EN EL EJEMPLO $p = 0{,}25$

$$\begin{cases} n_0 \approx n \cdot p(0) = n \cdot (1-p) \\ n_1 \approx n \cdot p(1) = n \cdot p \end{cases}$$

- EL NÚMERO DE SECUENCIAS DISTINTAS ES DEL ORDEN DEL COEFICIENTE BINOMIAL DE NEWTON:

$$\text{NÚMERO} \quad N \approx \binom{n}{n \cdot p},$$

- USANDO LA APROXIMACIÓN DE STIRLING, $\ln(n!) \approx n \cdot \ln n - n$

$$\log\binom{n}{n\cdot p} = \log\left(\frac{n!}{np!\,(n-np)!}\right) = \log\left(\frac{n!}{np!\,n(1-p)!}\right) \approx$$

$$\log\binom{n}{n\cdot p} \approx n\cdot\log(n) - n - \left[np\cdot\log(np) - np + n(1-p)\cdot\log(n\cdot(1-p)) - n(1-p)\right]$$

$$\log_2(N) \approx n\left[-p\cdot\log_2(p) - (1-p)\cdot\log_2(1-p)\right] \approx n\cdot H(X)$$

$$N \approx 2^{nH(X)}$$

SECUENCIAS DE TAMAÑO $\approx nH(X)$

- CUANDO $n \to \infty$ P(SECUENCIA TÍPICA) $\to 1$
- CON UNA PEQUEÑA PROBABILIDAD, LA MONEDA PRODUCIRÁ UNA SECUENCIA ATÍPICA, EN CUYO CASO LA COMPRESIÓN FALLA.
- POR ESTA RAZÓN, ESTA COMPRESIÓN ES LLAMADA DE COMPRESIÓN SIN RUÍDO.
- EN GENERAL, EL OBJETIVO DE LA COMPRESIÓN DE DATOS ES PARA QUE 2 PARTES (ALICE Y BOB), ES DECIR, A Y B SE COMUNIQUEN EN UN CANAL INICIALMENTE SIN RUÍDO.
- LA COMPRESIÓN DE DATOS FUNCIONA COMO UN GRAN BLOQUE DE INFORMACIÓN QUE SE PUEDE DIVIDIR EN SECUENCIAS TÍPICAS (QUE CONTIENEN CARACTERES PROBABLES) Y SECUENCIAS ATÍPICAS QUE CONTIENEN SECUENCIAS DE CARACTERES IMPROBABLES.

- SEA X^n n MUESTRAS INDEPENDIENTES DE UNA VARIABLE ALEATÓRIA X Y SEA $x_1; x_2; x_3; \ldots; x_n$ UNA SECUENCIA DE VALORES OBTENIDOS DE X^n.
- $H(X) = E_{p(x)}(-\log(p(x)))$
- LA LEY DOS GRANDES NÚMEROS (VERSIÓN DÉBIL) NOS DICE QUE PARA CADA $\delta > 0$ Y $\varepsilon > 0$, EXISTE UN n (SUFICIENTEMENTE GRANDE), TAL QUE, CON UNA PROBABILIDAD DE AL MENOS $1-\varepsilon$, UNA SECUENCIA ε-TÍPICA $x_1; \ldots; x_n$ OCURRA SATISFACENDO

$$2^{-n(H(x)+\delta)} < p(x_1, x_2, \ldots, x_n) < 2^{-n(H(x)-\delta)}$$

DE $2^{-n(H(X)+\delta)} < p(x_1, x_2, \dots, x_n)$

$$|N(\varepsilon,\delta)| \leq 2^{n(H(X)+\delta)}$$

$n(H(X)+\delta)$ BITS SON SUFICIENTES PARA UN ERROR MÁXIMO DE ε.

PROPRIEDADES DE LA ENTROPÍA DE SHANNON

- LA ENTROPÍA DE SHANNON ES NO NEGATIVA $H(X) \geq 0$

COMO $0 \leq P(X) \leq 1$, ENTONCES $H(X) = E_{p(x)}\left(\log\left(\frac{1}{P(X)}\right)\right) \geq 0$

DESIGUALDAD DE JENSEN P/ FUNCIONES CÓNCAVAS

SEA X UNA VARIABLE ALEATORIA Y f UNA FUNCIÓN ESTRICTAMENTE CÓNCAVA, ENTONCES

$$E_{p(x)}(f(X)) \leq f(E_{p(x)}(X))$$

- PRINCIPIO DE MAXENT

SEA $f\left(\frac{1}{p(x)}\right) = \log\left(\frac{1}{p(x)}\right)$

$f\left(\frac{1}{p(x)}\right) = I(p(x))$

$$E_{p(x)}(f(X)) \leq f(E_{p(x)}(X))$$

$$\sum_{x=1}^{n} p(x)\cdot \log\left(\frac{1}{p(x)}\right) \leq \log\left(\sum_{x=1}^{n} \cancel{p(x)}\left(\frac{1}{\cancel{p(x)}}\right)\right)$$

$$H(p(x)) \leq \log n$$

EL VALOR MÁXIMO OCURRE P/ EL VECTOR $p(x) = \frac{1}{n} \ \forall \ n$.

INFORMACIÓN MÚTUA DE SHANNON Y OTRAS ENTROPÍAS

- ENTROPIA RELATIVA (DIVERGENCIA DE KULLBACK-LEIBLER) ES UNA MEDIDA DE COMO UNA DISTRIBUCIÓN P ES DIFERENTE DE UNA OTRA DISTRIBUCIÓN DE REFERÊNCIA Q

$$D_{KL}(P\|Q) = \sum_{x} P(x) \log\left(\frac{P(x)}{Q(x)}\right)$$

MIDE LA SIMILITUD ENTRE P Y Q.

- UNA INTERPRETACIÓN DE D_{KL} ES UNA "DISTANCIA" ENTRE 2 DISTRIBUCIONES, AUNQUE NO SEA UNA MÉTRICA ($D_{KL}(P\|Q) \neq D_{KL}(Q\|P)$ EN GENERAL). TAMBIÉN NO SATISFACE LA DESIGUALDAD TRIANGULAR AUNQUE EXISTA UNA DESIGUALDAD "DEL TIPO DE PITÁGORAS" GENERALIZADA.

- ENTROPÍA CONJUNTA $H(X,Y)$, DONDE $X,Y = X \cup Y$

QUE MIDE LA INFORMACIÓN COMBINADA EN 2 (O MÁS) VARIABLES ALEATORIAS

$$H(X,Y) = -\sum_{x}\sum_{y} p(x,y)\log_2(p(x,y))$$

$$H(X,Y) \geq 0$$

$$H(X,Y) \geq \text{MÁX}\{H(X), H(Y)\}$$

COVER Y THOMAS, 1995

RECORDE QUE LAS DISTRIBUCIONES MARGINALES SON:

$$P(X) = \sum_{y} P(X,Y)$$

$$P(Y) = \sum_{x} P(X,Y)$$

$$\sum_{x} P(X) = \sum_{x}\sum_{y} P(X,Y) = 1$$

$$\sum_{y} P(Y) = \sum_{y}\sum_{x} P(X,Y) = 1$$

ENTROPÍA CONDICIONAL: $H(X|Y)$ y $H(Y|X)$

MIDE LA INFORMACIÓN CONTENIDA EN UNA VARIABLE ALEATORIA DADO QUE SE CONOCE EL RESULTADO DE OTRA VARIABLE ALEATORIA

$$H(Y|X) = -\sum_{x,y} P(x,y)\log\left(P(y|x)\right) = -\sum_{x,y} P(x,y)\log\left(\frac{P(x,y)}{P(x)}\right)$$

$$H(Y|X) = -\sum_{x,y} P(x,y)\log\left(\frac{P(x,y)}{P(x)}\right)$$

$$H(X|Y) = -\sum_{x,y} P(x,y)\log\left(P(x|y)\right) = -\sum_{x,y} P(x,y)\log\left(\frac{P(x,y)}{P(y)}\right)$$

$$H(X|Y) = -\sum_{x,y} P(x,y)\log\left(\frac{P(x,y)}{P(y)}\right)$$

INFORMACIÓN MUTUA $I(X:Y)$

MIDE LA "CORRELACIÓN" ENTRE 2 VARIABLES ALEATORIAS, EN TÉRMINOS DE CUANTO CONOCIMIENTO DE UNA VARIABLE ALEATORIA REDUCE LA ÍNFORMACIÓN OBTENIDA AL APRENDER LA OTRA VARIABLE.

MIDE LA DEPENDENCIA ENTRE 2 VARIABLES

$$I(X:Y) = \sum_x \sum_y P(x,y) \log\left(\frac{P(x,y)}{P(x)\cdot P(y)}\right)$$

- SI $P(x,y) = P(x)\cdot P(y)$ (VARIABLES INDEPENDIENTES), $I(X:Y) = 0$
- $I(X:Y) \geq 0$ y $I(X:Y) = I(Y:X)$

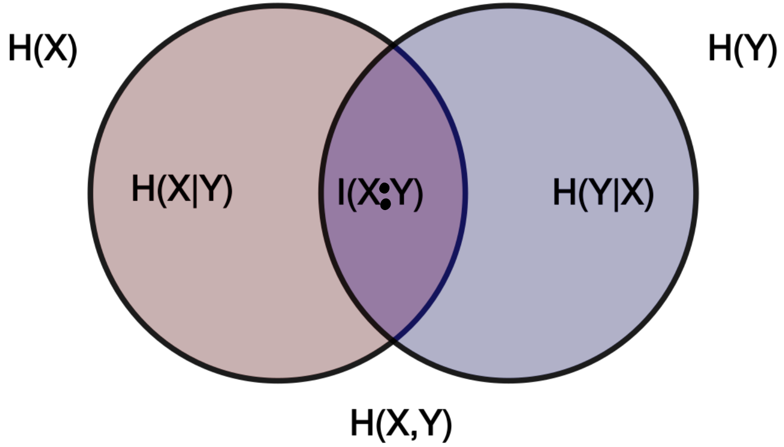
PROPIEDADES Y RELACIONES DE LAS ENTROPÍAS

CONDICIONAR DISMINUYE LA ENTROPÍA

$$H(X) \geq H(X|Y) \geq 0$$

$$H(Y) \geq H(Y|X) \geq 0$$

LA IGUALDAD OCURRE CUANDO X Y Y SON INDEPENDENTES



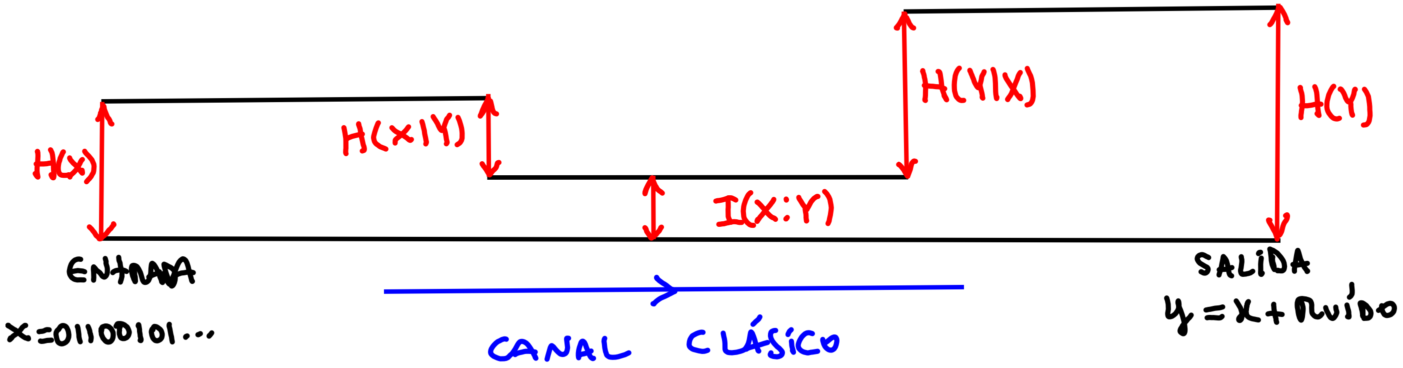
$$H(X,Y) \leq H(X) + H(Y)$$

$$I(X:Y) = H(X) - H(X|Y)$$

$$I(X:Y) = H(Y) - H(Y|X)$$

$$I(X:Y) = H(X) + H(Y) - H(X,Y)$$

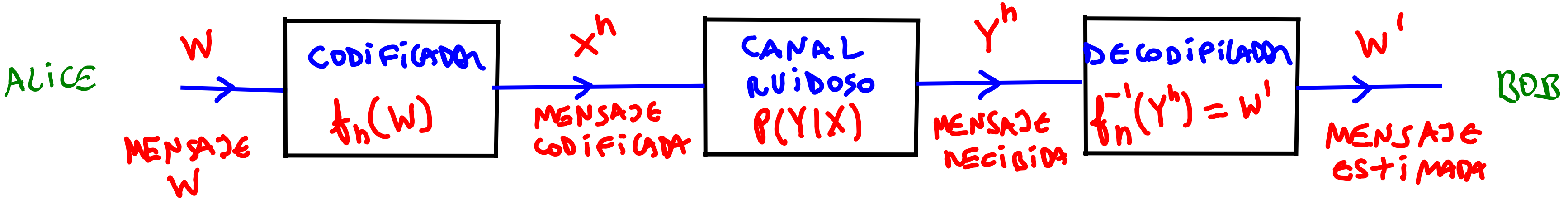
$$I(X:Y) = H(X,Y) - H(X|Y) - H(Y|X)$$



CAPACIDAD DE UN CANAL RUIDOSO

- CUANDO ALICE TRANSMITE UN MENSAJE A BOB, ELLA ENVÍA UNA CADENA DE BITS QUE PUEDE SER CONSIDERADA COMO UNA VARIABLE ALEATORIA X.
- SI EL CANAL ES RUIDOSO, ESTE SEÑAL SE CAMBIARÁ DEBIDO A ERRORES

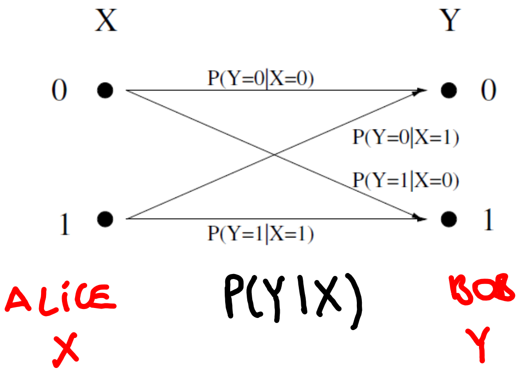
- EL MODELO MATEMÁTICO DE UN PROTOCOLO DE COMUNICACIÓN ES



- $P(Y|X)$ ES COMO MODELAMOS EL CANAL RUIDOSO

- SI LA CORRELACIÓN DE Y CON X ES CERO, BOB NO RECIBE EL MENSAJE DE ALICE.

ALICE INTENTA ENVIAR X P/ BOB



SIN EMBARGO, BOB RECIBE Y POR EL RUIDO DEL CANAL

TEOREMA DE SHANNON

- EL TEOREMA DE SHANNON MUESTRA CÓMO CALCULAR LA CAPACIDAD DE UN CANAL CLÁSICO A PARTIR DE LA DESCRIPCIÓN ESTADÍSTICA DE UN CANAL RUIDOSO CON UNA CAPACIDAD C, CON INFORMACIÓN TRANSMITIDA A UNA TASA R.

- SI $R < C$, EXISTE UNA TÉCNICA DE CODIFICACIÓN QUE PERMITE QUE LA PROBABILIDAD DE ERROR SE HAGA ARBITRARIAMENTE PEQUEÑA.

- SI $R > C$, LA PROBABILIDAD DE ERROR AUMENTA SIN LÍMITE CUANDO R AUMENTA.

- UNA FUENTE X CON ENTROPÍA $H(X)$ GENERARÁ $\approx 2^{nH(X)}$ SECUENCIAS TÍPICAS EN n PASOS. EL NÚMERO DE MENSAJES ÚTILES (SIN REDUNDÁNCIA) ES

$$2^{n(H(X)-H(X|Y))}$$

COMO $\begin{cases} I(X:Y) = H(X) - H(X|Y) \\ I(X:Y) = H(Y) - H(Y|X) \end{cases}$ $\in$ $N \approx 2^{nI(X:Y)}$

CAPACIDAD DE UN CANAL

$$C = \sup_{\substack{P(X) \\ X}} \{ I(X:Y) \}$$

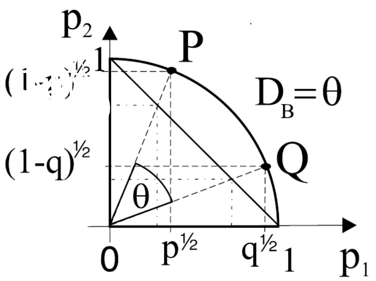
ALGUNAS MEDIDAS IMPORTANTES

FIDELIDAD

$$F(\vec{P}, \vec{Q}) = \sum_i \sqrt{p_i q_i}$$

c/ $\vec{P} = (p_1, p_2, \dots, p_N)^T$

$\vec{Q} = (q_1, q_2, \dots, q_N)^T$



$$
\begin{cases}
F(\vec{P},\vec{Q}) = \sqrt{\vec{P}}\cdot\sqrt{\vec{Q}} = \left|\sqrt{\vec{P}}\right|\cdot\left|\sqrt{\vec{Q}}\right|\cos\theta = \cos\theta \\
\text{PORQUE} \quad \sum_x \left(\sqrt{P_x}\right)^2 = \sum_x \left(\sqrt{Q_x}\right)^2 = 1
\end{cases}
$$

LA INTERPRETACIÓN DE LA FIDELIDAD ES EL PRODUCTO ESCALAR ENTRE LOS VECTORES $\sqrt{\vec{P}}$ y $\sqrt{\vec{Q}}$ EN UNA ESFERA UNITARIA

DISTANCIA ℓ_1

$$d(\vec{P},\vec{Q}) = \frac{1}{2}\sum_i |p_i - q_i|$$

ES EL EQUIVALENTE A LA DISTANCIA DE LA TRAZA

2.2) $\vec{P} = (p, 1-p)$ $\quad d(\vec{P},\vec{Q}) = \frac{1}{2}\left\{ |p-q| + |(\cancel{1}-p)-(\cancel{1}-q)| \right\}$

$\vec{Q} = (q, 1-q)$

$$d(\vec{P},\vec{Q}) = \frac{1}{2}\left\{ |p-q| + |-p+q| \right\} = \frac{1}{2}\left\{ 2|p-q| \right\}$$

$$d(\vec{P},\vec{Q}) = |p-q|$$

DISTANCIA ℓ_2 (NORMA EUCLIDEANA) TEOREMA DE PITÁGORAS

$$d_2(P,Q) = \|\vec{P}-\vec{Q}\|_2 = \sqrt{(P_1-Q_1)^2 + \cdots + (P_n-Q_n)^2}$$

DISTANCIA ℓ_∞

$$d_\infty(P,Q) = \lim_{k\to\infty}\left[\sum_i |p_i - q_i|^k\right]^{\frac{1}{k}} = \max_i(|p_i - q_i|)$$

DISTANCIA DE BHATTACHARYYA

(UNA DEFINICIÓN POSIBLE):

$$d_{\text{BHATTACHARYYA}} = -\ln\left[F(P,Q)\right] = -\ln\left[\sum_i \sqrt{p_i q_i}\right]$$

DIVERGENCIA DE JANSEN-SHANNON J

$$J(P\|Q) = \frac{1}{2} D_{KL}(P\|M) + \frac{1}{2} D_{KL}(Q\|M), \quad \text{CON} \quad M = \frac{P+Q}{2}$$

DISTANCIA DE WOOTTERS

$$D_W = \arccos(F)$$

DISTANCIA DE HELLINGER

$$H(P,Q) = \frac{1}{\sqrt{2}} \left\| \sqrt{P} - \sqrt{Q} \right\|_2 = \frac{1}{\sqrt{2}} \sqrt{\sum_i \left(\sqrt{p_i} - \sqrt{q_i}\right)^2}$$

ES INTERESANTE MIRAR QUE

$$H^2(P,Q) = \frac{1}{2}\sum_i \left(\sqrt{p_i} - \sqrt{q_i}\right)^2$$

$$H^2(P,Q) = \frac{1}{2}\sum_i p_i - 2\sqrt{p_i}\sqrt{q_i} + q_i$$

$$H^2(P,Q) = \cancelto{1}{\frac{1}{2}\sum_i (p_i + q_i)} - \frac{1}{\cancel{2}}\sum_i \cancel{2}\sqrt{p_i}\sqrt{q_i}$$

$$F(P,Q) = \sum_i \sqrt{p_i q_i} = 1 - H^2(P,Q)$$

$$\frac{1}{2}\sum_i (p_i + q_i) = 1$$

$$\frac{1}{2}\sum_i p_i + \frac{1}{2}\sum_i q_i = 1$$

PROBABILIDADES "NEGATIVAS" - INTERFERENCIA

Quantum mechanics permits the cancellation of possibilities.

Nick Herbert, Quantum Reality.

- COMO HEMOS VISTO, HAY VARIOS TIPOS DE NORMAS QUE INDUCEN VARIAS DISTANCIAS DIFERENTES.

- HAY LA NORMA ℓ_1: $\|\vec{P}\|_1 = \sum_i |p_i| = \sum_i p_i = 1$ $\longrightarrow$ $d(\vec{P},\vec{Q}) = \frac{1}{2}\sum_i |p_i - q_i|$

  NORMA DE LA TEORÍA DE PROBABILIDADES

  LA DISTANCIA ℓ_1 O LA NORMA DE LA TRAZA

- OTRA NORMA POSIBLE ES LA NORMA EUCLIDIANA

  ℓ_2: $\|\vec{P}\|_2 = \sqrt{\sum_i p_i^2}$ $\longrightarrow$ $d_2(P,Q) = \|\vec{P} - \vec{Q}\|_2 = \sqrt{(P_1-Q_1)^2 + \cdots + (P_n-Q_n)^2}$

  LA DISTANCIA ℓ_2 ES EL TEOREMA DE PITÁGORAS

- CUALES SON LAS CONSECUENCIAS SI INTENTAMOS CONSTRUIR UNA "OTRA" TEORÍA DE PROBABILIDADES BASADA EN LA NORMA ℓ_2?

- VOY A INTENTAR CONVENCEROS QUE ES LA MECÁNICA CUÁNTICA QUE SURGE NECESARIAMENTE.

- LA UNIDAD BÁSICA DE LA INFORMACIÓN ES EL BIT

BIT PROBABILÍSTICO $\begin{cases} P(0) = p \\ P(1) = 1-p \end{cases}$ $\qquad \vec{p} = (p, 1-p)^T; \ \|\vec{p}\|_1 = 1$

- VAMOS CAMBIAR LA NORMA ℓ_1 POR ℓ_2 Y MANTENER LOS VECTORES DE PROBABILIDAD

$$\vec{p} = (\alpha, \beta)^T \in \mathbb{R}^2$$

- CON LA NORMA ℓ_2, NO PODEMOS ASIGNAR DIRECTAMENTE UNA PROBABILIDAD PARA REALIZARNOS UN EXPERIMENTO Y OBTENER UN RESULTADO "0", POR EJEMPLO, NECESITAMOS DEFINIR ALGO COMO UNA AMPLITUD DE PROBABILIDADES

- ESTAS AMPLITUDES α y β, EN EL CASO BINÁRIO, SON TALES QUE

$$\begin{cases} p(\alpha) = \alpha^2 \\ p(\beta) = \beta^2 \end{cases} \qquad \vec{p} = (\alpha, \beta)^T \in \mathbb{R}^2$$

- SI INTENTAMOS USAR LA NORMA ℓ_1 PARA MEDIR ESTE VECTOR, VAMOS OBTENER UM VECTOR NO-NORMALIZADO, PORQUE $\alpha + \beta \neq 1$. (EN GENERAL $\alpha + \beta \in \mathbb{C}$)

- SIN EMBARGO, COMO ESTA TEORÍA ES MEDIDA POR LA NORMA ℓ_2, Y NO LA NORMA ℓ_1,

$$\|\vec{p}\|_2 = \sqrt{\sum_i p_i^2} = \sqrt{\alpha^2 + \beta^2} = \sqrt{p(\alpha) + p(\beta)} = \sqrt{1} = 1$$

- ENTONCES, POR QUÉ NO OLVIDAR LAS AMPLITUDES α y β y DESCRIBIR EL BIT CUÁNTICO SOLO POR SUS PROBABILIDADES?

- LA DIFERENCIA ES CÓMO SE TRANSFORMA ESTE VECTOR CUANDO REALIZAMOS OPERACIONES SOBRE ÉL.

- LAS MATRICES ESTOCÁSTICAS SON LAS MATRICES QUE PRESERVAN LA NORMA ℓ_1 INALTERADA Y SON ESTAS TRANSFORMACIONES QUE LOS BITS CLÁSICOS Y ESTADOS DE LA TEORÍA DE PROBABILIDAD EN GENERAL

$$\|\vec{p}\|_1 = 1$$

SEA S UNA MATRIZ ESTOCÁSTICA (SUS COLUMNAS O LÍNEAS SUMAN UNO)

PRUEBA P/ $\dim(S) = 2$

$$S = \begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix} \quad c/ \quad \begin{cases} 0 \leq p \leq 1 \\ 0 \leq q \leq 1 \\ 0 \leq s \leq 1 \end{cases}$$

$$\vec{p} = \begin{pmatrix} s \\ 1-s \end{pmatrix}$$

$$S\cdot\vec{p} = \begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix}_{2\times 2} \cdot \begin{pmatrix} s \\ 1-s \end{pmatrix}_{2\times 1} = \begin{pmatrix} ps + (1-q)(1-s) \\ q(1-s) + (1-p)\cdot s \end{pmatrix}_{2\times 1}$$

$$\|S\cdot\vec{p}\|_1 = ps + (1-q)(1-s) + q(1-s) + (1-p)\cdot s = \cancel{ps} + 1 - \cancel{s} - \cancel{q} + \cancel{qs} + \cancel{q} - \cancel{qs} + \cancel{s} - \cancel{ps} = 1$$

- PERO, EN LA NORMA ℓ_2, LA MATRIZ MÁS GENERAL QUE SIEMPRE MAPEA UN VECTOR NORMALIZADO EN ESTA NORMA ES UNA MATRIZ UNITARIA U

$$U^\dagger U = UU^\dagger = \mathbb{I}.$$

$$(|\psi'\rangle)^\dagger = (U|\psi\rangle)^\dagger$$

$$\langle\psi'| = \langle\psi|U^\dagger$$

- SI $\||\psi\rangle\|_2 = \sqrt{\langle\psi|\psi\rangle} = 1$

$|\psi'\rangle = U|\psi\rangle$ TAMBIÉN VA A ESTAR NORMALIZADO

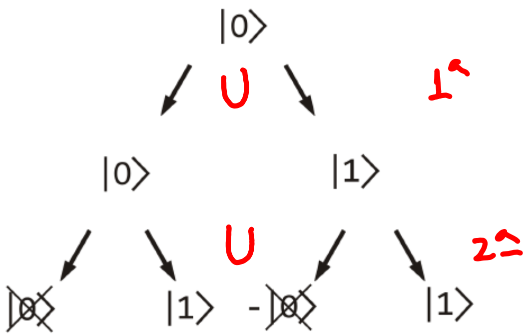
$$\||\psi'\rangle\|_2 = \sqrt{\langle\psi'|\psi'\rangle} = \sqrt{\langle\psi|U^\dagger U|\psi\rangle} = \sqrt{\langle\psi|\,\mathbb{I}\,|\psi\rangle} = 1$$

- NUESTRA "TEORÍA DE PROBABILIDADES" DEFINIDA POR MEDIO DE LA NORMA ℓ_2 EVOLUCIONA POR MEIO DE MATRICES UNITARIAS. ESTO GENERA UN TIPO DE CANCELACIÓN DE LAS AMPLITUDES DE PROBABILIDAD.

- COMO EJEMPLO DE ESTE FENÓMENO, IMAGINA UN QUBIT INICIALMENTE EN EL ESTADO $|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ Y VAMOS USAR UNA "MONEDA" CUÁNTICA

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}$$

LA EVOLUCIÓN ES DADA POR ESTE DIAGRAMA

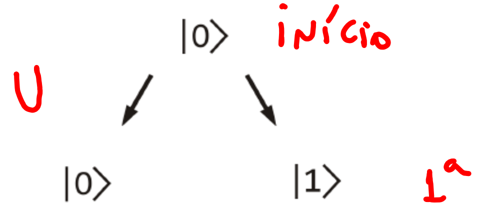


VAMOS APLICAR U DOS VECES Y INTERPRETAMOS LOS RESULTADOS EN CADA PASO.

- Como se discutió anteriormente, supongamos que nuestro qubit está inicialmente en $|\psi\rangle = |0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$. $\qquad |1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$

- Si aplicamos $U|\psi\rangle = U|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}$

$$P(|0\rangle) = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2} \quad \text{y} \quad P(|1\rangle) = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}$$



$$|\psi'\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

- Recuerda que con esta nueva norma las amplitudes no suman a 1: $\left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \neq 1\right)$.

- Por esta razón

$$\|\vec{P}\|_2 = \sqrt{\sum_i P_i^2} = \sqrt{\alpha^2 + \beta^2} = \sqrt{P(\alpha) + P(\beta)}$$
$$= \sqrt{1} = 1$$

- PERO, OBSERVE QUE SI APLICAMOS U UNA VEZ MÁS, OBTENDREMOS $|1\rangle$ CON 100% DE CERTEZA!

- PODRÍAMOS PENSAR EN LA PRIMERA OPERACIÓN U COMO UN LANZAMIENTO DE UNA "MONEDA". DESPUÉS DE LA PRIMERA OPERACIÓN CUÁNTICA, TENÍAMOS 2 CAMINOS DIFERENTES QUE SEGUIR.

- PERO, SI APLICAMOS 2 VECES UNA OPERACIÓN ALEATORIA PRODUCIREMOS UN RESULTADO DETERMINISTA $|1\rangle$!

- ESTE ES EL FENÓMENO DE INTERFERENCIA.

LOS CAMINOS NO OBSERVADOS QUE CONDUCEN AL RESULTADO $|0\rangle$ INTERFIEREN DESTRUCTIVAMENTE Y CANCELAN ENTRE SÍ Y NO SON OBSERVADOS, MIENTRAS LOS CAMINOS P/ $|1\rangle$ INTERFIEREN CONSTRUCTIVAMENTE!