

Información: cuantificar alternativas y dependencia

IDE-002 · Ideas fundamentales

Ensayos conceptuales para El universo y sus observadores

Autor: Juan Segurola

Versión: v0.6 ES · 19 de septiembre de 2026

Estado: Versión de publicación

Forma del artículo: Ensayo conceptual

Escala / contexto: Información estadística y física a través de escalas

Límite: Información cuantitativa y codificación; las señales físicas y la señalización se tratan en IDE-024

Trabajos relacionados: IDE-003, IDE-008, IDE-024

Perfil epistémico: Conocimiento establecido

Licencia en publicación: Todos los derechos reservados

Resumen del artículo

La información es un lenguaje cuantitativo para alternativas distinguibles, incertidumbre estadística y dependencia. En la teoría de Shannon, una fuente se representa mediante mensajes posibles con probabilidades; la entropía resume la incertidumbre media, mientras la información mutua mide cuánto reduce la observación de una variable la incertidumbre sobre otra. Estas magnitudes fijan límites rigurosos para compresión y comunicación fiable, pero por sí solas no aportan significado semántico, dirección causal ni una sustancia física llamada información. Un bit es una unidad definida por logaritmos en base dos, y la misma distinción lógica puede encarnarse en muchos sistemas físicos. La implementación física, sin embargo, importa: la memoria debe representarse, manipularse y a veces borrarse, y las operaciones lógicamente irreversibles tienen consecuencias termodinámicas. Este artículo separa la medida matemática de sus soportes físicos y de la señalización causal, fijando la frontera con IDE-024, que pregunta qué es una señal y cómo se transmite físicamente una influencia.

En 60 segundos

Empieza con un conjunto de resultados posibles y sus probabilidades. La entropía de Shannon, medida en bits cuando se usan logaritmos en base dos, da la incertidumbre media de esa distribución; la información mutua mide dependencia estadística entre dos variables. No son metáforas vagas. Sustentan teoremas de codificación: los resultados frecuentes pueden recibir descripciones más cortas, los canales ruidosos tienen capacidades finitas y procesar datos no puede crear información sobre una variable anterior que no estuviera en la entrada.[1,2] Pero la matemática abstrae deliberadamente el significado. Dos cadenas de bits pueden tener las mismas estadísticas de Shannon y diferir por completo en lo que una persona interpreta que significan.[1] La información tampoco es un fluido sin peso independiente de su representación. Un bit lógico puede almacenarse mediante voltaje, magnetización u otro grado de libertad físico, mientras borrar estados lógicamente distinguibles bajo condiciones

térmicas conlleva un coste termodinámico mínimo en el límite ideal.[3,4] Por último, información compartida no demuestra comunicación ni causalidad: causas comunes e interacciones previas pueden correlacionar variables sin una señal presente entre ellas.[5]

Por qué importa esta pregunta

La palabra información viaja con facilidad entre telecomunicaciones, física, biología, computación y lenguaje cotidiano, y esa portabilidad es útil y peligrosa a la vez. Sin declarar variables aleatorias, probabilidades y supuestos de codificación, una afirmación numérica sobre información puede quedar indefinida; sin un modelo causal, la información mutua puede confundirse con influencia; sin un modelo físico, el principio de Landauer puede citarse erróneamente como coste energético universal de cada paso computacional. Un tratamiento disciplinado permite usar las mismas herramientas matemáticas entre campos manteniendo separadas preguntas distintas: cuánta incertidumbre tiene una fuente, cuánto dependen dos variables, cómo se codifica un mensaje y cómo viaja realmente una señal física.

1. Qué mide la información de Shannon

El punto de partida de Shannon no es una sustancia, sino un conjunto de mensajes posibles y sus probabilidades. Para una variable aleatoria discreta X con resultados x y probabilidades $p(x)$, la entropía $H(X) = -\sum p(x) \log_2 p(x)$ es la sorpresa logarítmica esperada. La base del logaritmo fija la unidad: base dos da bits. La entropía es cero cuando un resultado es seguro y aumenta cuando la probabilidad se reparte entre alternativas.[1,2]

La definición es operacional porque conecta con la codificación. Para secuencias largas generadas por una fuente, la entropía fija la escala fundamental de la longitud media de descripción sin pérdida: las fuentes muy predecibles pueden comprimirse más que las muy inciertas. La teoría de Shannon para canales con ruido relaciona igualmente la comunicación fiable con la capacidad del canal y no con un alfabeto o hardware particulares.[1,2] La abstracción es potente precisamente porque ignora la forma material de los símbolos.

La información mutua plantea otra pregunta. $I(X;Y)$ cuantifica cuánto se aparta de la independencia la distribución conjunta de X e Y , o de forma equivalente cuánto reduce conocer una la incertidumbre sobre la otra. Es no negativa y se anula cuando las variables son estadísticamente independientes bajo la distribución especificada.[2] Un valor alto significa, por tanto, dependencia estadística fuerte, no que X haya causado Y ni que un mensaje haya viajado físicamente de X a Y .

En qué punto estamos

La teoría de la información empieza con alternativas explícitas y distribuciones de probabilidad y las convierte en magnitudes que restringen compresión y comunicación. Entropía e información mutua son propiedades de una descripción probabilística, no fluidos ocultos dentro de objetos. Esto hace inevitable la siguiente pregunta: si la matemática abstrae el hardware, ¿qué cambia cuando las mismas distinciones lógicas se implementan en dispositivos físicos reales?

2. Un bit es lógico, pero su implementación es física

Un bit es una distinción binaria, no una partícula microscópica. Los estados lógicos 0 y 1 pueden representarse por dos rangos de voltaje, dos direcciones de magnetización, dos estados ópticos o muchas otras configuraciones físicamente distinguibles. La teoría de la información puede tratar

estas realizaciones como equivalentes cuando solo importan las alternativas lógicas, mientras la ingeniería debe seguir además ruido, estabilidad, barreras energéticas, probabilidades de error y dinámica de lectura o cambio del estado.[1,2]

Landauer identificó un puente concreto entre irreversibilidad lógica y termodinámica. Cuando una operación física hace corresponder dos estados lógicos distinguibles con uno solo, como ocurre en el borrado idealizado, se descarta información sobre el estado previo. En un entorno térmico, realizar esa operación lógicamente irreversible de la forma más reversible posible exige aun así exportar entropía al entorno; para borrar un bit imparcial, el límite inferior ideal familiar es $kT \ln 2$ de calor disipado.[3]

El principio no afirma que toda operación lógica deba disipar $kT \ln 2$. Bennett mostró que, en principio, la computación general puede organizarse mediante pasos lógicamente reversibles, conservando información suficiente para recorrer hacia atrás estados intermedios y evitando el borrado hasta que sea realmente necesario.[4] Los ordenadores reales disipan mucho más por numerosas razones de ingeniería, pero la lección conceptual es precisa: la información es suficientemente abstracta para admitir muchas encarnaciones, mientras las operaciones sobre información encarnada siguen sujetas a la ley física.

En qué punto estamos

El mismo bit lógico puede habitar hardware muy distinto, por lo que la información no puede identificarse con un soporte material único. Sin embargo, la implementación no es irrelevante: almacenar, leer y borrar distinciones son procesos físicos con ruido y restricciones termodinámicas. Esto separa la información como relación cuantitativa entre alternativas de la señal que transporta físicamente un cambio entre lugares, tema reservado a IDE-024.

3. Información no es lo mismo que significado

Shannon diseñó explícitamente su teoría de la comunicación de modo que los aspectos semánticos de los mensajes no fueran necesarios para el problema de ingeniería.[1] Un código puede conservar una secuencia con error arbitrariamente bajo sin comprender si esa secuencia es un poema, un genoma, un archivo de software o una salida de apariencia aleatoria. La matemática atiende a probabilidades de símbolos, longitudes de código, estadísticas del canal y recuperabilidad bajo una tarea especificada.

Esto no afirma que el significado sea irreal. Es una frontera sobre lo que miden las magnitudes de Shannon. El significado depende de usuarios, convenciones, referencia, objetivos y modelos más amplios que $H(X)$ por sí sola no determina. Dos mensajes pueden tener longitud y entropía similares y diferir radicalmente en contenido semántico; a la inversa, la misma proposición puede codificarse en muchas cadenas con propiedades estadísticas distintas. Las medidas cuantitativas de información no se convierten, por compartir una palabra, en una teoría universal de la semántica.

El contexto entra, no obstante, en la matemática de forma disciplinada. Hay que elegir variables, un alfabeto o espacio de estados, un modelo probabilístico y a menudo un protocolo de codificación u observación antes de calcular entropía o información mutua. Elecciones distintas pueden responder preguntas legítimas distintas. Una vez fijadas, sin embargo, la magnitud resultante no es

arbitraria: el modelo probabilístico y las definiciones determinan el valor y los teoremas restringen qué pueden hacer las transformaciones.[2]

En qué punto estamos

La información de Shannon responde deliberadamente a una pregunta más estrecha que el significado del lenguaje ordinario. Cuantifica incertidumbre y dependencia una vez especificadas variables y probabilidades; no infiere a qué se refieren los símbolos ni por qué importan a un organismo u observador. Esa frontera evita un error de categoría frecuente y prepara el siguiente: puede haber dependencia estadística fuerte sin influencia causal ni comunicación presente.

4. Dependencia no es causalidad ni señalización

La información mutua detecta dependencia estadística. Si dos variables comparten una causa común, quedaron correlacionadas por una interacción anterior o son ambas efectos de un tercer proceso, observar una puede reducir incertidumbre sobre la otra aunque ninguna envíe actualmente nada a la otra. La teoría de la información por sí sola no puede decidir si la dependencia surgió por influencia directa, causa común, selección u otra estructura causal.[2,5]

La inferencia causal requiere supuestos adicionales o intervenciones. El marco gráfico de Pearl lo hace explícito al combinar distribuciones de probabilidad con información estructural sobre qué variables pueden afectar directamente a cuáles; solo bajo tales supuestos los datos observacionales pueden identificar efectos causales, y a veces no pueden hacerlo.[5] Esta es una tarea distinta de medir dependencia. La diferencia importa siempre que frases como «la información fluye de A a B» se usen como si una correlación no nula ya hubiera demostrado una ruta física.

La desigualdad de procesamiento de datos aporta una restricción útil pero limitada. Si X influye en Y y después Y se procesa para formar Z mediante una cadena de Markov $X \rightarrow Y \rightarrow Z$, el procesamiento ordinario no puede aumentar la información mutua que Z contiene sobre X más allá de la que ya contenía Y.[2] El teorema restringe la estructura estadística dado un modelo; no establece que la cadena causal asumida sea la historia física verdadera. Modelo y medida deben mantenerse separados.

En qué punto estamos

La información compartida puede surgir sin una señal contemporánea, de modo que dependencia, causalidad y comunicación no deben colapsarse en una sola palabra. La teoría de la información restringe distribuciones de probabilidad; la inferencia causal añade supuestos sobre intervención y dirección; la física de señales pregunta qué interacción puede transportar realmente un cambio. Separados esos papeles, queda precisar qué unifica la información sin convertirla en sustancia universal.

5. Qué unifica el concepto y dónde se detiene

La teoría de la información unifica problemas que pueden expresarse como alternativas, probabilidades, restricciones de codificación y dependencia estadística. Las mismas herramientas formales pueden describir compresión de fuentes, comunicación ruidosa, inferencia y partes de la física estadística porque cada problema contiene una estructura de estados distinguibles e

incertidumbre.[1,2] Esa matemática común es sustantiva, pero no implica que cada uso de la palabra información nombre el mismo mecanismo físico.

El concepto tiene también fronteras dentro de la matemática. La entropía diferencial puede ser negativa y cambia bajo transformaciones de coordenadas y cambios de escala, mientras que la información mutua tiene propiedades de invariancia más fuertes.[2] Afirmaciones como «el sistema contiene N bits» requieren por tanto especificar representación, resolución y modelo probabilístico.

La división del trabajo en esta serie es deliberada. IDE-003 trata la entropía como concepto más amplio, incluido su papel termodinámico; IDE-008 trata la probabilidad en sí; IDE-024 pregunta qué es una señal, centrándose en soportes físicos, propagación y condiciones de señalización. IDE-002 se detiene en la información cuantitativa: cómo se representan y restringen alternativas, incertidumbre, dependencia y codificación. El relevo es conceptual, no probatorio, y por ello no necesita cita.

En qué punto estamos

La información es una familia rigurosa de relaciones cuantitativas, no un ingrediente material universal. Su potencia procede de separar lo que depende de probabilidad y codificación de lo que exige semántica, estructura causal o un mecanismo físico de transmisión. Esa frontera hace complementarios, y no redundantes, los artículos vecinos: entropía, probabilidad y señal heredan cada uno parte de la historia, pero ninguno es simplemente otro nombre para información.

Lo que sabemos y cómo lo sabemos

- Para fuentes discretas, la entropía de Shannon cuantifica incertidumbre media y fija la escala fundamental de la compresión sin pérdida.[1,2]
- La información mutua mide dependencia estadística y obedece la desigualdad de procesamiento de datos bajo la estructura probabilística especificada.[2]
- El borrado lógicamente irreversible tiene un límite termodinámico inferior en el límite térmico ideal, mientras la computación lógicamente reversible puede en principio evitar pagar ese coste en cada paso.[3,4]
- La dependencia estadística por sí sola no identifica dirección causal; las afirmaciones causales requieren supuestos estructurales adicionales o intervenciones.[5]

Lo que todavía no sabemos

- Cómo conectar mejor magnitudes de teoría de la información con nociones semánticas o funcionales en sistemas biológicos y cognitivos concretos sin confundir definiciones.
- Qué medidas de información resultan más informativas para sistemas muy no estacionarios y de alta dimensión cuando las propias variables de estado relevantes son inciertas.
- Hasta qué punto futuras arquitecturas físicas de computación podrán acercarse a los límites de computación reversible sin perder velocidad, fiabilidad y fabricabilidad.

Qué podría hacernos cambiar de opinión

- Una violación reproducible de los límites de codificación de fuente sin pérdida o de codificación de canal bajo sus supuestos declarados obligaría a revisar el marco correspondiente de teoría de la información.
- Un experimento controlado de borrado que, tras contabilizar todos los reservorios y errores, disipara repetidamente menos que el límite de Landauer para un bit imparcial en el régimen térmico cuasiestático pondría en cuestión la explicación termodinámica estándar.[3]
- Un procedimiento de inferencia causal que identificara de forma única la dirección a partir de distribuciones conjuntas observacionales arbitrarias sin supuestos adicionales, y funcionara también con causas comunes y equivalencia observacional, derribaría la frontera aquí trazada entre dependencia y causalidad.[5]

Si solo recuerdas una cosa

La información cuantifica incertidumbre y dependencia estadística bajo elecciones explícitas de variables y probabilidades; no es significado semántico, influencia causal ni una sustancia física.

Artículos relacionados

- IDE-003 — Entropía
- IDE-008 — Probabilidad
- IDE-024 — ¿Qué es una señal?

Referencias

- [1] Claude E. Shannon. A Mathematical Theory of Communication. Bell System Technical Journal 27, 379-423 and 623-656 (1948). DOI 10.1002/j.1538-7305.1948.tb01338.x; 10.1002/j.1538-7305.1948.tb00917.x.
- [2] Thomas M. Cover; Joy A. Thomas. Elements of Information Theory, 2nd ed. Wiley (2006). DOI 10.1002/047174882X.
- [3] Rolf Landauer. Irreversibility and Heat Generation in the Computing Process. IBM Journal of Research and Development 5(3), 183-191 (1961). DOI 10.1147/rd.53.0183.
- [4] Charles H. Bennett. Logical Reversibility of Computation. IBM Journal of Research and Development 17(6), 525-532 (1973). DOI 10.1147/rd.176.0525.
- [5] Judea Pearl. Causal diagrams for empirical research. Biometrika 82(4), 669-688 (1995). DOI 10.1093/biomet/82.4.669.