
INFORMACIÓN, GENÉTICA Y ENTROPÍA¹

JULIO ERNESTO RUBIO BARRIOS

ABSTRACT. INFORMATION, GENETICS AND ENTROPY

The consolidation of the informational paradigm in molecular biology research concluded with a system that converts one epistemic object into an operational technological item and a stable epistemic product. However, the acceptance of the informational properties of genetic acids failed to clarify the meaning of the concept information. "Information" as a property of the genetic molecules remained as an informal notion that allows the description of the mechanism of inheritance, but was not specified in a logic-semantic structure. The metaphorical implications associated with the idea of genes as molecules with meaning questioned the linguistics that seemed too foreign to molecular biology. A reformulation of the concept of information in molecular biology was developed upon the theory of Claude Shannon. The node for the structural coupling between biology, physics and information theory was the identification of an analog structure between the coded messages of Shannon's theory.

KEY WORDS. Theory of information, molecular biology, Maxwell's demon, genetic information, inheritance, coded messages, genetic information transfer, coding.

1. INTRODUCCIÓN: TEORÍA DE LA INFORMACIÓN DE SHANNON

El concepto de información se ha convertido en una noción importante para muchos niveles del conocimiento a partir de las elaboraciones que Claude Shannon (1948, 1949) realizó, a finales de los años cuarenta, para optimizar los procesos de transmisión de señales codificadas. Aunque el título original que Shannon dio a su trabajo fue *teoría matemática de la comunicación*, este nombre fue sustituido en la práctica por el de *teoría de la información*, dado su enfoque en el proceso de optimización de la codificación, es decir, en la especificación del contenido informacional de las señales, más que en el proceso de generación y transmisión de las señales físicas, que se asocia de modo general en ingeniería eléctrica con el término 'comunicación'. La teoría de la información no trata directamente sobre las señales físicas sino sobre los mensajes codificados. Así, es posible

Decano de la Escuela de Humanidades y Educación, Tecnológico de Monterrey, Región Ciudad de México. / jerb@itesm.mx

realizar un análisis matemático de “... la medida de la información, la capacidad de un canal de comunicación para transferir información y la codificación como un medio de utilizar los canales a toda su capacidad” (Carlson y Crilly, 2009, p. 698).

Shannon estaba interesado en los principios de diseño de los sistemas de transmisión y recepción de señales que minimizaran la probabilidad de error en el proceso. Así, concibió una definición de información en función de la probabilidad de ocurrencia de un mensaje: la información (I) se definió como el logaritmo (base b) del inverso de la probabilidad de ocurrencia del mensaje (A):

$$I = \log_b (1/P_A)$$

Veamos los presupuestos básicos de esta definición: (i) La información depende exclusivamente de la probabilidad de ocurrencia del mensaje y *no del contenido semántico*. Si un mensaje es poco probable, contiene mucha información; si es muy probable, contiene poca información. En los casos extremos, un mensaje con probabilidad uno contiene cero información; por el contrario, un mensaje con probabilidad cero contiene *infinita* información. (ii) La definición ofrecida tiene sentido si es posible asignar una probabilidad a los mensajes, lo que implica la existencia de un conjunto de mensajes posibles (espacio muestral) donde podemos hacer la asignación de probabilidades, es decir, hablamos de probabilidades calculables *a priori* sobre la ocurrencia específica de un mensaje, pero *a posteriori* sobre el conjunto de señales posibles.

Si observamos la definición de información que la teoría de Shannon ofrece, nos encontramos, antes que nada, con una definición matemática. El uso de las matemáticas se encuentra, por supuesto, en áreas que no están restringidas a aplicaciones tecnológicas. Las matemáticas constituyen un elemento fundamental del trabajo científico y, particularmente, de las ciencias naturales. La distinción entre el uso matemático y el uso científico de las matemáticas puede explorarse desde la perspectiva pragmática que caracteriza a los procesos de producción tecnológica. Las matemáticas son, en una disciplina tecnológica, una herramienta de cálculo que se integra en el proceso productivo del aparato técnico. El cálculo que se realiza no tiene sentido en sí mismo si no contribuye a la producción de productos tecnológicos. Por otro lado, nos encontramos en la ciencia con una utilización epistémica de las matemáticas como lenguaje descriptivo. Las matemáticas funcionan en este caso como el lenguaje idealizado para modelar la realidad. Sin embargo, la distinción no implica que no haya utilización pragmática de las matemáticas en la ciencia, o que la tecnología no incluya elementos de producción cognoscitiva. Por el contrario, la tecnología modifica la realidad y crea espacios significativos que transforman nuestros cuerpos de conocimiento. Por su parte, la ciencia produce conocimiento como resultado de procesos de orden tecnológico que utilizan

eventualmente modelos matemáticos. Así, la frontera entre ciencia y tecnología es difusa, especialmente en la ciencia contemporánea, y la relación no se reduce al papel de la tecnología como mecanismo de validación-experimentación, o resultado práctico de aplicación del conocimiento científico.

La producción material de objetos tecnológicos en el laboratorio, según una lógica de disección *ad-hoc* de la naturaleza, nos muestra el primer orden de aproximación para entender el papel de la tecnología en la construcción del conocimiento científico. De este modo la maquinaria molecular se develó bajo las restricciones impuestas por los dispositivos y objetos tecnológicos. Otro tipo de acoplamiento estructural entre ciencia y tecnología que ocurre en la construcción del concepto de información genética se da mediante la transferencia analógica de estructuras teóricas generadas en la teoría de la información. En este caso no hablamos de restricciones tecnológicas materializadas en el uso de dispositivos para la manipulación del sistema experimental, sino de estructuras comunicacionales generadas en un campo de trabajo teórico de orientación tecnológica, que son importadas por la biología molecular para dar sentido a una noción interna.

La teoría de la información es un producto tecnológico en tanto está orientada originalmente al análisis y optimización del proceso generativo de señales físicas en canales de comunicación que forman parte de sistemas materiales de manipulación de tipos específicos de energía. Además, esta teoría es una producción comunicacional, simplemente porque toda teoría es una construcción lingüística que opera como estructura comunicacional de una comunidad para la codificación de un fenómeno.

Después de su aparición en un contexto tecnológico, la teoría de la información encontró aplicación en otros campos teóricos. La posibilidad de formalizar nociones particulares en términos matemáticos pareció ser la principal razón para el interés que otras disciplinas tuvieron en esta teoría. El lenguaje científico tiene como una de sus orientaciones la construcción de códigos cerrados. El concepto de información de Shannon ofreció la posibilidad de acotar el concepto de información en aquellas aéreas que manejaran ideas asociadas con flujos de estructuras codificadas. Algunos sistemas particulares de conocimiento con tales características se acoplaron estructuralmente con la teoría de la información, y reestructuraron algunos de sus propios contenidos en función de esa interacción. El uso del concepto de información de Shannon se ha discutido en el análisis de disciplinas sociales como la pedagogía (Frank, 1966); psicológicas, como la teoría de la percepción de Abraham Moles (1966); tecnológicas, como la cibernética (Wiener, 1966); biológicas, como la genética molecular (Lwoff, 1962; Eigen, 1992) además en ciertas aproximaciones

a problemas propios de la física, como la paradoja del demonio de Maxwell.

Sin embargo, ya desde 1966 se discutía también la capacidad de la teoría de la información para producir analogías útiles en la ciencia. Mandelbrot (1966, p.61) declaraba entonces que la teoría de la información debía abandonarse como forma de aproximación a nuevas problemas interdisciplinarios: "Volviendo a la teoría de la información en sentido estricto, cuando se me pregunta hoy cómo aprenderla para poder participar en actividades interdisciplinarias... respondo sin titubear que verdaderamente no vale la pena aprenderla".

Como ocurre usualmente en la dinámica de la ciencia, principios normativos que no surgen de la dinámica comunicativa de la ciencia misma son generalmente ignorados. La utilización del concepto de información en otros campos no fue detenida por criterios limitantes como el ofrecido por Mandelbrot. En biología molecular, la teoría de la información proporcionó la estructura para la aplicación analógica de un paradigma estructuralista a la interpretación de la información genética.

2. TRANSFERENCIA ANALÓGICA DEL CONCEPTO DE INFORMACIÓN DE SHANNON A LA BIOLOGÍA MOLECULAR

Quizá el caso más espectacular de transferencia analógica fue la aplicación del concepto de información de Shannon en la biología molecular. El descubrimiento de la estructura bioquímica de los ácidos nucleicos (ADN y ARN) en los años cincuenta señaló el mecanismo fundamental de la herencia y abrió un territorio de investigación en biología. Se encontró que las secuencias genéticas se formaban por el encadenamiento químicamente arbitrario de bases nucleicas (en cuatro posibilidades: dos purinas y dos pirimidinas). La transmisión de caracteres hereditarios se interpretó entonces como un resultado causal de la decodificación de la información contenida en las secuencias genéticas. No tardó en hablarse de las cuatro bases como el alfabeto del código genético y de las secuencias genéticas como los programas de desarrollo de los sistemas biológicos. En tal contexto, resultó completamente natural la aplicación del concepto de información de Shannon. Las secuencias genéticas podían verse como mensajes escritos en un código especificado, donde las posibilidades combinatorias del alfabeto podían calcularse, formalizando de ese modo el concepto de información genética.

No obstante, la transferencia analógica del concepto de información desde la teoría de la información requería condiciones estructurales adecuadas para que la analogía encontrara un espacio en el código comunicacional de la biología. Estas condiciones estaban dadas en la biología con relación a la cuestión de la herencia, que se preguntaba por el contenido

de tipo informacional de la sustancia responsable de la transmisión generacional de caracteres. La cuestión ya se había planteado desde el siglo XIX por autores como Lamarck, el propio Darwin y, por supuesto, Mendel. Estas investigaciones siguieron su curso con el trabajo de Weismann y, después, con los trabajos en biología molecular que culminaron con el descubrimiento de la estructura del ADN y la elucidación de los mecanismos de la herencia a nivel molecular. El subsistema biológico de la herencia se encontró en un acoplamiento estructural entre la biología molecular y la teoría de la información cuando se notó la estructura isomórfica entre el ADN y los mensajes digitales, reforzada por el uso —contingente— del término *información* para designar la teoría de Shannon.

La aplicación de la definición de la teoría de la información al caso de las macromoléculas genéticas estableció condiciones aparentes para la clausura conceptual de un objeto epistémico hasta un nivel poco frecuente en la biología. No solamente se conocía la estructura molecular de la sustancia de la herencia y la secuencia de reacciones bioquímicas que convertían los contenidos hereditarios en proteínas, sino también se formalizaba matemáticamente el concepto de información investigado. Una mayor clausura de la definición no podía lograrse. Aun así, la matematización de la información genética no se clausuró fácilmente. Por el contrario, la definición matematizada de la información genética se convirtió en el corazón de una controversia sobre el significado de la información.

3. INFORMACIÓN SEMÁNTICA Y CONTEXTO

En lenguaje ordinario, la *información* se entiende como el conjunto de datos que extraemos de alguna experiencia. Esta experiencia puede ocurrir con un agente verbal o no verbal. Podemos obtener información de un noticiero y también podemos obtenerla de la observación de un árbol, de un paisaje, o de una pieza mineral. ¿Cuál es la fuente de la información? ¿Ésta se encuentra en el objeto que funciona como referencia, en el observador, o en el proceso?

Si pensamos en términos de lenguaje ordinario, hablamos como si la información se encontrara en el objeto y nosotros simplemente la extrajésemos. No obstante, la información no resulta de la mera percepción del objeto, sino de cierto tipo de procesamiento de la experiencia perceptiva. Cuando decimos que un objeto se conforma de tal y cual manera, sus dimensiones son tales y demás, es decir, cuando describimos el objeto, utilizamos un código verbal para realizar la conversión de la experiencia específica en información. Por supuesto, el código no se encuentra en el objeto. Es el agente quien codifica una experiencia para producir información. La generación de información requiere de un código; la información se obtiene mediante la codificación de una experiencia. No podemos decir

que la realidad en sí porta un código, aunque algunas experiencias son en sí mismas experiencias codificadas (v.gr., el intercambio verbal entre agentes humanos). De manera general, la adquisición de información requiere del uso de un código que permita la traducción de una experiencia en información.

La codificación implica la traducción de la experiencia singular según algún contexto simbólico de nivel más general que incluye una gama de posibilidades significativas. Por ejemplo, si observamos una hoja de árbol y concluimos esta hoja es verde estamos utilizando un código que contiene la gama de colores que somos capaces de identificar. Así, el proceso requiere de un contexto de interpretación de nivel más general que la experiencia singular codificada. Bateson, analizando la estructura de la explicación cibernética, declara:

La jerarquía de contextos dentro de contextos es universal en el aspecto comunicacional (o "émico") de los fenómenos y lleva siempre al hombre de ciencia a buscar la explicación en unidades cada vez más amplias. En la física puede (quizá) ser verdad que la explicación de lo macroscópico deba buscarse en lo microscópico. En la cibernética suele ser verdad lo opuesto: sin contexto no hay comunicación (Bateson 1967, p. 340).

La naturaleza contextual, apuntada por Bateson para explicar procesos comunicacionales cibernéticos, es en el caso humano de orden *significativo*. Esto es, una experiencia con contenido informacional tiene sentido dentro de la red conceptual que el individuo posee. No podemos hablar del significado como si éste fuera un objeto o cosa. El significado existe no con relación exclusiva al objeto al que atribuimos la información sino en el conjunto de relaciones comunicacionales que el individuo tiene con su comunidad lingüística y en los patrones conceptuales con los que codifica sus experiencias cognoscitivas.

Podemos entonces concluir esta pequeña caracterización semántica de la información estableciendo los niveles participantes en cualquier proceso de adquisición de información:

- 1) El nivel del objeto físico al que atribuimos contenido informacional: un fenómeno concreto o señal física (una imagen, un conjunto de sonidos, o una configuración de algún tipo de energía).
- 2) El nivel de la codificación de la señal.
- 3) El nivel del contexto de interpretación significativa que da sentido al código.

Cualquiera que sea el caso de la fuente de información semántica, encontraremos los tres niveles indicados: una manifestación energética, un código y un contexto de generación e interpretación del código.

Desde la acepción contextual de la información, Bateson entendió el flujo de información como la detección o noticia de la diferencia y definió la información como una diferencia que hace una diferencia. El autor lo explica así:

... cuando hay información o comparación hay para mí un proceso mental... Un órgano sensorial es un órgano que compara, es un dispositivo que responde a una diferencia. Desde luego, el órgano es material, pero es su condición de responder a la diferencia lo que nos permite distinguir su funcionamiento como "mental"...nuestras explicaciones, nuestros tratados sobre la materia inanimada, están colmados de información. Pero esa información es toda nuestra, es parte de nuestros procesos vitales (Bateson, p. 30).

Dado nuestro interés por explorar el sentido del concepto de información, utilizó la cita para confirmar el aspecto contextual ineliminable de la información. Lo que Bateson llama aquí *lo mental* representa la dimensión contextual ineliminable de un proceso donde hay flujo de información. Bateson está tratando de construir un concepto de información aplicable en la trama conceptual de la cibernética. Esta disciplina se ocupa de sistemas autocontrolados que requieren capacidad de detección de su ambiente y generación de respuestas. Parece aplicar en este caso la idea de la noticia de la diferencia como una propiedad de un mecanismo que se autorregula. Por otro lado, el contexto significativo humano que se requiere para la interpretación de códigos apunta también al problema de la conciencia. La cibernética se encuentra justo entre la concepción organicista de una máquina y una concepción mecanicista de un ser vivo. La cuestión relevante es entonces si la naturaleza de la información es la misma en el caso humano que en el caso de un dispositivo mecánico autorregulado y, más en general, en un sistema cualquiera que es capaz de responder a las diferencias que una señal porta. La respuesta debemos buscarla en el nivel del contexto de interpretación. Las diferencias van a surgir de algún arreglo físico y el sistema detector tendrá un código de transformación de la señal en alguna forma de respuesta que ejecuta una acción material concreta, pero es la naturaleza del código y su contexto generativo lo que será distinto cuando hablamos de un dispositivo retroalimentado, de un ser humano, de una mariposa, o de la maquinaria celular de decodificación del código genético.

La aproximación informacional en la biología molecular reconoce la existencia de un código genético como punto nodal de la explicación de la herencia biológica, y el uso ordinario de términos de orientación semántica procede de tal reconocimiento. Por otro lado, la teoría de la información de Shannon proporcionó una estructura para formalizar el concepto de información genética en términos matemáticos y composicionales. La

búsqueda de un sentido para el concepto de información en biología molecular acopló estructuralmente la noción proveniente de la teoría de la información y la noción contextual de la información proveniente del sentido semántico de la codificación. Con relación a ambas nociones, la información genética presenta diferencias definitivas. A menos que desdijemos las particularidades del lenguaje humano en la dimensión de la comunicación y de la vida mental, o que establezcamos un sentido claro para las nociones semánticas en biología molecular, resulta cuestionable la asignación de interacciones significativas entre las macromoléculas biológicas. La codificación genética no cumple la condición teleológica de los códigos semánticos humanos (a menos que postulemos un agente teleológico oculto: Dios, la naturaleza y demás). Aún menos podemos hablar de “significados” sin distinguir su sentido en el caso de la interacción entre moléculas. No obstante, la información semántica y la información genética comparten la necesidad de una aproximación funcional para su interpretación. Por el otro lado, aunque el ADN presenta una estructura isomórfica con los mensajes codificados de Shannon, la designación informacional de las interacciones biomoleculares tiene un aspecto contextual que no corresponde a la definición estructural de la información en la teoría de Shannon. La “información contextual” sólo tiene sentido en una interpretación funcionalista del sistema que demarca el contexto del flujo informacional, sea éste de orden semántico o de orden biológico. La “información” de Shannon es totalmente independiente del contexto (aunque éste sea necesario para la generación del código) y depende exclusivamente de las posibilidades composicionales de la codificación. La coexistencia de tales usos lingüísticos estableció una tensión entre la interpretación funcionalista (contextual) y estructuralista (composicional) de la información genética.

4. TEORÍA DE LA INFORMACIÓN Y BIOLOGÍA MOLECULAR: ¿FUNCIÓN CONTEXTUAL O PROBABILIDAD ESTRUCTURAL?

El concepto de Shannon parecía una solución natural a la definición de la información genética, primero, porque la estructura de las macromoléculas correspondía a la estructura de los mensajes codificados de la teoría de la información; segundo, porque la ecuación descriptiva de la información correspondía isomórficamente con el negativo de la entropía termodinámica, y tercero, porque la definición matematizada de la teoría de Shannon cumplía con los requisitos del paradigma newtoniano de clausura operacional del sistema de representación. Aun así, las reacciones críticas contra los alcances de la ecuación de Shannon en la biología molecular no tardaron en ocurrir. Por ejemplo, en 1962, Lwoff, que había formado parte del grupo del fago, reaccionó a las ideas que empezaban a plantearse en

el sentido de proponer la ecuación de Shannon para medir la *cantidad de orden* (definida como entropía negativa y calculada en función del número de posibilidades combinatorias de las bases del código) en un organismo:

... ciertos biólogos han tratado de aplicar la idea de neguentropía... al organismo vivo. Puede calcularse la probabilidad para que una secuencia dada de ácidos nucleicos se realice mezclando las bases al azar. Pero creo que ese cálculo no tiene ningún sentido... En realidad, el aspecto funcional del organismo vivo no debe descuidarse, el cálculo de la neguentropía utilizando las fórmulas de Shannon no se aplica en absoluto al ser vivo (Lwoff 1966, p.128).

La cantidad de información en una secuencia de ADN se calcula a partir del número de combinaciones posibles de bases nucleicas. Si se tienen N unidades en la secuencia de ADN, la cantidad de combinaciones posibles es $(4)^N$ ya que hay 4 posibilidades para cada posición. La cantidad de información es proporcional a la inversa del número de posibilidades. Lwoff (1962) hace el cálculo para la bacteria *Escherichia Coli*, utilizando la constante de Boltzmann como constante de proporcionalidad, y obtiene un valor de 2×10^{-9} unidades de entropía. La pregunta relevante es acerca del significado de la cantidad obtenida en un contexto biológico. Lwoff defiende que la ecuación de Shannon no describe el aspecto funcional que ha sido utilizado tradicionalmente para construir explicaciones en biología. Desde esta perspectiva, el cálculo de Shannon aplicado a las secuencias genéticas no dice absolutamente nada acerca de la operación biológica.

Cuando se calcula el valor de la entropía en un canal de comunicación donde fluyen señales codificadas, el cálculo de la información en un mensaje es una medida de su improbabilidad (o de su probabilidad si se prefiere) de ocurrencia. El proceso de comunicación de señales transcurre en el tiempo y la cantidad de información de un mensaje dado tiene sentido en el ensamble de señales transmitidas. La cantidad de información de la señal nos dice cuál es la frecuencia de ocurrencia del mensaje. Ahora bien, no hay un fenómeno equivalente en el caso biomolecular. La cantidad de información en una secuencia de ADN no mide la frecuencia de ocurrencia del mensaje contenido en una secuencia particular, pues si este fuera el caso, la probabilidad de ocurrencia del mensaje es prácticamente igual a uno, dado que en los procesos internos de un organismo, el ADN permanece casi siempre inmutable. Por lo tanto, la información calculada en términos de Shannon es prácticamente igual a cero. Habría que especificar el ensamble de mensajes posibles para asignar probabilidades de ocurrencia y de esa forma obtener un cálculo con algún sentido. La cantidad de información obtenida en función de las combinaciones nucleicas no tiene sentido en el uso original de la ecuación de Shannon, porque no existe un ensamble de secuencias de ADN posibles

donde la información de una secuencia particular nos indique cuál es su probabilidad de ocurrencia en el ensamble.

La especificidad funcional del orden macromolecular no puede ignorarse en la aclaración del significado del concepto de información genética. Lwoff ejemplifica radicalmente este punto en el caso de las mutaciones:

Como resultado de un accidente, a veces una base nucleica del material genético es remplazada por otra. Esta alteración puede causar la muerte del organismo. Si, por ejemplo, una molécula dada de guanina en un gen es remplazada por una molécula de adenina, la información, la negaentropía del sistema, es la misma. Para el físico, aun cuando la mutación sea letal, nada ha cambiado: el contenido de negaentropía sigue siendo el mismo (Lwoff, 1962, p. 113).

Como vemos, el argumento de Lwoff nos muestra que el sentido de la información genética presenta una dimensión que no está presente en el mero orden estructural del sistema físico. ¿Existe información genética en una secuencia de ADN biológicamente inoperante?

Cuando nos referimos cualitativamente a la información contenida en el ADN suponemos la posibilidad de generación de un organismo a partir de una estructura orgánica mínima que contiene una secuencia de ADN biológicamente operante. Una secuencia arbitraria de ADN o ARN que no participa en la generación de un sistema biológico funcional no puede ser designada como información genética. No obstante, la secuencia aun tendría una cantidad de información asociada, en el sentido de Shannon, aunque careciera de sentido en términos biológicos.

Concluimos, entonces, que la aplicación de la ecuación de Shannon al cálculo del contenido informacional en las macromoléculas genéticas no produce ningún conocimiento relevante del lugar de la información en el orden biológico, específico y funcional. La aplicación de la ecuación de Shannon para intentar cuantificar la información genética está mal fundamentada en una confusión del tipo de contextos operantes en ambos casos. No es posible trasladar el contexto estadístico, que da sentido a la información en la teoría de Shannon, al contexto funcional de la biología molecular.

5. TRANSFERENCIA ANALÓGICA DEL CONCEPTO DE INFORMACIÓN DE SHANNON A LA FÍSICA

Aparentemente, la develación del mecanismo molecular de la herencia había mostrado la inexistencia de las nuevas leyes de la física que habían sido sugeridas por Böhr y buscadas por el grupo del fago. De tal modo, se excluía la posibilidad de reducir la biología a la física en el campo específico de la genética molecular, donde las posibilidades de un reduccionismo

interteórico tienen un territorio natural de prueba, porque se dan en condiciones de máxima especificación estructural y mínimo nivel de orden funcional. Los mecanismos de transferencia de la información genética parecían demostrar la naturaleza funcional irreductible de las operaciones biológicas, aun en el nivel molecular. La terminología que Zuckerland y compañía proponían destacaba el orden funcional de operación de las macromoléculas genéticas: si el ADN contenía la información codificada que regulaba el orden orgánico, otras moléculas tendrían la función de decodificar el mensaje para el resto del organismo.

Sin embargo, la aplicación de la definición matemática de información de Shannon al caso genético reavivó la tensión reduccionista entre física y biología. La ciencia persigue la obtención de modelos clausurados de reconstrucción lingüística de los objetos de estudio. En este sentido, la física clásica proporcionó un modelo de cientificidad que se fundamentó en la construcción de sistemas cerrados de variables, al que se conoce como paradigma newtoniano (Rosen, 1985, Martínez, 1990). La aplicación de este modelo de construcción representacional, podemos observarla en el caso del concepto de información genética como la aplicación del modelo composicional y estructural de la teoría de la información al caso molecular.

Con todo, la aplicación del paradigma composicional no fue el único modo de acoplamiento comunicacional entre física y biología molecular. A partir de las definiciones de Shannon, se inició una nueva defensa de la posibilidad de reducir eventualmente las teorías biológicas a teorías físico-químicas. La posible reducción de la biología a la física encontró una zona de exploración teórica en la existencia de otro isomorfismo estructural, esta vez no entre mensaje digital y secuencia genética, sino entre la forma matemática de la ecuación para calcular la cantidad de información y la ecuación para calcular la entropía de un sistema. La entropía se asocia con la cantidad de desorden existente en un sistema termodinámico; por el contrario, la información genética podía verse como el origen mismo del orden biológico. La información genética se planteó así como medida de la conversión termodinámica de sistemas físicamente desordenados a sistemas biológicamente organizados.

La relación entre entropía e información no aparece inicialmente con relación al problema específico de la información genética. En 1929, Szilard proponía una relación entre la producción de entropía y la intervención de seres inteligentes en el funcionamiento de dispositivos que materializaran el comportamiento paradójico de sistemas donde el hipotético "demonio de Maxwell" redujera la entropía. Después, el problema es abordado por varios teóricos (por ejemplo Brillouin en 1951, Schrödinger en 1944) que introducen explícitamente el concepto de información de la teoría de Shannon con relación al problema del demonio de Maxwell. No es hasta después de las especulaciones hechas alrededor del demonio de

Maxwell que el concepto de información de Shannon se aplica a contextos biológicos y, después, a las secuencias genéticas moleculares.

El tratamiento clásico de Szilard no se ocupa directamente del concepto de información, como su trabajo es interpretado *a posteriori*. Szilard desea mostrar que la intervención de seres inteligentes en un dispositivo de disminución de entropía por intervención del demonio de Maxwell produce necesariamente una cantidad de entropía al menos de magnitud equivalente al decremento de entropía del sistema. Para estructurar su argumento, Szilard habla de procesos de *medición (measurement)*, *memoria (memory)* o *acoplamiento de parámetros (coupling)* y solamente menciona la noción de información en la introducción de su artículo:

A perpetual motion machine therefore is possible if—according to the general method of physics—we view the experimenting man as a sort of *deus ex machina*, one who is continuously informed of the existing state of nature and who is able to start or interrupt the macroscopic course of nature at any moment without the expenditure of work (Szilard, 2003, p. 111).

En su demostración, Szilard formaliza la noción de medición como acoplamiento entre variables que representan, por un lado, el estado del sistema y, por otro, un estado del dispositivo de medición. La noción de Shannon, que aparecerá hasta 1948, no tiene un antecedente directo en el trabajo de Szilard pero llegará posteriormente a formar parte de la discusión e incluso será atribuida al propio Szilard. Por ejemplo, Brillouin (2003, p. 134) escribe acerca de la paradoja del demonio de Maxwell: “The paradox was considered by generations of physicists, without much progress in the discussion, until Szilard pointed out that the demon actually transforms ‘information’ into ‘negative entropy’”.

A pesar de la interpretación posterior, Szilard nunca hace una declaración tan específica en su trabajo. Por lo contrario, Szilard define rigurosamente las condiciones de su demostración rechaza la formalización de los sistemas inteligentes o biológicos, y sustituye la participación de éstos por dispositivos modelables físicamente. Szilard trabaja un problema termodinámico en términos termodinámicos. El concepto de entropía del que Szilard habla es el concepto perteneciente a la física y no se propone una equivalencia entre información (o inteligencia u otra condición de conciencia) y entropía negativa. No obstante, su referencia explícita a los sistemas biológicos abre claramente la aplicación de su demostración a este tipo de sistemas y señala el terreno conceptual para el acoplamiento que se daría posteriormente entre teoría de la información y biología molecular.

El acoplamiento entre teoría de la información y física no es generado exclusivamente desde la sugerencia informal de Szilard sobre la conversión de información en entropía negativa. Desde la teoría de la informa-

ción, Shannon impulsa la identificación analógica entre información y entropía negativa, y construye su definición de entropía en comunicaciones a partir del isomorfismo entre la ecuación de la entropía termodinámica y las ecuaciones descriptivas de sistemas de señales. La ecuación de entropía en teoría de la información y en termodinámica es

$$H = -k \sum P_i \log P_i$$

Donde k es una constante y P_i representa la probabilidad de ocurrencia de un microestado (en termodinámica) o la probabilidad de ocurrencia de un mensaje (en teoría de la información).

Tal isomorfismo ya había aparecido desde 1928 en las ecuaciones que Hartley había desarrollado para la transmisión de información telegráfica. Dos décadas después, Shannon retomó y desarrolló la noción de Hartley para convertirla en una elegante teoría sobre la transmisión de mensajes codificados. Sin embargo, el isomorfismo entre ecuaciones no es suficiente explicación para la aplicación de un mismo término a dos ecuaciones que representan fenómenos distintos. La elección de Shannon del término "entropía" respondió también a intereses específicamente comunicacionales. Por recomendación de von Bauer, Shannon escogió el término "entropía" (Wicken, 1987), con el fin de explotar la carga semántica que el término ya poseía en la física.

La estrategia comunicativa de von Bauer tuvo un éxito innegable e inmediato. Wiener (1948) parece ser el primero que propone la identificación entre información y entropía negativa. El contexto de trabajo es la cibernética y la noción de información es aquella que puede aplicarse a sistemas retroalimentados. Un sistema cibernético se comporta teleonómicamente en función de su capacidad para captar una señal que indica la adecuación del sistema con su salida esperada. La señal retroalimentada es comparada con el nivel esperado y el comportamiento del sistema es modificado con relación al resultado de la comparación. El sistema sostiene entonces su orden operacional gastando energía para verificar constantemente su estado. La información podía asociarse en cibernética con flujos operacionales físicos en sistemas concretos y una definición matemática tendría la capacidad de clausurar el sistema de conceptos de la naciente cibernética de modo tal que estableciera vínculos validatorios con la física, además de proporcionar un concepto que expresara también el orden propio de sistemas biológicos y psicológicos que la cibernética se aprestaba a modelar.

Después de la aparición de la teoría de Shannon, la asociación entre información y entropía negativa se volvió un punto de referencia para la discusión sobre la naturaleza del orden desde el nivel físico hasta el nivel social. En el caso de la relación entre física y biología, la teoría de la información funcionó como un puente para el acoplamiento estructural entre esas dos disciplinas que volvían a encontrarse en una tensión de tipo

reduccionista. La identificación entre información y entropía negativa ha sido un punto de conjunción que ha mantenido esta tensión fundamental.

6. COMUNICACIÓN INTERDISCIPLINARIA ENTRE TEORÍA
DE LA INFORMACIÓN, FÍSICA Y BIOLOGÍA:
EL CONCEPTO DE ENTROPÍA

La identificación entre información y entropía (termodinámica) negativa enfrenta por sí misma una tensión donde la biología no aparece necesariamente. En 1951, Brillouin había señalado el problema de la identificación entre información (codificada) y entropía (física) negativa en objetos singulares que coincidirán en la cantidad absoluta de entropía negativa. Al hablar sobre los planteamientos que Wiener hace en cibernética, Brillouin declara:

Take an issue of *The New York Times*, the book on cybernetics, and unequal weight of scrap paper. Do they have the same entropy? According to the usual physical definition, the answer is "yes." But for an intelligent reader, the amount of information contained in these three bunches of paper is very different. If "information means negative entropy," as suggested by Wiener, how are we going to measure the new contribution to entropy? (Brillouin 2003, p. 123).

La identificación entre información y entropía negativa es el blanco de la réplica de Brillouin: la especificidad del valor de la información (codificada) puede ser muy distinta en casos particulares que, no obstante, pueden poseer la misma entropía física ². La mera identificación lingüística de entropía con desorden e información con orden es el fondo de esta tensión. La tensión ha permanecido como objeto epistémico de un sistema de investigación que ha producido originales elaboraciones teóricas y ha generado otro acoplamiento estructural entre física y biología.

Layzer (1990) propuso una modificación importante a la identificación original entre información y entropía negativa. Este autor hace notar que no es posible equiparar formalmente distintas manifestaciones del orden (físico, biológico, lingüístico, etc.), pero que sí es posible identificar una manifestación general de desorden. El orden en cada tipo y nivel de organización es específico, pero el desorden es el mismo. Partiendo de esta premisa, Layzer propone que la información, como noción de orden, no puede entenderse como mera entropía negativa, pero puede formularse en términos del desorden de un sistema. La información no depende directamente de la entropía del sistema, sino de la diferencia entre la entropía máxima (desorden máximo) y la entropía observada (desorden de facto):

Randomness is a numerical measure of disorder. In the same way, information, defined by the formula $Information = maximum\ randomness - actual\ randomness$ is a numerical measure of order. Information is thus potential randomness... Processes that increase randomness destroys information; processes that decrease randomness generates information (Layzer, 1990, pp. 27-8).

Luego, la condición de aleatoriedad (randomness) es definida como el logaritmo del número de microestados que corresponden a un macroestado dado. Esta es precisamente la definición física de entropía, desarrollada por Boltzmann en 1873, donde los microestados del sistema se establecen en función de una colección de variables dinámicas.

Layzer aplica su definición a las macromoléculas genéticas. La aplicación directa de la fórmula de Shannon a este caso —dice Layzer— confunde macroestados con microestados; solo unas pocas de las combinaciones de bases son funcionalmente adecuadas y contienen por tanto información. El resto de las combinaciones son informacionalmente vacías. El autor identifica los macroestados de la secuencia con las combinaciones funcionalmente adaptadas de bases nucleicas, y los microestados con las combinaciones posibles de bases. De esta manera, Layzer intenta atrapar la funcionalidad biológica en su ecuación del cálculo de la información y trata de superar la crítica constante que se ha hecho a la aplicación de la ecuación de Shannon al contexto biológico. Layzer intenta dar sentido a la aplicación de la ecuación física de entropía de Boltzmann a la genética molecular. Pero la relación fenoménica entre información y entropía no es trabajada por Layzer. El autor parece sugerir que la relación se da simplemente por la estructura matemática y no por sus contenidos significativos, aunque, con relación a la aplicación de la fórmula de Boltzmann a la genética, Layzer (1990, p. 31) afirma "... But Boltzmann's definition of randomness, like every mathematical definition, is a pure form, devoid of content. Let's see if we can't find an appropriate biological content for it".

Con todo, la formulación de Layzer no llega a establecer claramente la relación entre las ecuaciones que propone y la *interpretación biológica* de las ecuaciones. La distinción fenoménica entre orden termodinámico, biológico e informacional es introducida por el autor, pero la confusión de niveles de organización permanece en la equivalencia del desorden (*randomness*) de los distintos niveles. La orientación hacia una clausura unificatoria del sistema conceptual prevalece sobre la distinción de niveles fenoménicos. De cualquier modo, Layzer hace una aportación central al sistema de investigación de la información: la distinción entre macroestados y microestados. Esta distinción operó como una variación de la reproducción del sistema de investigación que permitió nuevos desarrollos en torno al problema de la información.

Al utilizar las ideas de Layzer y de otros autores (en especial Collier). Wiley y Brooks (1980) intentan desarrollar un aparato formal para establecer las relaciones entre información y entropía. En realidad, la intención de estos autores va mucho más allá de la relación entre información y entropía. Wiley y Brooks (1986, pp. 32-3) intentan construir una teoría general de la entropía como principio unificador en todos los niveles de organización propios de la biología. Con este propósito, desarrollan una teoría jerárquica de la información que trata de superar los defectos de la aplicación un tanto tosca que se hizo en el origen de la identificación entre entropía e información. Los autores toman la idea fundamental de Layzer sobre la distinción entre macroestados y microestados, y desarrollan un formalismo que abarca distintos niveles de organización anidados y pretenden resolver de este modo las relaciones físicas entre información biológica y entropía: "... biological information (in all of its hierarchical manifestations) is subject to the constraints of the second law of thermodynamics in the same way that energy flows are subject to the law in its more traditional thermodynamic manifestations".

La cuestión de la información genética aparece de nuevo en el trabajo de Wiley y Brooks. La definición de Shannon les parece a los autores insuficiente como una descripción completa de la información biológica y construyen por tanto una tipología que distingue la información genética como "información instruccional" (siguiendo la terminología de Collier). La información instruccional es una condición física del ADN que es postulada como una expresión intrínseca de los organismos, esto es, una condición autorreferencial del organismo que no es estructurada por el entorno, sino solamente por el sistema mismo. Además, la información instruccional del ADN es una condición física del sistema porque no es dependiente del observador. A diferencia de la interpretación semántica de la información, la información instruccional no requiere, de acuerdo con los autores, ser definida por un observador. Este tipo de información es una condición material localizable en los procesos internos de los organismos. La aclaración es más que pertinente para el propósito general de la teoría de Wiley y Brooks. Si han de buscarse relaciones específicas entre información biológica y entropía termodinámica, la información ha de ser planteada en términos fisicalistas. La fisicalización de la información biológica aparece como un medio que podría permitir a los autores alcanzar la clausura del sistema histórico de relaciones entre información y entropía. El sistema de relaciones que Wiley y Brooks proponen incluye el concepto de la información de Shannon, que intenta la clausura de relaciones entre las tres disciplinas involucradas (teoría de la información, biología y física). Wiley y Brooks (1986, p. 91) definen entonces la operación biológica de la información como un sistema de comunicaciones: "Because biological arrays are also communication systems, biological

array information is also a subset of message information (from communications theory)".

El concepto de información que usan Wiley y Brooks surge de la asociación entre teoría de la información, biología y física. La novedad que encontramos en la teoría de Wiley y Brooks es el nivel de elaboración de su trabajo teórico y de autoconciencia epistemológica. Los autores parten de la teoría de la información de Shannon, y complementan las definiciones utilizando la distinción de Layzer entre macroestados y microestados, que no aparece en los trabajos de Shannon o de Brillouin. Además, intentan capturar la complejidad de los sistemas biológicos considerando la organización jerárquica de los niveles de organización participantes, y entonces se proponen llevar a sus últimas consecuencias la relación entre información y entropía. La teoría jerárquica de la información que desarrollan Wiley y Brooks propone una serie de distinciones entre clases de información que serviría el propósito de generar las analogías formales para establecer los principios puente entre física y biología:

As a formal analogy with statistical thermodynamics, HIT [hierarchical information theory] implies a number of terminological equivalences... The total information capacity of the system is equivalent to the total entropy (H_{max}) of the system. The information content of the constraints is equivalent to the entropy of the constraints ($H_{max}-H_{obs}$), or to the macroscopic information..., or to the macroscopic portion of the total entropy... The information content of the system is equivalent to the entropy of the system (H_{obs}) or to the microscopic information... or to the microscopic portion of the total entropy... We must also show that these equivalences are useful and meaningful (Wiley y Brooks, 1986, pp. 72-3).

El proyecto teórico de Wiley y Brooks puede clasificarse entre los grandes proyectos unificatorios de corte fisicalista a la manera en que Oppenheim y Putnam (1958) postulan en su artículo clásico "Unity of science as a working hypothesis", esto es, un proyecto epistemológico unificador donde las teorías científicas de cualquier nivel de organización específico sean sustentadas por teorías en el nivel de organización inferior, de tal modo que finalmente toda teoría científica pueda sustentarse en la física. Wiley y Brooks estarían contribuyendo a este proyecto si efectivamente lograran la unificación que pretenden entre física y biología. La identificación entre información y entropía sería el principal puente que permitiría la reducción. Asimismo, el objeto epistémico del sistema de investigación sobre la relación entre información genética y entropía alcanzaría la clausura reduccionista buscada por Brillouin, Wiener y otros. A pesar de estos nobles propósitos, el sistema está lejos de haber encontrado su punto de clausura en la teoría de Wiley y Brooks por varias razones que ahora discutiremos.

Wicken (1987) ha presentado con mucha lucidez el problema de comunicación interdisciplinaria en el caso de la relación información-entropía. El concepto de entropía, nos dice el autor, ha pasado por varios campos teóricos que han hecho distintos usos significantes del término. La noción de entropía surge en la termodinámica clásica para indicar la dirección restringida de la dinámica de un sistema macroscópico. La entropía (S) fue definida por Clausius como la razón de calor (Q) a temperatura (T), del siguiente modo:

$$dS = dQ/T$$

La segunda ley de la termodinámica postula que esta razón siempre aumenta en cualquier proceso irreversible. Este es el sentido original del concepto "entropía", y una generalización del concepto, afirma Wicken, debe compartir esta característica fundamental.

Boltzmann introduce en termodinámica la distinción entre macroestados y microestados como una manera de construir una explicación causal que partiera de la dinámica de los componentes microscópicos (moleculares) de un sistema. Así, Boltzmann llega a su conocida ecuación, de la que es posible derivar matemáticamente la fórmula macroscópica de Clausius:

$$H = k \log W$$

donde k es la constante de Boltzmann y $\log W$ es el logaritmo natural del número de microestados equiprobables que corresponden al macroestado observado del sistema. Después de Boltzmann, Gibbs extiende el resultado para microestados no equiprobables:

$$H = -k \sum P_i \log P_i$$

donde P_i es la probabilidad de ocurrencia de cada microestado i. Esta forma matemática de la entropía física es la que da lugar a la analogía formal con la ecuación de entropía de la teoría de la información, como aparece en los libros de texto de teoría de la comunicación:

$$H = \sum P_i \log_2 (1/P_i)$$

En esta fórmula, P_i es la probabilidad de ocurrencia de cada mensaje del sistema, la entropía H es la cantidad de información promedio en el sistema de señales. La ecuación es la misma, dado que k es una constante de proporcionalidad que ajusta el sistema de unidades, la base del logaritmo es arbitraria y $\log_2 (1/P_i) = -\log_2 P_i$.

Es la analogía matemática entre esta ecuación y la fórmula de Gibbs la que da sustento estructural a la relación entre información y entropía. Sin embargo, la equivalencia matemática no tendría por qué garantizar la equivalencia conceptual, como hace notar Wicken. Los contenidos conceptuales de las ecuaciones, en la teoría de la información y en la mecánica estadística, son fundamentalmente distintos —defiende el autor. Wicken señala que los microestados termodinámicos son estados cinéticos de las partículas de un sistema cuya incertidumbre radica en la fluctuación estadística de las partículas a través del espacio de fase. En cambio, las

probabilidades de la teoría de la información indican incertidumbre con respecto al ensamble de mensajes posibles, pero, una vez que un mensaje ha sido enviado, la incertidumbre desaparece como una condición del estado del mensaje. Por el contrario, el microestado de las partículas termodinámicas permanece fluctuando en el espacio de fase. Wicken concluye que la distinción entre microestado y macroestado que se aplica en la mecánica estadística carece de sentido en la teoría de la información.

De la misma distinción entre macroestado y microestado se desprende la diferencia conceptual entre la entropía termodinámica y la entropía en la teoría de la información. Wicken lo expone así:

What allows us to assign a thermodynamic system an entropy is that any measurable macrostate in which it resides can be expressed in a variety of alternative microstates. Since these are all accesible by the system, there is an essential uncertainty in knowing its microstate at any instant...A message, in contrast, cannot possess entropy. It is what has been said, a *fait accompli* (Wicken, 1987, p. 181).

Wicken indica el lugar fundamental de la incertidumbre de los microestados termodinámicos para mostrar que sin tales microestados, la entropía informacional es una cantidad conceptualmente ajena a la entropía termodinámica. Los mensajes de la teoría de la información son estructuras definidas sin microestados alternativas. No existe, por tanto, la relación fundamental entre microestado y macroestado que da sentido a la noción de entropía termodinámica, y no es posible usar la analogía formal para establecer una correspondencia fenoménica. Las estructuras especificables carecen de microestados alternativas y por tanto su entropía es cero. Tal es el caso de un mensaje especificado en la teoría de la información o de una secuencia genética: no poseen entropía en sentido termodinámico.

Por otro lado, la entropía de un sistema informacional, según la fórmula de Shannon, es una propiedad estadística de un ensamble que mide la expectativa promedio de ocurrencia de los mensajes del ensamble. Tal noción no es aplicable al contexto termodinámico porque no captura la distinción macroestado-microestado que Wicken ya señala como la condición fundamental de la entropía en mecánica estadística. Aquí debemos agregar que tampoco es el caso de las secuencias genéticas en biología. Las macromoléculas genéticas son estructuras estables. No existe un ensamble estadístico de mensajes genéticos. El ADN es un "mensaje" o un conjunto de mensajes que operan funcionalmente en procesos definidos por la estructura molecular, y no en procesos estadísticos.

Entonces, la doble traducción que lleva de la entropía termodinámica a la entropía informacional, y luego a la entropía genética, descansa en mapeos interteórico inválidos. La identificación de información y entropía negativa mezcla distintos niveles fenoménicos. La cantidad de informa-

ción medida por la fórmula de Shannon se aplica a ensambles estadísticos de secuencias codificadas. La información depende de la cantidad de unidades utilizadas para la codificación y de la cantidad de mensajes posibles, así como de su probabilidad de ocurrencia. La entropía de la teoría de la información no cuantifica una propiedad física de las señales, sino una propiedad matemática del conjunto de mensajes, es decir, la entropía negativa de Shannon es una propiedad de un arreglo lógico-matemático de signos cuya estructura física es irrelevante. Esta confusión fenoménica agrega una segunda dimensión problemática a la identificación entre entropía negativa e información. La entropía de Shannon sería la misma en un ensamble de señales materializadas físicamente por un alfabeto de ondas electromagnéticas de distintas frecuencias, señales de humo de distinta duración o barcos de tamaño distinto. En tanto el alfabeto se use en un proceso estadístico de emisión de las señales codificadas, la entropía de Shannon mide una propiedad del sistema de signos y no tiene relación directa con la entropía física del sistema de señales físicas utilizadas.

La aplicación de la definición matematizada de la información a las macromoléculas genéticas acusa ambos problemas. La "comunicación" entre macromoléculas biológicas no es un proceso estadístico de emisión de señales con distintas probabilidades asignadas. Las probabilidades de cada interacción entre moléculas es prácticamente uno porque cada contexto determina funcionalmente la interacción en curso. Por el otro lado, una entropía informacional de las secuencias genéticas, calculada a partir de la estructura de unidades asignadas como signos del código, sería la misma aunque el alfabeto base no fuera formado por ácidos nucleicos sino por banderas, en tanto la longitud de la cadena de signos base sea la misma y las probabilidades correspondientes a cada signo sean también las mismas. Evidentemente, la entropía física de cada arreglo no tiene relación alguna con la estructura de signos. La entropía termodinámica es una propiedad que solamente tiene sentido en el nivel físico. La entropía (o información promedio) de un ensamble estadístico de señales sólo tiene sentido en el nivel de organización de la representación matemática. La información genética es una propiedad de un sistema que existe en un nivel de organización intermedio entre los dos anteriores y no corresponde a ninguno de ellos. ¿Por qué entonces la persistencia de la identificación entre información y entropía (negativa)? Podemos sugerir que la utilización de las ecuaciones de Shannon permite la formalización matemática de conceptos que por su naturaleza funcional se oponen a la reducción estructural. La definición matemática permite la clausura del sistema conceptual bajo las condiciones de científicidad impuestas por el paradigma newtoniano de explicación: identificación de variables de estado y establecimiento de ecuaciones del sistema. Además, la formalización matemática de la información genética parece funcionar como un principio

puente que sustenta principios biológicos en principios físicos, lo que resuelve en apariencia la tensión reduccionista entre física y biología que ha acompañado la historia de la biología molecular.

El concepto de entropía es la moneda de cambio comunicacional utilizada para perseguir la clausura de un sistema de comunicación científica. Ello queda evidenciado por las razones que llevaron a Shannon a utilizar el término “entropía” en su teoría. Según explica Wicken:

To be fair: “entropy” was not a term that Shannon himself had decided on simply by virtue of the formal similarity of his equation to Boltzmann’s... he did so at the urging of von Newmann for very practical reasons. Entropy was an established concept that his theory could draw on for salability to the scientific community (Wicken, 1987, p. 183).

Los principios reduccionistas en el trabajo científico han operado como mecanismos orientados a la clausura operacional al establecer expectativas muy específicas sobre la estructura del conocimiento. Una de estas estructuras, que acompaña sobre todo a las ciencias naturales³, es el paradigma newtoniano de explicación. La construcción del concepto de información genética incorporó esta estructura de expectativas al transferir la definición matematizada de Shannon de la teoría de la información a la genética molecular. Otra estructura de expectativas en la ciencia proviene de las corrientes reduccionistas que plantean proyectos unificatorios de teorías pertenecientes a distintos niveles de organización. También encontramos este caso en la historia de la información genética. Las relaciones analógicas que se establecen entre entropía informacional y entropía termodinámica tienen precisamente el propósito de encontrar el mecanismo de unificación entre física y biología. La fuerza sociológica de estas expectativas puede ser tan fuerte que resista, como ha ocurrido en el caso del concepto de información genética, el condicionamiento de estructuras epistémicas particulares de los campos teóricos en juego. Las analogías que han conducido a las concepciones informacional y neguentrópica de la información genética pudieron seguir su propia dinámica una vez transferidas de la física y la teoría de la información; por su parte, la estructura de expectativas reduccionistas ha permanecido en el sistema de investigación de la información, como podremos aun ejemplificar en una teoría reciente sobre el origen de la vida.

7. EL SISTEMA DE INVESTIGACIÓN DE LA INFORMACIÓN GENÉTICA EN TIEMPOS RECIENTES:

LA TENSIÓN FUNCIONAL-ESTRUCTURAL Y LA METÁFORA SEMÁNTICA

Abordaremos en esta sección algunas ideas recientes sobre la información genética, a fin de mostrar que la investigación sobre la naturaleza biológica

de la información genética ha mantenido en las últimas décadas ⁴ la tensión entre la aproximación funcional y la aproximación estructural.

El concepto de información de la teoría de Shannon encontró un espacio de aplicación en el campo de estudio del origen de la vida y en la modelación de propiedades de los sistemas vivientes ⁵. Recientemente, algunos autores han concentrado sus ideas sobre el origen de la vida alrededor del concepto de información porque éste permite, aparentemente, establecer relaciones conceptuales entre el nivel de organización físico-químico y el nivel biológico, mediante el concepto de entropía negativa, y porque ofrece una plataforma formal para modelar entidades primigenias en función de cadenas de unidades de codificación. Manfred Eigen (1992, y Schuster 1979, et. al., 1981) postuló una teoría sobre el origen de la vida a la que se conoce como *teoría del hiperciclo*. En esta teoría, el concepto de información se coloca como la clave para entender el surgimiento de la vida sobre la Tierra. Eigen hace equivalente el origen de la vida con el origen de la información genética y caracteriza el proceso de evolución molecular en términos informacionales. Desde este punto de partida, el autor intenta construir una explicación del origen de la vida sustentada en principios físico-químicos.

Eigen estructura el problema del origen de la vida en función del concepto de información: resolver el problema del origen de la vida es equivalente a encontrar el origen de la información genética. Además, la información de las entidades biológicas surge, según Eigen, bajo la extensión del principio de selección natural a sistemas moleculares protobiológicos. El problema que el autor se plantea es la develación del principio de organización que conduce al surgimiento de la vida. La teoría es llamada *teoría del hiperciclo* en referencia al mecanismo de interacción primigenia entre secuencias genéticas y secuencias peptídicas (enzimas), dado en ciclos de replicación contenidos en ciclos de mayor orden (hiperciclos), donde cada secuencia participa en la reacción de replicación de alguna otra de las secuencias componentes. El escenario planteado propone la acción de fuerzas selectivas sobre poblaciones de secuencias primigenias formadas por condiciones físico-químicas carentes todavía de información biológica. La evolución aparece entonces a nivel molecular y conduce a la selección y optimización de las cadenas genéticas primordiales. El concepto de información es el eje conceptual de esta elaboración. De hecho, el origen de la vida se hace equivalente a la emergencia de la información genética: "All the varieties of life have a common origin. This origin is the information that, in all living beings, is organized according to the same principle" (Eigen, 1992, pp. 16-7).

Eigen define las posibilidades combinatorias de las secuencias de bases moleculares como estados informacionales del sistema. Cada estado informacional del sistema corresponde con una secuencia de bases cuyo con-

tenido informacional puede calcularse utilizando las ecuaciones de la teoría de la información. Además, esta concepción de la información permite la construcción de un espacio de estado para el sistema donde se representan los estados posibles, y permite también el tratamiento matemático en sistemas de ecuaciones diferenciales que modelan las entidades que se utilizan como posibles precursores de la vida orgánica. Específicamente, se construyen poblaciones estadísticas de secuencias llamadas *quasiespecies*, y poblaciones estadísticas de quasiespecies llamadas *hiperciclos*.

Sin embargo, a pesar de las intenciones reduccionistas de la teoría del hiper ciclo, el autor reconoce el límite impuesto a la noción de información absoluta de la teoría de Shannon porque no ha sido posible caracterizar el contenido informacional de las secuencias genéticas, de modo que puedan plantearse de forma satisfactoria las interacciones moleculares entre las secuencias genéticas y las moléculas funcionales sin recurrir a algún principio no contenido en la estructura molecular de los polinucleótidos.

El atractivo del concepto de información de la teoría de la información de Shannon radica en su formulación estructural de estados matematizados. Esto permitió la adecuación de la teoría del hiper ciclo al paradigma newtoniano de explicación que supone la posibilidad de descomposición de un sistema, y de planteamiento del estado del sistema total como suma de los estados de las partes.

La tendencia a privilegiar la explicación estructural en términos de estados sobre la explicación funcionalista es planteada explícitamente por Charles Bennett (1985) cuando intenta elaborar una definición de la organización biológica fundamentada en la analogía entre un sistema biológico y la capacidad de una máquina de cómputo (por ejemplo, la máquina de Turing) para modelar el proceso generativo de un sistema cualquiera ⁶. En consecuencia, Bennet (1985, p. 217) intenta una definición de sistema organizado como aquel susceptible de ser modelado en una máquina universal de cómputo, dadas las condiciones iniciales del sistema, pero descalifica su propia definición por no estar planteada en términos estructurales de estado: "... computational universality is an unsuitable complexity measure for our purposes because it is a functional property of systems rather than a structural property of states".

La fuerza persuasiva del paradigma newtoniano descansa en la posibilidad de obtener condiciones de clausura de los modelos realizados por definición de conjuntos cerrados de variables de estado que permitan la manipulación estable del modelo (Earman, 1985). A final de cuentas, la teoría del hiper ciclo de Eigen no logra resolver la tensión entre la naturaleza funcional o estructural de la información genética. El autor plantea la ecuación de Shannon como una posibilidad de tratar formalmente el problema del origen de la información genética, y traza incluso las vinculaciones del problema con la identificación entre entropía e información

negativa, pero no llega a darle ningún uso concreto al cálculo de la cantidad de información de Shannon. Incluso el propio Eigen aceptar la limitación de la ecuación de Shannon para describir el comportamiento biológico a nivel molecular. Al hablar de las reglas estadísticas del comportamiento de las secuencias genéticas, entendidas como mensajes escritos en un código, Eigen (1992, p. 12) afirma: "There are also grammatical and syntactical rules. The spectrum of rules and conditions covers even the meaning of the sentence, that is, the semantic information, which is dependent upon specific premises that cannot be described by general statistical frequency laws".

Por un lado, Eigen afirma la posibilidad de reducir el origen de la vida a leyes fisicoquímicas; por otro, sostiene la limitación de la descripción estadística (fundamentada en la teoría de la información) para describir la naturaleza de la información genética. La tensión que aparece en la teoría de Shannon no es tanto un accidente de la propia teoría, como un resultado de las condiciones científico-comunicacionales en que se genera. La teoría de Eigen hereda una tensión, que no logra resolver, entre la aproximación funcional y la aproximación estructural a la interpretación del comportamiento biológico de las macromoléculas genéticas. Las nociones de Eigen muestran la permanencia de la metáfora lingüística como descripción del orden informacional-biológico. Nuevamente se habla de una dimensión semántica del comportamiento biomolecular. La metáfora lingüística expresa el sentido contextual del término "información" y el orden funcional aparentemente irreductible de la operación biológica.

Otro intento para formalizar el concepto de información genética ha sido planteado por Bernd-Olaf Küppers (1990) en el libro *Information and the Origin of Life*. Las ideas presentadas por Küppers nos muestran la continuidad de la tensión entre explicación estructural y funcional alrededor del concepto de información en biología molecular. Küppers propone la existencia de tres aspectos de la información: sintáctico, semántico y pragmático. El aspecto sintáctico contempla la estructura y posibilidades cuantificables de la información; el aspecto semántico apunta a los contenidos significativos de la información que permiten la integración funcional de los distintos elementos de un sistema biológico; el aspecto pragmático considera los mecanismos observables de operación de las estructuras informacionales.

La formulación de Shannon persiste en el planteamiento de Küppers, aunque su aplicabilidad es explícitamente restringida a una descripción sintáctica de la información biológica, es decir, a su aspecto estructural. El genoma de un organismo posee una longitud característica que permite aplicar la ecuación de Shannon en función de la cantidad de alternativas que pueden formarse con la misma longitud. La persistencia de las ecuaciones de Shannon puede explicarse por la misma razón de fondo que ya

hemos señalado: el peso del paradigma newtoniano de cientificidad como estructura comunicacional incuestionable. El propio Küppers (1990, p. 33) sostiene la utilidad de la definición de Shannon en la posibilidad de la matematización del problema: “In order to make an information-processing operation accessible to mathematical analysis, we need a quantitative measure of the amount of information that is contained in the sequence of symbols”.

No obstante —Küppers aclara— la medida de la información proporcionada por la ecuación de Shannon es poco útil en el tratamiento del orden biológico. Pone como ejemplo la secuencia de bases genéticas de la bacteria *Escherichia Coli*, y afirma:

... Yet only a few of these carry biologically “meaningful” information, that is, information that can guarantee the sustenance of the functional order of the bacterial cell. We now begin to see that the structural information in a nucleotide chain that is given by the Shannon information measure is hardly able to make useful statements about biological complexity and order (Küppers, 1990, p. 48).

Este argumento no es novedoso. De hecho, es el ya viejo alegato que apareció desde los primeros intentos por aplicar la teoría de la información a la biología molecular y que vimos explicitado por Lwoff desde 1962. La cita de Küppers nos permite ver que la tensión funcional-estructural del concepto de información genética se ha mantenido fundamentalmente en los mismos términos. De igual modo, nos permite explicar el retorno de metáforas semánticas como un medio de señalar el orden funcional irreductible de las estructuras biológicas. Küppers lo expone del siguiente modo:

Everything about biological structures that works “according to a plan,” that is, everything that is controlled by information, has a “meaning” and a “significance” in the context of the functional order that we find in living systems. This means that biological information is associated with defined semantics. When in the rest of this chapter we refer to “biological” information, it will be precisely this semantic aspect of information that is meant (Küppers, 1990, p. 31).

Las categorías lingüísticas empleadas por Küppers aparecen esta vez en la conceptualización de la información genética, sustentadas por la autoconciencia epistemológica. Se trata precisamente de rescatar el orden funcional de la información que, como hemos dicho, se encuentra contenido en el sentido de lo semántico.

Además de los casos de Eigen y Küppers, podemos citar al menos a otro autor fundamental en la biología teórica contemporánea, quien también hace uso de categorías lingüísticas para tratar el problema de la informa-

ción biológica: Stuart Kauffman (1992, 1993). Este autor elabora una teoría de grandes alcances que intenta aproximarse de manera novedosa a una serie de problemas de la biología: el origen de la vida, el desarrollo ontogenético de los organismos y, en general, los procesos de generación y continuidad de la autorganización de los seres vivos. El concepto de información es abordado desde una perspectiva orientada al enfoque semántico. Kauffman nos da otra aproximación al problema del origen de la vida (el modelo de conjuntos autocatalíticos de polímeros) que parte de consideraciones informacionales inspiradas por la teoría del hiperciclo y el concepto de información de Shannon. En su teoría, Kauffman elabora un concepto de las reglas de interacción entre compuestos moleculares que llama "gramáticas aleatorias". Nuevamente, la metáfora lingüística aparece como un medio de representar el orden funcional del comportamiento biomolecular. Evita el énfasis en la molécula maestra informacional que parece ser el más serio obstáculo para la robustez de la teoría del hiperciclo (Wicken, 1985), Kauffman propone sistemas de polinucleótidos y polipéptidos que interactúan según un sistema de reglas abiertas que se transforman en la evolución del sistema como totalidad.

8. CONCLUSIÓN

Aunque el paradigma newtoniano sigue funcionando como un modelo de cientificidad que propone la construcción de espacios de fase matemáticos como traducción del objeto de investigación, la aplicación de la ecuación de Shannon (o de sus versiones hijas, como las ecuaciones de Layzer o Wiley y Brooks) se ve confrontada permanentemente por el orden específicamente funcional de la información biológica. La limitación de la aproximación matemático-estructural para describir el comportamiento biológico a nivel molecular, y el fracaso de los proyectos reduccionistas para subordinar la explicación biológica a la explicación fisicalista, han dejado el espacio abierto para las analogías lingüísticas en el tratamiento del problema de la información. En el lenguaje humano, el contenido semántico del intercambio simbólico es de orden funcional porque los sujetos participantes en un proceso comunicacional se acoplan estructuralmente a través de los significados compartidos de tal modo que un sistema emergente —la comunicación misma— se sostiene en el acoplamiento funcional de las partes. El lenguaje es la estructura que sostiene la continuidad del proceso comunicacional. Del mismo modo, el aspecto sintáctico (estructural) de la información biológica está subordinado funcionalmente a su aspecto semántico.

El sistema de investigación que ha tenido como objeto epistémico la pregunta por la naturaleza de la información genética no ha resuelto la tensión entre la aproximación estructural y funcional a su objeto de

estudio. La teoría de la información ha proporcionado los elementos de intercambio comunicacional mejor atrincherados de la aproximación estructural. No obstante, las ecuaciones de Shannon no han logrado ser interpretadas en un sentido biológico específico. Por otro lado, la metáfora semántica ha proporcionado los elementos de intercambio comunicacional mejor atrincherados de la aproximación funcionalista. Con todo, tampoco los conceptos de procedencia lingüística han llegado a un punto de resonancia conceptual que permita su atrincheramiento en un sentido exclusivamente biológico.

NOTAS

- 1 Artículo original: "Information, genetics and entropy", *Principia. An International Journal of Epistemology*. Brasil, 2015, vol. 19, num. 1, pp. 121-146. [1808-1711]. Versión en español del propio autor.
- 2 El argumento es fundamentalmente el mismo que Lwoff replica antes a la aplicación de las ecuaciones de Shannon a la genética molecular: la confusión entre el orden funcional y estructural.
- 3 Las ciencias sociales, sin embargo, no se encuentran al margen de estas estructuras de expectativas como muestra el caso de los programas conductistas en psicología o el uso extensivo de herramientas estadísticas en el análisis sociológico.
- 4 Por otro lado, la naturaleza de la información genética es un problema meta-teórico clausurado con respecto a las prácticas específicas en los diversos campos de la biología molecular, y seguramente desconocido para muchos de los científicos de esta disciplina. La biología molecular se ha convertido en una disciplina muy estable, cuyos fundamentos epistémicos son robustos. La localización de la información hereditaria en los ácidos nucleicos es un presupuesto epistemológico incuestionable en la biología molecular contemporánea, aunque recientemente se está acumulando evidencia contra la validez del dogma central (Torres, 1992). No obstante, la biología molecular es un campo de trabajo bien atrincherado en su cuerpo de conocimiento teórico (cajas negras) y técnicas de manipulación de sus objetos de estudio (objetos tecnológicos). Ni siquiera el eventual derrumbe del dogma central parecería una amenaza a la estabilidad de la biología molecular que ha dejado de depender de la robustez de este criterio desde hace décadas.
- 5 Estos dos problemas se encuentran estrechamente relacionados, pues para conocer el origen de la vida hay que especificar qué es lo viviente.
- 6 La caracterización computacional de los sistemas biológicos parece descansar en la concepción del ADN como programa generativo y la analogía que se establece con los programas de ordenador. Esta vertiente constituye otra fuente epistemológica de inspiración tecnológica para el concepto de información.

REFERENCIAS

- Bateson, G. (1985), "La explicación cibernética", in *Pasos hacia una ecología de la mente*. México: Planeta.
- Bennet, C. (1988), "Dissipation, information, computational complexity and the definition of organization", in D. Pines (ed.), *Emerging Synthesis in Science*. Santa Fe: Westview Press.
- Brillouin, L. (2003). "Maxwell's demon cannot operate: information and entropy", in H. Leff, A. Rex (eds.), *Maxwell's Demon: Entropy, Classical and Quantum Information, Computing*, London: Institute of Physics Publishing.
- Carlson, A. B.; Crilly, P. (2009). *Communications Systems*. New York: McGraw-Hill.
- Earman, J. (1985). *A Primer on Determinism*. Dordrecht: Reidel Publishing Company.
- Eigen, M. (1992). *Steps Towards Life: a Perspective on Evolution*. Oxford: Oxford University Press.
- Eugen, M; Schuster, P. (1979), *The Hypercycle: A Principle of Natural Self-Organization*. Heidelberg: Springer-Verlag.
- Frank, H. (1966), "Información y pedagogía", in M. Gueroult (ed.), *El concepto de información en la ciencia contemporánea. Coloquios de Royaumont*. Barcelona: Siglo XXI.
- Kauffman, S. A. (1993), *The Origins of Order. Self-organization and Selection in Evolution*. Oxford: Oxford University Press.
- Küppers, B. O. (1990), *Information and the Origin of Life*. Cambridge: MIT Press.
- Layzer, D. (1990), *Cosmogogenesis. The Growth of Order in the Universe*. Oxford: Oxford University Press.
- Lwoff, A. (1962), *Biological Order*. Cambridge: MIT Press.
- Mandelbrot, B. (1966), "¿Es útil "todavía" la teoría de la información?", in M. Gueroult (ed.), *El concepto de información en la ciencia contemporánea. Coloquios de Royaumont*. Barcelona: Siglo XXI.
- Martínez, S. (1990), "Más allá de la presuposición Newtoniana: propiedades genuinamente disposicionales en la mecánica cuántica", *Crítica* 22(66): 25–37.
- Moles, A. (1966), "Teoría informacional de la percepción", in M. Gueroult (ed.), *El concepto de información en la ciencia contemporánea. Coloquios de Royaumont*. Barcelona: Siglo XXI.
- Oppenheim, P. and Putnam, H. (1958), "The unity of science as a working hypothesis", in H. Feigl, et al., (eds.), *Minnesota Studies in the Philosophy of Science*. Vol. 2. Minneapolis: Minnesota University Press.
- Rosen, R. (1985), "Organisms as causal systems which are not mechanisms: an essay into the nature of complexity", in R. Rosen (ed.), *Theoretical Biology and Complexity*. Orlando: Academic Press.
- Schrödinger, E. (1944), *What is life?* Cambridge: Cambridge University Press.
- Shannon, C. E. (1948), "A mathematical theory of communication", *Bell System Technical Journal* 27: 379–423, and 623–56.
- (1949), "Communication in the presence of noise", *Proc. IRE* 37: 10–21.
- Szilard, L. (2003). "On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings", in H. Leff; A. Rex (eds.) *Maxwell's Demon: Entropy, Classical and Quantum Information, Computing*. London: Institute of Physics Publishing.
- Kourany, J. (ed.) (1987), *Scientific Knowledge*. Minneapolis: University of Minnesota Press.

- Torres, J. M. (1993), "Current notions of genotype and phenotype vis a vis new evidence in molecular biology. A methodological study", *Biology Forum* 86-2: 181-197.
- Wicken, J. S. (1985), "An organism critique of molecular Darwinism", *Journal of Theoretical Biology* 117: 545-61.
- (1987), "Entropy and information: suggestions for common language", *Philosophy of Science* 54: 176-93.
- Wiener, N. (1948), *Cybernetics or Control and Communication in the Animal and the Machine*. New York: John Wiley & Sons.
- (1966), "El hombre y la máquina", in M. Gueroult (ed.) *El concepto de información en la ciencia contemporánea. Coloquios de Royaumont*. Barcelona: Siglo XXI.
- Wiley, E. O.; Brooks, D.R. (1986), *Evolution as Entropy. Toward a Unified Theory of Evolution*. Chicago: Chicago University Press.
- Zuckerland, E.; Pauling, L. (1965), "Molecules as documents of evolutionary history", *Journal of Theoretical Biology* 8: 357-66.