



La verdadera revolución de la información aún no ha llegado

Descripción

Con motivo de su reciente fallecimiento, en febrero de este año, la figura del matemático e ingeniero Claude E. Shannon está siendo objeto de un renovado interés. **Alberto Solana Ortega** repasa lo que fue la vida y obra de este científico, centrándose especialmente en su contribución a la nueva formulación de la ciencia entendida como teoría de inferencia probabilista.

lvrdianh11.jpg type unknow

Aunque poco conocido hasta la fecha por el gran público, incluso por parte del especializado, entre otras razones debido a su temprano retiro y a que la mayoría de sus trabajos han estado clasificados como secretos durante largo tiempo, hay acuerdo en destacarlo como uno de los fundadores de varias de las disciplinas científicas modernas indispensables para el desarrollo tecnológico y cultural de la segunda mitad del siglo XX. Especial relevancia tienen sus aportaciones a la teoría matemática de la comunicación y a la teoría de la computación. En este sentido se le ha llegado a considerar el padre de la llamada Sociedad de la Información y se ha propuesto situarlo como uno de los grandes investigadores de la Historia.

Analizaremos aquí su trabajo desde una perspectiva diferente, más amplia que la habitual. Nos interesa explorar la idea de que, a largo plazo, su significación mayor quizá no resida tanto en los aspectos que son elogiados en la actualidad, como en el hecho de haber sido uno de los pioneros en el camino hacia una nueva concepción de la ciencia entendida como teoría de inferencia lógica probabilista.

Nacido en Estados Unidos en 1916, Shannon creció en un ambiente intelectualmente estimulante en el que resultó esencial su abuelo, inventor de la lavadora y de diversas máquinas agrícolas. Desde pequeño mostró una atracción paralela por las matemáticas, a la vez que por todo tipo de aparatos mecánicos y radios; una inclinación que mantendría durante toda su vida. Cursó estudios de Ingeniería eléctrica y Matemática en la Universidad de Michigan, obteniendo la licenciatura en ambas disciplinas en 1936. Posteriormente se doctoró en ellas en el Instituto de Tecnología de Massachussets (MIT). Allí conoció y trató con Norbert Wiener, creador de la cibernética.

Mientras estudiaba, trabajó como asistente para controlar el Analizador Diferencial de Vannevar Bush, la máquina calculadora más avanzada de la época, constituida por un sistema mecánico dirigido por un circuito de relés, y utilizada para resolver ecuaciones diferenciales. Su primera publicación, en 1938, fue su propia tesis en Ingeniería. En ella elaboró las ideas sobre la relación entre la lógica simbólica y el funcionamiento de circuitos, las cuales constituyen los principios operativos de los ordenadores actuales. Proponía, por un lado, la discretización de todo circuito elemental en dos estados, cerrado y abierto, y su representación mediante sendos valores (0 y 1) que se hacen corresponder respectivamente con la falsedad o verdad de una proposición lógica. Sugería, por otro, el empleo del álgebra de Boole para el análisis de circuitos más complejos. Sus resultados tuvieron un gran impacto y por ellos recibió el premio más importante que concedían las sociedades de ingeniería de Estados Unidos a menores de treinta años, el Alfred Noble.

En 1940 presentó su tesis en matemáticas, titulada *An Algebra for Theoretical Genetics*, en la cual investigaba cómo emplear la lógica para organizar el conocimiento genético, de modo análogo a como había hecho con los circuitos con interruptores y relés.

Al año siguiente completó una estancia con el matemático Herman Weyl en el Instituto para Estudios Avanzados de Princeton, y a continuación, en 1941, fue contratado en los Laboratorios Telefónicos Bell, donde ya había colaborado algunos veranos.

Allí trabajó durante los siguientes quince años, en compañía de científicos de primerísima fila, entre los cuales cabe citar a H. Nyquist, especialista en teoría de señales, J. Pierce, experto en comunicaciones por satélite, Brattain, Bardeeb y Shockley, inventores del transistor, y G. Stibitz, quien había construido, ya en 1938, un ordenador con relés. En este periodo investigó en áreas diversas, principalmente en el campo de la transmisión eficiente de la información. Después fue profesor en el MIT, y *fellow* entre 1957-58 en el Centro para el Estudio del Comportamiento en Palo Alto, California. Su personalidad resulta sumamente atractiva y ha dado origen a innumerables anécdotas. Siempre se le consideró como un «cacharrero» que, a la hora de resolver un problema, lo mismo proponía una idea técnica que sacaba la sierra y el martillo (en este sentido quizás habría que colocarlo en un lugar próximo a Edison, de quien, por cierto, era primo lejano). Aficionado al malabarismo, era famoso en los laboratorios por encerrarse durante el día en su despacho y montar de noche en monociclo

realizando figuras por los pasillos. La lista de sus inventos y patentes es interminable, al igual que la de premios y honores que le fueron concedidos.

En 1949 se había casado con Elizabeth Moore -licenciada en Matemática y colaboradora como analista numérico en el grupo de J. Pierce -; con ella tuvo tres hijos.

Shannon se jubiló a los cincuenta años, un poco para huir de la pesada tarea de divulgar sus pensamientos y, sobre todo, para poder dedicarse a sus ingenios. En una de sus últimas entrevistas dijo: «siempre he perseguido mis intereses sin prestar mucha atención a su valor financiero o para el mundo; he dedicado muchas horas a cosas totalmente inútiles». Murió tras varios años de lucha con el Alzheimer. El lector interesado puede encontrar dos biografías, junto con la mayor parte de su obra, en el libro *C. E. Shannon: Collected Papers*, editado por N. J. A. Sloan y A. D. Wyner (IEEE Press, 1993).

INVESTIGACIÓN MATEMÁTICA

Con respecto a su vida profesional, sin duda alguna sus trabajos más importantes fueron dos artículos seminales. Uno es el ya citado de 1938, titulado *A Symbolic Analysis of Relay and Switching Circuits*; y el otro, *A Mathematical Theory of Communication*, aparecido en 1948 y publicado en 1949 en forma de libro junto con otro artículo del sociólogo Warren Weaver bajo un título que ya es clásico: *The Mathematical Theory of Communication* (nótese el cambio de «una» a «la» teoría). En ellos se encuentran prácticamente todas las ideas que desarrollaría más adelante, a excepción quizá de su modelización matemática de los juegos malabares, la cual establece teoremas que ligan cantidades como el tiempo que un objeto está en el aire con otras como el número de manos que se utilizan para jugar o la velocidad de lanzamiento. Dichas ideas dieron lugar, junto con contribuciones de otros autores, a nuevas ramas científicas tales como la teoría de la computación, la robótica y la *teoría matemática de la transmisión de información*, mal llamada, por elipsis, como iremos discutiendo, *teoría de información*.

En relación con las primeras, baste destacar aquí la construcción de ordenadores y programas de cálculo, de algunas de las primeras computadoras para jugar al ajedrez, de diversos autómatas y robot, (como, por ejemplo, el ratón Teseo, capaz de encontrar la salida de laberintos), así como de máquinas «para leer la mente».

Por su parte, el libro sobre comunicación produjo una auténtica sacudida, y es que había una necesidad apremiante de una base teórica para las tecnologías de la comunicación existentes y en pleno desarrollo, debido al aumento de la complejidad de las redes telefónicas y de teletipos, y a la extensión de los sistemas de radio. Pero el impacto no se circunscribió a la ingeniería. El número de publicaciones de otros autores relacionadas con las nuevas propuestas creció rápidamente y su influencia llegó a todos los ámbitos del saber, notándose incluso en disciplinas artísticas tales como la pintura y la música. A este respecto es necesario subrayar el papel de Weaver como intérprete de la teoría, precisando algunos de sus aspectos filosóficos, y como impulsor de su difusión fuera del campo de las comunicaciones técnicas, delimitando sus posibles niveles de aplicación. No obstante, hay que señalar también que el empuje resultó excesivo y, en contrapartida, la expresión «teoría de información» llegó a convertirse en un comodín en toda conversación pseudointelectual.

Pero, ¿cuál es el contenido de una teoría con semejante capacidad? La de Shannon estudia los procesos de comunicación desde un enfoque, conviene recalcarlo, exclusivamente técnico en un principio. Con anterioridad habían existido otras visiones de estos procesos, pero la suya constituye el primer análisis formal de los mismos que va más allá de lo puramente descriptivo.

La entidad clave es el sistema de comunicación. Shannon, en primer lugar, propone un modelo lineal de transmisión en el que el sistema se considera formado por varios elementos, entre otros: una fuente emisora, que selecciona símbolos de un alfabeto dado, o bien un mensaje de un conjunto prefijado de mensajes posibles, para formar un mensaje o una secuencia de mensajes; un transmisor o codificador, que transforma los mensajes en señales adecuadas para una transmisión eficiente a través de un medio físico denominado «canal de comunicación», *empleando un código previamente acordado*; un descodificador, el cual traduce los mensajes codificados a mensajes en el alfabeto original, de modo que puedan resultar comprensibles cuando llegan al receptor. Suele añadirse además una fuente de ruido, cuyo efecto consiste en distorsionar los mensajes haciendo que, en general, no tenga por qué haber correspondencia entre el mensaje recibido y el enviado, lo que se entiende como un error del proceso.

A continuación, tomando la idea de Wiener de describir un sistema de comunicación en términos estadísticos, Shannon representa probabilísticamente cada uno de sus elementos. Así, por ejemplo, el receptor queda caracterizado por las probabilidades con que puede recibir un mensaje entre varios posibles, asumiéndose que éstas son conocidas de antemano, aunque no haya certeza sobre el mensaje concreto que le va a llegar.

Los objetivos de la teoría pueden resumirse en dos. Primero, investigar los límites teóricos fundamentales que afectan a la capacidad y eficiencia de transmisión cuando se utiliza un esquema de codificación de una clase determinada, examinando para ello hasta qué punto la presencia de ruido

restringe, la tasa de transmisión sin reducir la fiabilidad de la comunicación. Y segundo, diseñar códigos eficientes que permitan un comportamiento razonablemente bueno (en comparación con el máximo permitido).

Las investigaciones de Shannon se centraron exclusivamente en el primer objetivo. Los resultados que obtuvo se expresan como una serie de teoremas relativos a la transmisión eficiente y a la codificación óptima para diversos tipos de fuentes emisoras. Por ejemplo, uno de ellos establece que, bajo determinadas condiciones, si una fuente transmite mensajes a través de un canal con una capacidad de transmisión fijada, existe un código para el cual la frecuencia de errores puede hacerse tan pequeña como se desee, aunque no anularse.

Para poder llegar a dichos resultados se vio en la necesidad de introducir un nuevo concepto, llamémosle de momento «S», cuyo significado intentaremos aclarar. Es este punto precisamente el más importante de toda la teoría, y el que más malentendidos ha provocado. Veamos en primer lugar las razones de éstos.

Shannon plantea un problema concreto en el cual se considera una serie de sucesos posibles (por ejemplo, el conjunto de mensajes alternativos que pueden llegar a un receptor), cuyas probabilidades de ocurrencia son conocidas, siendo eso todo lo que se sabe. Seguidamente busca una función «S» que cuantifique el estado de conocimiento del receptor y que cumpla ciertas propiedades de consistencia lógica, y la obtiene aplicando un procedimiento axiomático, sin hacer referencia al contexto de la teoría de la comunicación. Una vez hallada, Shannon intenta atribuirle un nombre basándose en su interpretación dentro de la teoría de comunicación: «pensé en llamarle información, pero era una palabra demasiado utilizada, así que decidí denominarla incertidumbre. Cuando lo discutí con John von Neumann, él tuvo una idea mejor: Deberías llamarle entropía, por dos razones: primero porque tu función de incertidumbre ha sido utilizada en mecánica estadística con ese nombre, así que ya tiene un nombre; y en segundo lugar, y esto es más importante, porque nadie sabe lo que es la entropía realmente, de modo que en una discusión siempre tendrás ventaja».

La confusión estaba servida. La interpretación de la función como medida de información condujo a su rechazo inmediato en algunos campos, pues no tenía nada que ver con la idea intuitiva de información como contenido semántico. Por otro lado, el nombre de entropía hizo que se identificase sin ninguna justificación con la entropía termodinámica. Hoy en día la función de Shannon se denomina «entropía informativa» para distinguirla de la termodinámica, y se suele interpretar de diversas maneras. Es proporcional al número de sucesos alternativos posibles y alcanza su valor máximo cuando todos los sucesos son igualmente probables. Por ello se ha tomado, por ejemplo, como medida de sorpresa en la recepción de un mensaje, que es mayor cuantos más mensajes alternativos se consideren; de libertad en la fuente emisora a la hora de elegir mensajes; como medida de incertidumbre acerca de cuál de las alternativas potenciales es la que se puede enviar o recibir; o bien como medida de información. Las interpretaciones más correctas son las dos últimas, en donde «incertidumbre» e

«información» se pueden contemplar como emparejadas, siendo la entropía, en el contexto de la comunicación, una medida de la información que falta para eliminar o reducir la incertidumbre inicial con respecto a la emisión o recepción de un mensaje.

Junto con la función de entropía Shannon introdujo una unidad de información, denominada *bit*, acrónimo de *binary digit*, a propuesta de J. Tukey. Esta unidad expresa la información asociada a la situación de selección más elemental que cabe concebir, en la cual se consideran dos alternativas **equiprobables**.

Las ideas de Shannon sobre codificación fueron asimiladas rápidamente y aplicadas a aquellos procesos cuyas variables de interés eran entonces continuas, esto es, a las comunicaciones a través de radio, televisión y teléfono. Sin embargo, sus resultados para sistemas con variables discretas fueron demasiado prematuros para la tecnología de la época. Hasta la llegada de los circuitos integrados de alta velocidad en los años setenta no fueron explotados por los ingenieros. Actualmente son la base de los algoritmos que soportan las telecomunicaciones digitales así como de los sistemas de compresión de datos y de los protocolos de eliminación de errores en la transmisión de información por la red, pero en un principio fueron utilizados casi únicamente en lingüística y criptología.

El hecho de reconocer que todo lenguaje posee una estructura probabilista, que determina el encadenamiento de símbolos en los mensajes comunicados, condujo al estudio de la redundancia existente en los distintos idiomas y en su papel en la estrategias para superar el ruido. Se descubrió, por ejemplo, que un hablante inglés tiene una libertad de elección a la hora de producir mensajes de aproximadamente sólo el 50 por ciento, estando restringidas la mitad de sus opciones por los esquemas probabilísticos, los cuales pueden incluir aspectos gramaticales, propiedades de eufonía, etc. Como ilustración de lo anterior, piénsese en la expresión «*redundanci*»; las probabilidades de que, una vez transmitida, la siguiente letra en la cadena sea una «a» son enormemente altas en castellano, con lo que la inventiva del emisor y las expectativas del receptor se ven muy reducidas. Que todo esto ocurra permite que, aunque el ruido afecte a parte de los mensajes, la información llegue finalmente al receptor. El diseño de discos compactos y otros registros similares está basado justamente en este principio de la consideración de esquemas de codificación artificiales para aumentar la redundancia de los mensajes originales que se quieren comunicar.

Por otra parte, el uso inverso de la codificación, buscando que el ruido sea mayor mediante la introducción de una clave de tipo estadístico que mezcla el mensaje original, ha sido la base de muchos sistemas criptográficos. El propio Shannon desarrolló este tipo de aplicaciones y fue nombrado por ello consultor en materia de criptografía del gobierno de Estados Unidos, siendo precisamente su sistema el que emplearon Churchill y Roosevelt para mantener conferencias transoceánicas durante la Segunda Guerra Mundial.

CRÍTICAS

La teoría de Shannon y Weaver ha recibido numerosas críticas, sobre todo cuando ha intentado salirse de su ámbito técnico de aplicación. Éstas han partido especialmente de los lógicos, los lingüistas y los expertos en comunicación animal, social y de masas, quienes han considerado muy certeramente que es insuficiente como modelo general para la comunicación y que tampoco es válida como una teoría de la información. Se ataca fundamentalmente, por un lado, que ignora la posibilidad de distintos contextos (situacionales, sociales, históricos, etc); y, por otro, que su concepción de la información no guarda relación con las nociones intuitivas de información entendida como dato, contenido sensible o contenido semántico, lo que es aparentemente paradójico, puesto que uno de los fines de la comunicación es precisamente la producción de significados.

Además, se señala que el modelo propuesto es del tipo denominado «modelo de transmisión» y está basado en la metáfora de la comunicación como transporte, más concretamente como transporte postal, considerando por tanto que los mensajes preexisten como paquetes con un significado asociado, son emitidos secuencialmente y llegan al receptor, quien no es más que un elemento puramente pasivo en el sistema. Este modelo presupone asimismo que las probabilidades son conocidas de antemano, cosa que en la práctica no sucede; y que los códigos son previamente acordados por emisor y receptor, algo imposible sin una comunicación anterior de la cual en ningún caso se habla. Como consecuencia, el aprendizaje no costaría esfuerzo, pues consistiría en una simple descodificación y recepción de mensajes. Sin embargo, como muestran los resultados experimentales sobre comunicación social, es más verosímil que sea el receptor quien construye el significado de un mensaje, y quien cree el mensaje situándolo en un contexto. Concluyen, así, que no debe contemplarse la teoría como «la ciencia que subyace bajo todo intercambio de información» o «la que incluye todos los procesos por los que una mente puede afectar a otra», tal y como sostenían algunos de sus defensores.

No obstante, con respecto a estas cuestiones hay que indicar que fue el mismo Weaver el primer investigador en distinguir en todo proceso de comunicación varios niveles: el sintáctico, asociado al problema técnico de la transmisión; el semántico, y el pragmático, correspondiente al estudio de su intencionalidad. Era, pues, consciente de las limitaciones de la teoría.

En relación con esto, él y Shannon manifestaron repetidamente que su modelo pertenece al nivel técnico o sintáctico y que los aspectos semánticos son irrelevantes para el diseño de los sistemas de comunicación, ya que interesa que éstos sean eficientes independientemente del tipo de mensaje comunicable o de su valor. Aclararon también que el concepto de información en su teoría es formal y tiene un sentido especial distinto al de significado o contenido de los símbolos comunicados. A diferencia de la información semántica, la información de Shannon no se asocia a mensajes individuales, sino que caracteriza globalmente el conjunto de mensajes alternativos posibles. En el

acto de transmisión de un mensaje la información está ligada a todos los mensajes que podían haber sido enviados. Por tanto, se refiere más a «lo que podría decirse» que a «lo que se dice», es decir, a los mensajes en tanto que potencialmente enviabiles, y no a un mensaje enviado concreto. Aceptando que la teoría no es totalmente general, lo que sí defendieron sus autores es que los teoremas obtenidos establecen restricciones sobre los otros niveles, semántico y pragmático.

A pesar de las críticas hay unanimidad en atribuir a Shannon un papel relevante dentro del campo de la comunicación y como precursor de muchos de los desarrollos tecnológicos actuales. Dentro del marco de análisis usual se le atribuyen valiosas contribuciones, algunas de las cuales ya hemos señalado. Sin embargo, al ampliar dicho marco se vislumbra una imagen distinta y se descubre que el potencial asociado a sus ideas mucho mayor que el reconocido hasta ahora.

UN CONCEPTO CLAVE: LA ENTROPÍA INFORMATIVA

Si se analiza su libro de 1949, se observa que hay referencias continuadas a la comunicación en prácticamente todas las secciones. Pero es muy interesante descubrir que hay una en la que no ocurre esto. Se trata del capítulo dedicado a la introducción de la entropía informativa. En él Shannon se sale de dicho contexto y es mucho más general y abstracto. Presenta la entropía como una propiedad asociada a probabilidades. De este modo la entropía pasa a ser un concepto perteneciente a la teoría de la probabilidad, no a la de comunicación. Se puede utilizar por tanto para caracterizar o cuantificar toda situación de incertidumbre. Así pues, cabe concluir que la importancia fundamental de Shannon consiste en haber dado con un concepto esencial para la inferencia científica que faltaba en todas las teorías epistemológicas anteriores. El significado del nuevo concepto se vuelve más claro: la entropía informativa mide la incertidumbre global en todo problema de conocimiento, por ejemplo, en aquella situación en que la evidencia disponible no permita conocer la verdad o falsedad de proposiciones sobre observaciones potenciales de un fenómeno cualquiera.

En este punto hay que resaltar la figura de Edward T. Jaynes, otro pensador de la talla de Shannon, fallecido también recientemente (1998). Éste adoptó la entropía informativa como medida de incertidumbre y dio la vuelta a las ideas del primero. En lugar de suponer, como se hacía en teoría de la comunicación, que las probabilidades asociadas a los distintos mensajes son conocidas, Jaynes utilizó la conexión establecida por Shannon y presentó, en 1957, un método basado en la entropía informativa justamente para buscar y asignar las probabilidades en una situación de incertidumbre.

Dicho método es conocido como el Método de Máxima Entropía y constituye un procedimiento de inferencia científica. Una vez formalizado, Jaynes lo aplicó a la termodinámica y a la teoría de la comunicación y, sorprendentemente, consiguió derivar de manera muy sencilla todas las expresiones teóricas que se habían obtenido en dichas áreas presuponiendo todo tipo de hipótesis más o menos

restrictivas. Estableció además, por vez primera, una conexión precisa entre la entropía informativa y la termodinámica. Desde entonces, muchas de las relaciones que anteriormente se entendían como principios o leyes de la naturaleza hay que contemplarlas como manifestaciones de reglas del razonamiento probabilista.

Se comprende ahora por qué hemos evitado hablar de «teoría de información» para referirnos a la modelización de Shannon. Este término debe ser reservado para algo más general que una teoría de la comunicación; debería asociarse a la lógica del razonamiento, es decir, a una teoría del conocimiento. Desde este nuevo punto de vista, las tecnologías actuales, llamadas de la información, no constituyen más que desarrollos derivados de las ideas concebidas en los años 40 y 50. Ha aumentado la conectividad, los ordenadores son más veloces y las redes de comunicación más complejas. Sin embargo, su funcionamiento es esencialmente el que se describe en trabajos de esa época. Los avances técnicos han sido debidos a la explotación industrial, pero a nivel teórico no ha habido cambios. Éstos están por venir. En este sentido, lo principal en Shannon es que, a través de Jaynes, abre las puertas hacia un pensamiento científico unificador nuevo.

La verdadera revolución de la información todavía no ha llegado, debiendo considerarse la situación presente más bien como un preludio. La próxima revolución no sólo tendrá aplicaciones tecnológicas, sino que conllevará una transformación radical en la manera de concebir el mundo. A este respecto hay que citar al físico John A. Wheeler, quien en su frase «*it from bit*» ha simbolizado la idea de que la realidad se construye desde actos elementales de participación de un observador que formula preguntas cuyas respuestas proporcionan información, enlazando de este modo con las teorías sociales constructivistas. El objetivo es, pues, la formalización de la inferencia científica como teoría del conocimiento. La contribución de Shannon en esta dirección ha sido precisamente la de proporcionar herramientas conceptuales imprescindibles que anteriormente no existían.

Fecha de creación

29/09/2001

Autor

Alberto Solana Ortega