

# TAREA 1. INVESTIGACIÓN

## ¿Quién fue Claude Shannon?

Claude Elwood Shannon fue un matemático y criptógrafo americano quien tiene fama de ser "El padre de la teoría de la información" esto sobre todo por su teoría matemática de la comunicación de 1948, además de que también es famoso por ser uno de los primeros en sentar los fundamentos teóricos en circuitos digitales (Esto a los 21 años en 1937)[1].

De manera general podemos atribuir que Shannon fue uno de los primeros en relacionar el campo del álgebra booleana con la relación lógica y sobre todo numérica de lo que sería el futuro bit dentro de las computadoras.

## ¿Qué es la teoría de la información?

La teoría de la información surgió a finales de la Segunda Guerra Mundial a través de un artículo publicado en el Bell System Technical Journal por Claude E. Shannon, en este artículo, se menciona acerca de una teoría matemática de la comunicación, esto sobre todo porque al apenas tener menos de medio siglo con las telecomunicaciones se necesitaban asentar principios y fundamentos a las nuevas tecnologías e innovaciones traídas por la guerra.[1]

Con base a esto, sabemos que de manera general esta teoría está basada con aspectos matemáticos relacionados con la medición de información además de la representación de esta misma dentro de sistemas de comunicación.

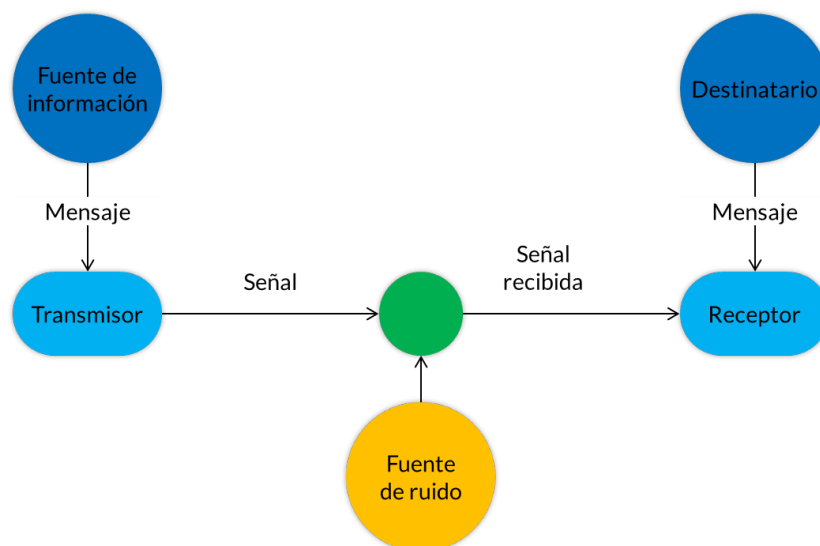


Imagen 1. Diagrama general de la comunicación entre un transmisor y receptor dentro de un sistema de comunicaciones establecido por Shannon y Weaver.

De esta forma es como esta teoría fue una primera premisa a aspectos relevantes como la forma en que se transmitirá la información además de aspectos mucho más futuros como la forma en la que se modificaría e incluso almacenaría.

### ¿Qué es la entropía?

El concepto de “Entropía” fue usado por primera vez por Rudolf Clausius en 1851 cuando estaba formulando la segunda ley de la termodinámica, que si bien en ese contexto hace referencia de que es una magnitud que mide la parte de la energía que no puede utilizarse para producir un trabajo, es decir “una forma de medir el desorden de un sistema”, la realidad es que en otros contextos, puede variar como “Es la medida de incertidumbre promedio, la cual se calcula a partir de la probabilidad de ocurrencia de cada uno de los eventos”

### ¿Qué es la entropía de Shannon?

Para entender la entropía de Shannon podemos asimilar de la siguiente forma:  
Podemos definir la autoinformación de un evento  $X = x$  como

$$I(x) = -\log P(x)$$

Por tanto, nuestra definición de  $I(x)$  está escrita en unidades de nat. Un nat es la cantidad de información obtenida al observar un evento de probabilidad  $1/e$ .

Es decir, que nosotros podemos cuantificar la cantidad de incertidumbre en una distribución de probabilidad completa usando la **entropía de Shannon**.

$$H(x) = \mathbb{E}_{x \sim P} [I(x)] = -\mathbb{E}_{x \sim P} [\log P(x)]$$

Además sabemos que la entropía de Shannon de una distribución es la cantidad esperada de información en un evento extraída de esa distribución, dando como límite inferior en el número de bits (si el logaritmo es base 2, de lo contrario las unidades son diferentes) necesarios en promedio para codificar símbolos extraídos de una distribución.[5]

### ¿Por qué es útil esa entropía para clasificar un documento observando sus caracteres?

Porque es una forma en la que podemos analizar la complejidad y la probabilidad de desorden con la que tendremos que enfrentarnos al analizar un texto en un determinado idioma, por ejemplo y como se menciona Ravin Charan[6] en su escrito, Shannon en sus primeros análisis de la entropía del idioma inglés notó que podría estar entre 0.6 y 1.3 bits por carácter, sin embargo, con herramientas como GPT-2 de OpenAI sabemos que es aproximadamente de 1 bit por carácter, y esto en ¿Qué se traduce ?.

En que mientras un sistema tenga una mayor entropía resultará mucho más complejo de analizarlo o estimarlo debido a esa incertidumbre que tiene.

Particularmente el uso de la entropía para clasificar documentos mediante caracteres hace o considera el grado de entropía del idioma a través de estimaciones como las previas mencionadas, y mediante esto es como se puede abordar de una mejor forma los problemas a resolver e incluso estimar el tiempo que tardaría en resolverse.

## Referencias:

- [1] "Full Page Reload", *IEEE Spectrum: Technology, Engineering, and Science News*, 2021. [Online]. Available:  
<https://spectrum.ieee.org/tech-history/cyberspace/claude-shannon-tinkerer-prankster-and-father-of-information-theory>. [Accessed: 15- Mar- 2021].
- [2] "Introduction to entropy", *Khan Academy*, 2021. [Online]. Available:  
<https://www.khanacademy.org/science/ap-chemistry/thermodynamics-ap/entropy-tutorial-ap/v/introduction-to-entropy>. [Accessed: 15- Mar- 2021].
- [3] *Catarina.udlap.mx*, 2021. [Online]. Available:  
[http://catarina.udlap.mx/u\\_dl\\_a/tales/documentos/mosl/rojas\\_b\\_js/capitulo4.pdf](http://catarina.udlap.mx/u_dl_a/tales/documentos/mosl/rojas_b_js/capitulo4.pdf). [Accessed: 15- Mar- 2021].
- [4] *Deeplearningbook.org*, 2021. [Online]. Available:  
<https://www.deeplearningbook.org/contents/prob.html> [Accessed: 15- Mar- 2021].
- [5] "The intuition behind Shannon's Entropy", *Towards Data Science*, 2021. [Online]. Available:  
<https://towardsdatascience.com/the-intuition-behind-shannons-entropy-e74820fe9800>. [Accessed: 15- Mar- 2021].
- [6] "The relationship between perplexity and entropy in NLP", *Towards Data Science*, 2021. [Online]. Available:  
<https://towardsdatascience.com/the-relationship-between-perplexity-and-entropy-in-nlp-f81888775ccc>. [Accessed: 15- Mar- 2021].