



COMPRENDIENDO EL LENGUAJE DESDE LA TERMODINÁMICA Y LOS FRACTALES

Dra. Daniela López De Luise

Introducción

Este escrito es un resumen que corresponde a la presentación realizada por la autora en la disertación del día 24 de agosto de 2021 a las 18 h hora local, para la comunidad asistente al evento organizado por la Academia Nacional de Ciencias de Buenos Aires, en su ciclo de presentaciones CETI (Centro de Estudios en Tecnologías Inteligentes).

1. El lenguaje natural

El lenguaje humano es una compleja actividad del cerebro que le distingue de otras especies por su extraordinaria riqueza y flexibilidad. Es importante desde la infancia puesto que se ha probado que influye significativamente en el desarrollo cognitivo y en la maduración emocional y social del individuo. Varios autores han detectado que se rige por leyes precisas y muchos ya han avanzado en su estudio desde las matemáticas y la entropía. Parte de estos trabajos son mencionados en este escrito. Constituyen algunos pilares en los que se asienta la teoría introducida al final. La misma, a diferencia de otras contribuciones en el campo, no intenta explicar el lenguaje producido por los seres humanos, sino que se enfoca en las leyes que gobiernan la esencia de la producción de sentencias en contextos de diálogos.

2. Indicios matemáticos en el lenguaje

Desde hace años ya que los lingüistas y especialistas de la llamada Computación Lingüística trabajan en el análisis de las producciones del lenguaje. Si bien existen varias corrientes, una de ellas especialmente productiva sostiene que sigue reglas llamadas Gramática, algo que se suele representar con la tupla $G(\Sigma, N, R, S)$. En esta, Σ representa un conjunto de signos que reconocemos como vocabulario, las palabras que uno encontraría en el diccionario, por ejemplo. La N son signos denominados “no terminales”, R las reglas de formación de las sentencias y frases, y S se reserva como un signo de inicio de la generación de una sentencia. A partir de allí se puede hablar de una gramática generativa que justifica la formación de todas las frases posibles de un hablante probada la inclusión de todos los elementos necesarios para ello.

Tal vez resulte poco impactante pensar que una cuestión tan compleja como el habla pueda explicar sus producciones con algo que es relativamente tan simple como una tupla de elementos. En cierta manera es verdad que no explica completamente los aspectos semánticos (significado y esencia informativa), pero sí sus aspectos meramente formales. Y esto da cuenta de algún tipo de mecánica, una sistemática de formación a la que se agrega una esencia de significados. Esta dualidad también es objeto de estudio, y es lo que induce a pensar en que el lenguaje es finalmente algo impreciso e inmanejable.

Pero habiendo dado cuenta de esta dualidad, en esta presentación se avanza en la parte mecánica por un sendero que asocia hallazgos diversos, que podrían explicar en qué momentos la formalidad se enreda con esencia semántica.



3. Antecedentes y curiosidades

Dado que el habla acompaña a la humanidad desde hace mucho tiempo, su estudio no es nuevo y las aportaciones en la materia es variada, rica y extensa como cualquiera podría imaginar. De los estudiosos destaca el lingüista Ferdinand de Saussure, quien en 1998 afirma que el lenguaje es un a-priori social internalizado antes del desarrollo del razonamiento, una entidad rígida y de origen externo al individuo. Si bien se han discutido estas afirmaciones, es innegable el impacto que tuvo en la comunidad. Al día de hoy las aportaciones han evolucionado y entre otras nos interesa el feliz arribo de especialistas como Sala Torrent, que recientemente se han animado a afirmar que el lenguaje es un sistema y funciona bajo las leyes del caos (Sala Torrent, 2020). Es que el lenguaje muestra caos y fue el mismo Saussure que en sus notas póstumas afirma haber hallado unos anagramas en el latín que no se explican desde los conceptos tradicionales. No fue sino hasta mucho después que estos extraños comportamientos se relacionan con las leyes de la entropía y del caos. Hoy se pueden hallar trabajos como “Lingüística del caos” de Hakim Bey (1985) y “La teoría del caos y el lenguaje” (Kartchner, 2000).

Como otras actividades humanas el lenguaje sigue leyes matemáticas (entre ellas comportamientos fractales) y parece ser que existe una estrategia innata del cerebro del tipo logarítmico. La distribución fractal palabras (conocida como distribución Zipf-Mandelbrot) es una más de estas manifestaciones. El biólogo Ludwig von Bertalanffy, Claude Shannon, Karen Wynn, Rodríguez Santos, son algunos nombres que trabajaron en otras manifestaciones por el estilo.

4. La computadora cerebral

Si algo tiene de maravilloso el mundo fractal, es que se centra en definir invariancias persistentes a cambios de escala. Lo que lleva en términos de geometría a una concepción extraordinaria denominada Geometría fractal, que logra describir mucho mejor las figuras y comportamientos de la naturaleza, dejando a la Geometría euclidiana para describir artefactos artificiales creados por humanos. No es de extrañar que el lenguaje, también pueda ser bien descrito con los fractales. Pero la clave de todo yace en qué medir como fractales dentro del proceso de comunicación con el habla natural.

Una forma de rastrear la esencia fractal consiste en evaluar la producción de sentencias en contextos de diálogos. Y si se puede acotar a un sistema automático, es posible acotar aún más el análisis y concentrarse en los aspectos esenciales de construcción de sentencias como vehículo de comunicación.

Por ello el juego 20Q, que está en 20q.net, es un interesante objeto de estudio. Detrás de la red hay un sistema inteligente que aprende con una red neuronal. Aprende a adivinar palabras que el usuario piensa, por medio de una serie de preguntas (idealmente 20 preguntas). Esto es como decir que un conjunto de neuronas artificiales define un sistema de ecuaciones polinómicas N-dimensionales que representa la manera en que se asocian las palabras con ciertas características.



Luego de un estudio técnico apropiado, es posible ver que el sistema logra adivinar la palabra. Estudiar la interacción con los usuarios permite observar la ley detrás del caos: de las 93000 palabras del diccionario de la lengua española (RAE, 2014) debe adivinar una

y solo una. A veces son términos compuestos como “lancha a motor”) que complica aún más la adivinación. Pero lo logra en una sorprendente cantidad de casos, y en otros ideas u objetos de tipo muy cercano. ¿Cómo es posible?

No se trata de magia, sino de la puesta en escena de reglas muy precisas que todos manejamos: Reglas termodinámicas de la lengua.

Dos de ellas que aplican de inmediato son:

- Regla 1: *Regla del comportamiento completo*. Se puede observar en la manera en la curva de probabilidades para verbos y sustantivos dentro de las sucesivas sentencias a lo largo del juego y en la manera diferenciada en que evolucionan. También en la posibilidad de encontrar tipos de sentencias muy específicos en momentos clave del diálogo con la máquina.

- Regla 2: *Dimensión y Ritmo*. Es un poco más sutil que la anterior porque requiere que se estudie la dimensión fractal de la distribución de entropía (normalmente reconocida por la letra H) de las sentencias a través de sus palabras. Entre otras cosas se puede observar esta regla en la existencia de un comportamiento fractal en diálogos (López De Luise & Azor, 2015). Pero también existe un orden y ciclo en lo que sucede con H. Se observan también anomalías de flujo de H respecto a la presencia de verbos y sustantivos.

Siguiendo el rastro del ritmo de entropía de manera apropiada se descubren ciclos organizados con cierto ritmo, donde la información de una comunicación C permite estudiar perfectamente cómo fluye la información entre las partes de una comunicación. También permite explicar las estrategias empleadas (por ejemplo, para ganar el juego) y su grado de éxito.

Resumen

La información intercambiada en un diálogo se estructura según cierta dinámica precisa de entropía. Dichas reglas permiten que, a través del estudio de la entropía del lenguaje, se determine si existe un comportamiento completo en el intercambio de información y cierta dimensión y ritmo en el intercambio.

En estas reglas son esenciales consideraciones como la dimensión fractal para explicar la diferencia de precisión entre la información intercambiada y la entropía medida, discrepancia que ahora sí puede ser descrita con mayor precisión por estas y otras reglas adicionales.