



Lectura crítica: “How Language Programs the Mind”

Gary Lupyan and Benjamin Bergen, 2014

Dentro del ámbito de la Inteligencia Artificial, lecturas teóricas sobre ciencia cognitiva y aprendizaje humano son claves de incorporar dentro de nuestro bagaje disciplinar, ya que dan muchas luces de cómo se podrían resolver problemas actuales y persistentes sobre la simulación del razonamiento y aprendizaje. Me parece que el interés principal del texto de Lupyan y Bergen es la basta demostración de teorías que relacionan el lenguaje con nuestra capacidad humana de categorizar al mundo. Ante todo, este mismo concepto de etiquetas que tan fuertemente persiste en las técnicas de aprendizaje supervisado en aprendizaje profundo.

Sin embargo, la tesis de los autores sobre *language-based programmability* no logra desarrollarse de forma específica y requiere mayor clarificación. No queda claro si se ha llegado a entender cómo en la mente se van registrando las relaciones entre palabras, que bien insisten vienen en primer lugar de la unidad de las palabras alimentadas por contexto. Así mismo, al referirse a activación, no hacen referencia al fenómeno químico o biológico de la activación en el cerebro, sino más bien lo da por entendido como un proceso cognitivo. Más aún, se refieren de forma muy amplia al lenguaje como un sistema de control para programar la mente humana, y por alguna razón no fundamentada descartan la existencia de otra especie en nuestro universo conocido que tenga la misma capacidad.

En este sentido la postura de los autores viene de conclusiones *in extrema res* derivadas de una sociedad sistemática e institucionalizada ya construida en torno al lenguaje, cuando lo adecuado sería una aproximación teórica *ab ovo*. Creo que la referencia más importante de todo el argumento, y que podría construir el cuerpo de la hipótesis, se encuentra en la sección de las conclusiones cuando refiere al

estudio de Clark de 1998 sobre el rol computacional del lenguaje [1].

Fiona Macpherson hace una observación similar [2], indicando que el cerebro produce modelos generativos *top-down* o representaciones que son predicciones de cómo es el mundo, pero que no siempre la codificación predictiva mediante el lenguaje implica penetración cognitiva.

Una conexión interesante sin duda es la postura respecto al aprendizaje de categorías y cómo la biología humana y el ambiente sub determinan los conceptos posibles que puede adquirir el sujeto humano. Me hace pensar si esa misma limitación acaso no es evidente también en las técnicas supervisadas de aprendizaje profundo.

En este mismo sentido y conectando con el aprendizaje profundo, la idea introducida por Edmiston & Lupyan sobre la activación de representaciones concretas o abstractas dependiendo de la naturaleza de la etiqueta [3], hace pensar si acaso se podría agregar una capa que a su vez clasifique las etiquetas según su naturaleza. En este sentido ciertas etiquetas pueden conformar un entendimiento abstracto del concepto “lluvia”, mientras otras un entendimiento específico sobre ciertas instancias específicas de la lluvia. El desafío es sin dudas multimodal.

[1] Clark, A. (1998). Magic words: How language augments human computation. In P. Carruthers & J. Boucher (Eds.), *Language and thought: Interdisciplinary themes* (pp. 162–183). New York: Cambridge University Press.

[2] Macpherson, F. (2015) Cognitive Penetration and Predictive Coding: A Commentary on Lupyan. *Review of Philosophy and Psychology*, 6, 571–584.

[3] Edmiston, P., & Lupyan, G. (2015). What makes words special? Words as unmotivated cues. *Cognition*, 143, 93–100.