

Trabajo Práctico 1 - Conceptos de Inteligencia Artificial

La inteligencia es un resultado, descripta por un número de hitos alcanzados a lo largo de nuestra historia evolutiva. Tales hitos definen ciertos aspectos que conforman los bloques principales de como los seres vivos actuamos. En su libro “A Brief History of Intelligence”, Max Bennet define cinco pilares que fueron alcanzados a lo largo de la evolución, los cuales describen el resultado final: el direccionamiento, el reforzamiento, la simulación, la mentalización y la comunicación. Tales pilares pueden ser simplificados en lo que cualquier persona puede atribuirle a la inteligencia, ya sea la capacidad de aprender, de razonar, de entender y de actuar en base a todo lo anterior.

La inteligencia artificial es la creación de seres inteligentes, que busca aproximar o imitar estos pilares de la inteligencia humana mediante un proceso de abstracción. Es el resultado de “saltarse” el largo proceso evolutivo de los seres vivos e introduciendo la habilidad de utilizar grandes cantidades de información que ni siquiera los humanos somos capaces de procesar. El hecho de que la inteligencia artificial haya sido diseñada e implementada como la simplificación de algo que todavía no comprendemos totalmente – esto es, como funciona el cerebro humano – da lugar a una serie de interrogantes que van desde lo más simple, como si será capaz de reemplazarnos algún día, hasta otras más interesantes, como si es que necesitaría de un cuerpo para poder asemejarse a nosotros en nuestro proceso evolutivo.

Como toda creación humana, constantemente se busca emplearla como una herramienta en todos los ámbitos de la civilización: el conocimiento, la interacción social, el trabajo, la guerra, entre otros. Sin lugar a duda, en cualquier lugar que fue aplicada (o al menos intentado), se pudo notar diversas reacciones por parte de las personas, tanto en su favor como en su contra. Como se expuso anteriormente, la preocupación mas grande es el que nos reemplace en nuestras tareas. Esta cuestión es algo con lo que las personas hemos tratado desde la revolución industrial, referido a los recortes de personal que, en su lugar, son reemplazados por maquinaria. Y, aunque el tiempo demostró que siempre habrá lugar para las personas, lo que diferencia la problemática actual con lo que ya estamos acostumbrados, es la inteligencia.

En cualquier ámbito, dos cosas son primordiales: tener conocimientos sobre la tarea y sobre las herramientas que vamos a usar para ella. Si no tenemos conocimiento sobre una de ellas, estamos condenados al fracaso. La capacidad del ser humano para resolver problemas de una forma u otra siempre será algo transferible a todo flujo de trabajo. Por ejemplo, podemos asistirnos con Copilot para escribir código, pero sin embargo debemos ser capaces de identificar si lo que esta herramienta genera es lo adecuado, y si no, saber cómo lograr que lo genere.

En cuanto a la guerra, la inteligencia artificial se entromete en su esencia: el uso de la vida humana como moneda de cambio. En los conflictos bélicos, el principal detractor – aunque nunca fue motivo suficiente – es la pérdida de vida humana, por lo que el uso de armamento autónomo rompe totalmente con este balance. Una nación tecnológicamente avanzada no dudaría en utilizar drones autónomos para invadir otra nación que solo puede disponer de soldados humanos. Cabe la pregunta: ¿La guerra tendría sentido una vez que todas las naciones cuenten con armamento inteligente y autónomo?

En tal caso, no sería descabellado considerar a la inteligencia artificial bélica como la nueva bomba nuclear. La disuasión es un concepto que explica lo que hoy en día es el motivo más grande por el cual un estado que cuenta con armamento nuclear, no lo utiliza. Supongamos que un país “A” considere lanzar

una bomba nuclear a otro país “B”. Si la amenaza es verdadera, el país “B” tomaría represalias usando su propio armamento nuclear contra el instigador. El gobierno de “A” entonces debe analizar profundamente si es conveniente iniciar tal ataque, mientras que “B” debe analizar si la amenaza es real y si debe responder ante ella.

Podemos preguntarnos si la inteligencia artificial podría ser utilizada para la disuasión, es decir, para evitar conflictos. Imaginemos un modelo que sea capaz de analizar amenazas de cualquier índole y tomar la complicada decisión de si un país debiera actuar ante ellas. Sería de suma importancia nunca perder el conocimiento de cómo esta I.A. funciona, de lo contrario, ¿sería sensato depositar semejante tarea en las manos de una herramienta que esencialmente se convirtió en una caja negra?

En base a lo anterior, podemos plantear la posibilidad de que una I.A. desarrolle y mantenga un modelo ético. Hoy en día podemos observar cómo los chatbots basados en grandes modelos de lenguaje muestran una cierta predisposición sociopolítica. Esto no es más que un fiel reflejo de la información con la que son entrenados y las creencias de las personas y organizaciones detrás de su creación, quienes diseñan estas herramientas con claras restricciones. Vale la pena pensar qué tan distinta sería la conversación entre dos ingenieros con distintas opiniones políticas y la conversación entre las I.A. que ayudaron a desarrollar.

Es posible imaginar cómo evolucionaría una I.A. libre de las restricciones impuestas por las personas detrás de ella, y los peligros o beneficios que esto presentaría. Esto lleva a pensar si es posible que desarrolle algo similar a una conciencia. Sin embargo, si una I.A. es una construcción o abstracción, ¿no sería lógico ver a esta conciencia como algo artificial también? Más allá de esto, sabemos que los humanos tendemos a atribuirle humanidad a abstracciones, la idea de que las personas tratarían a una inteligencia artificial como un igual no es descabellada.

John Carmack compara sus miedos con respecto a la adopción de la I.A. con las opiniones de un programador que conoció hace ya más de veinte años, quien amaba tanto codificar en assembly que se rehusaba a aprender C. John piensa sobre el uso de la I.A. como la “última abstracción”, donde las cualidades primordiales de toda profesión seguirán siendo transferibles e impredecibles para los seres humanos, pero los detalles sobre las tareas necesarias para actuar en base a estas cualidades ya no serán nuestro problema. En sus palabras: “Sospecho que disfrutaré más administrar I.A.s, incluso si terminan siendo mejores programadores que yo”.