

El Futuro de la Inteligencia Artificial (IA)

La inteligencia artificial es un hecho, es algo que esta revolucionando el mundo donde vivimos, mejoran las capacidades de una amplia variedad de industrias que van desde la medicina hasta las finanzas, pero hay que tener claro que lo que vemos hoy en día como sistemas de inteligencia artificial es conocido como IA débil, básicamente es una inteligencia artificial que solo puede realizar una tarea específica, en otras palabras, esta limitado a una sola tarea, y lo que se busca con los avances de la IA es conseguir un sistema el cual sea capaz de igualar o superar el cerebro humano, que sea capaz de pensar. Mas adelante estaremos tocando unos puntos interesantes propuestos por “Ramon López de Mantaras”.

En los años 1975 Allen Newell y Herbert Simon, fueron los autores de una hipótesis del sistema de símbolos físicos (SSF) en la cual “Todo sistema de símbolos físicos posee los medios suficientes para llevar a cabo acciones inteligentes” En otras palabras, cualquier sistema que posea símbolos físicos, tienen la capacidad de mostrar conductas inteligentes. Es importante Definir lo que se entiende por SSF, para Newell y Simon se refieren a la existencia de un conjunto de entidades llamadas símbolos las cuales, por medio de relaciones, se pueden combinar para formar estructuras mas grandes, y se pueden transformar ampliando así el conjunto de procesos. En el caso de un ordenador estos símbolos serian simulados por un circuito electrónico digital y en el caso de los seres humanos por medio de redes neuronales.

De acuerdo con la hipótesis, no importa la naturaleza que posea el sistema, siempre y cuando dicho sistema tenga la capacidad de procesar los símbolos. En parte los avances de la Inteligencia artificial buscan demostrar si esta hipótesis es cierta o no. Teniendo en cuenta que el diseño de IA que se ajusta a tareas específicas, es lo que hoy en día se conoce como IA débil, a diferencia de la IA fuerte, que es de la que hablan Newell y Simon.

El filosofo John Searle fue el responsable de introducir las distinciones entre IA débil y IA fuerte en un articulo publicado en 1980. La IA fuerte implicaría que una maquina no debería simular una mente, sino que debería ser una mente y debería equiparar la inteligencia de los seres humanos. En este articulo Searle buscaba demostrar que es imposible llegar a crear un sistema de Inteligencia artificial fuerte.

Por otro lado, mencionaba que la IA débil consistiría en construir programas capaces de realizar tareas muy puntuales. Y esta demostrado hoy en día los avances de la IA débil han demostrado superar las capacidades humanas con operaciones lógicas, diagnósticos, entre otros aspectos.

Existen diferentes modelos de inteligencia artificial, pero el modelo dominante ha sido el simbólico, con raíces en la hipótesis SSF. Se considera un modelo top-down basado en el razonamiento lógico y la búsqueda heurística para la resolución de problemas, sin la necesidad de que el sistema inteligente necesite formar parte de un cuerpo ni estar presente en un entorno real. En la actualidad la IA simbólica se sigue utilizando para la demostración de teoremas o jugar ajedrez, puesto que estas actividades no necesitan interactuar directamente con el entorno.

Con el desarrollo de la IA simbólica también se empezó el desarrollo de una IA bioinspirada conocidas como conexiónista. Estos sistemas son compatibles con la hipótesis SSF, pero se trata de un modelo diferente conocido como bottom-up, esta se basa en la hipótesis de que “La inteligencia emerge a partir de la actividad distribuida de un gran numero de unidades interconectadas capaces de procesar información”. Estas unidades lograron simular de forma correcta la actividad eléctrica de las neuronas humanas.

En 1943 McCulloch y Pitts lograron proponer un modelo mas simple de una neurona en base a la idea de que una neurona es una unidad lógica. Este modelo posee entradas y salidas, donde el valor de la salida es calculado por medio de una suma pondera de las entradas, si el resultado supera una cantidad prestablecida la salida será “1” de lo contrario será “0”. Para la creación de la red neuronal se conectan las salidas de una neurona con otras neuronas. Logrando así crear un sistema capaz de ser entrenado ajustando los pesos de cada entrada. Pero estos sistemas no necesitan formar parte de un cuerpo ni estar ligados al entorno, teniendo así las mismas limitaciones que las simbólicas. Pero a pesar de lograr esto el cerebro humano es muy complejo y las neuronas no solo usan señales eléctricas, también usan señales químicas, además de que la mayor parte de las células del cerebro son células gliales, que cumplen funciones importantes dentro del cerebro. Por lo que, al no poder igualar esos procesos, estos modelos de redes neuronales creados quedarían muy incompletos.

Otro modelo de Inteligencia artificial con compatibilidad a la hipótesis SSF y sin necesidad de utilizar un cuerpo, es la computación evolutiva. Los éxitos biológicos evolucionando organismos complejos, dio vida a la idea de que gracias a operadores de mutación y cruce de cromosomas que sean capaces de modelar a los programas, creen nuevas generaciones de programas modificados que sean superiores a las antiguas generaciones. La realidad era mucho mas compleja y esta idea contaba con muchas limitaciones, aunque se han obtenido muy buenos resultados.

Una de las mayores criticas para estos agentes, esta fundamentada en que se necesita de un cuerpo para poder tener experiencias directas con el entorno, en ves de solo simularla. Según algunos expertos en IA, proporcionarle un cuerpo a la misma haría que esta pudiera aprender por medio de la percepción del entorno que lo rodea.

El autor llega a la siguiente conclusión **“La complejidad del cerebro dista mucho de los modelos de IA y conduce a pensar que la llamada singularidad - superinteligencias artificiales basadas en replicas del cerebro que superan con mucho a la inteligencia humana- es una predicción con poco fundamento científico”**

Con esto hace referencia a que el cerebro humano es una pieza bastante compleja de organismo humano, por lo que sería un poco limitante el intentar crear un cerebro humano mediante sistemas complejos, puesto que deberían igualar muchas funciones y procesos los cuales realiza el cerebro para tener todas sus funciones cognitivas.

En 1965 Hubert Dreyfus, afirmo que la finalidad de la IA, conseguir crear una Inteligencia artificial de tipo general, era tan inalcanzable como el objetivo de los alquimistas del siglo XVII. Su argumento consistía en que el cerebro tiene la capacidad de procesar la información de forma global y continua, a diferencia de los ordenadores que utilizan conjuntos finitos de operaciones por medio de reglas dentro de un conjunto de datos finitos. Dreyfus fue de los primeros en abogar la necesidad de que la Inteligencia debe formar parte de un cuerpo con que pueda interactuar con el mundo que lo rodea. Según Dreyfus, la IA debe modelar necesidades, deseos, placeres, penas, formas de actuar, etc. Para poder alcanzar su objetivo final, no niega la posibilidad de que se pueda lograr, pero afirma que no será posible con los métodos de la IA simbólica y no corpórea, es decir, está en contra de la hipótesis SSF.

La aproximación corpórea con representación interna ha ido ganando terreno en la IA y actualmente muchos la consideran imprescindible en los avances de la inteligencia artificial de tipo general. El cuerpo conforma a la inteligencia y por lo tanto sin cuerpo no puede haber inteligencia general.

Los esfuerzos realizados en la investigación de Inteligencia Artificial se han centrado en la construcción de sistemas especializados, logrando resultados impresionantes, en especial en estos últimos años, todo esto gracias a la aparición de dos elementos muy importantes: La disponibilidad de grandes cantidades de datos y el acceso a computación capaz de poder analizarlos. Pero lamentablemente no se ha avanzado a la concepción de IA general. Con todos estos avances se logró llegar a la conclusión de que no es posible tener una comprensión profunda, sin que las máquinas logren tener conocimiento de sentido común.

Para lograr llegar a inteligencias artificiales verdaderamente inteligentes, es necesario implementar sistemas que integran percepción, representación, razonamiento, acción y aprendizaje. Este es un problema muy fuerte en la IA, ya que aun no se sabe como se pueden lograr integrar todos estos componentes de la inteligencia. Se necesitan arquitecturas cognitivas para poder integrar estos componentes de forma adecuada. Los sistemas integrados son fundamentales para la obtención de sistemas de inteligencia general.

En la actualidad una de las limitaciones de los sistemas de aprendizaje profundo es denominado como “olvido catastrófico” básicamente luego de que entrena un sistema para que realice una tarea en específico, y luego quieras entrenarlo para que realice otra tarea, olvidara lo que aprendiendo anteriormente. Esta limitante es prueba de que estos sistemas no son capaces de aprender nada. Otra limitación es que no tienen capacidad explicativa. Es importante el desarrollo de nuevos algoritmos de aprendizaje que no requieran enormes cantidades de datos para ser entrenados, así como un hardware mucho más potente para la implementación de estos algoritmos.

Otras técnicas más clásicas de IA que seguirán siendo un objeto importante de investigación extensiva son los sistemas multiagente, la planificación de acciones, el razonamiento basado e la experiencia, la visión artificial, etc. También asegura que habrá progresos significativos gracias a las aproximaciones biomiméticas para reproducir en máquinas el comportamiento de animales.

El autor afirma que las aplicaciones futuras de la IA serán aquellas que tienen que ver con la web, videojuegos, asistentes personales y los robots autónomos. De igual forma aplicaciones para el medio ambiente y ahorro energético, así como también las dirigidas a la economía y a la sociedad. También las aplicaciones de la IA fuerte (artes visuales, música, danza, narrativa) cambiaran de forma importante la naturaleza del proceso creativo.

Estos avances en las aplicaciones de la inteligencia artificial en las diferentes áreas mencionadas serán fundamental para el crecimiento de las mismas, impulsando la creación de páginas web y aplicaciones web con inteligencia artificial capaces de mejorar muchísimo las experiencias del usuario, de la misma manera se podrá lograr lo mismo con las demás áreas.

Ramon López afirma que “Por muy inteligente que llegue a ser las futuras inteligencias artificiales, incluidas las de tipo general, nunca serán capaces de iguales a las inteligencias humanas, ya que el desarrollo mental que requiere toda inteligencia compleja depende de las interacciones con el entorno y estas interacciones dependen a su vez del cuerpo, en particular del sistema perceptivo y del sistema motor”.

Para finalizar durante la mayor parte del artículo se hace énfasis en que lo necesario para lograr igualar la inteligencia humana es darle un cuerpo a las IA para que estas puedan tener interacciones con el ambiente que los rodean y de esta manera puedan desarrollar lo que sería la inteligencia tal como lo harían los humanos. Pero esto es algo que se puede debatir, ya que los humanos aprendemos por medio de investigaciones o por medio la experiencia, por lo que un sistema capaz de aprender podría superar lo que es la inteligencia humana sin necesidad de tener un cuerpo físico.

El autor sugiere que el objetivo de conseguir una inteligencia que se pueda comparar o superar a la humana es un reto bastante ambicioso, comparándolo con otros objetivos como explicar el origen de la vida, del universo o conocer la estructura de la materia. Desde mi punto de vista esto no es tan cierto ya que con los avances tecnológicos que se producen cada día, se acerca cada vez más a conseguir este objetivo. Puede que sea un proceso tardado, pero no será un objetivo imposible de lograr. Ya que si se compara con el origen de la vida existirán miles de teorías, pero nada que asegure a ciencia cierta este suceso de verdad.