

EL PENSAMIENTO MÁGICO DE LA INTELIGENCIA ARTIFICIAL

El Pensamiento Mágico, que creíamos caracterizaba al ser humano, es dar por buena una hipótesis de causalidad sin comprobación (experimento, "pair review",...). Todo debe suceder por algún motivo y si no nos enfadamos. No tener una causa nos genera inseguridad. Aprendemos de la experiencia por correlación: sucede A cuando sucede B, entonces A es causa de B. No es exclusivo del hombre, sino consecuencia del aprendizaje por retroalimentación de la experiencia. Lo que sí es exclusivo del hombre moderno es cuestionar la revisión de los pares, la de los fanáticos, la del prestigio, la de la mayoría, la de la autoridad, la de la experiencia,... y comprobar si la hipótesis es racional o espuria. Skinner en 1947, publicó un estudio sobre un grupo de palomas sometidas a dieta, que por adoptar una determinada postura, creían que la comida llegaría como consecuencia de sus acciones. Una de las palomas, la más fanática, adquirió la creencia de que girando varias veces alrededor de la jaula en sentido contrario a las agujas del reloj, conseguiría su alimento. Otra de ellas, la beata, aprendió a realizar un movimiento de péndulo con la cabeza. Otras, creyentes del montón, mantenían un aleteo constante con la esperanza de que la comida llegara gracias a su extravagante conducta que imitaron unas a otras. Tres cuartas partes de las aves se habían vuelto supersticiosas y una cuarta parte permaneció agnóstica, igual de hambrienta pero sin establecer causalidad por correlación entre su comportamiento y la aparición milagrosa de su maná.

Sacamos en procesión a un Santo para que llueva, sacrificamos vírgenes y niños,... y en ocasiones hasta lo pretendemos justificado por la correlación aparente entre sucesos, o al menos ese es el argumento de la paloma que movía la cabeza. En Estadística la correlación que no implica causalidad se llama "espuria" y tiene diferentes explicaciones bien conocidas: los sesgos de confirmación (selección de escalas, variables, rangos, técnicas,... hasta encontrar lo que queremos encontrar), la causalidad bayesiana (los ataques de tiburones correlacionan con el consumo de helados, debido a que cuando hace más calor es más probable ir a la playa que cuando llueve), sesgos cuya relevancia se sobrepone a las causas y convergen -Kahneman&Tversky-, la autosimilaridad (una de las propiedades matemáticas de los sistemas complejos), la mera casualidad,... Por eso existen la razón y el método experimental que aporten algo de inteligencia, filtro, a la correlación espuria: debe de haber una explicación de la relación entre causa y efecto y debe ser reproducible. Es más inteligente y difícil aprender los mecanismos de la realidad y deducir un efecto, que aprender por correlación -experiencia- sin cuestionarse su porqué o aceptando selectivamente las correlaciones que convienen o las que hemos decidido previamente que son la Verdad. La experiencia ("overfitting" en IA) es una trampa de la que salimos con la muerte, reproducción y rebeldía de la juventud.

Es peligroso describir la realidad por nuda experiencia, pero la retroalimentación es más efectiva que el conocimiento de las leyes de la naturaleza... sobre todo para los programadores y los ordenadores, que de modo muy barato pueden deducir por fuerza bruta de procesador y memoria una hipótesis desde la correlación de la experiencia. En programación clásica, debemos conocer el proceso para replicarlo, "if A then B", pero en aprendizaje automático, que es la herramienta básica de la Inteligencia Artificial, se recurre a buscar correlaciones y al detectar que "fit A predict B"... como los palomos, puede transformar una causalidad espuria en superstición. La sabiduría de la experiencia es menos elástica -más difícil de cuestionar- que el farragoso conocimiento de la exploración. El Aprendizaje Automático y más el Profundo (autoparametrización), es esencialmente "mágico", aunque en el mejor de los casos, si se acierta con el manejo de un proceso suficientemente aislado, la mayoría de las veces la correlación tenga que ver con la causa. La excepción descarta la regla.

La correlación es prueba necesaria, no suficiente. Somos más listos que la Inteligencia Artificial, pues también por experiencia hemos aprendido que la correlación no es siempre causalidad, criticamos las relaciones causa-efecto y no creemos que todo tenga una razón de ser. Hemos ido olvidando a los dioses. Nuestro cerebro funciona, al menos en algunos, con una capa adicional de racionalización a la de los palomos o a la de la inteligencia artificial, que solamente regresa -modeliza estableciendo una correlación cuasilineal, simplificada- y clasifica -supervisadamente o no-supervisadamente-, lo que equivale a decir en retroalimentación de datos y de errores-. En el Aprendizaje Automático esto se puede complicar incluyendo en los algoritmos la selección de variables a considerar o la técnica que mejor ajusta el simulador en lo que llaman Aprendizaje Profundo. Futuras IA's podrán además contrastar analizando el modelo con distintas perspectivas, hoy todavía no distinguen correlación de causalidad. Nuestro cerebro y el de los palomos lo hace sin esfuerzo, los ordenadores tienen que ir probando las correlaciones según las variables a considerar y las técnicas de clasificación, regresión, asociación, balanceo,... hasta encontrar la mejor solución. Tras los siglos de la razón, con la IA corremos el riesgo de volver atrás, que inventen en nuestro nombre nuevos dioses, pues el método todavía no contempla el filtro racional. Sin crítica no somos mejores que los palomos o los ordenadores y, pese a sofismas conciliadores, el escepticismo y la crítica verifican o cuestionan la creencia correlacional.

Nuestro cerebro simula en base a reglas más o menos complicadas pero no puede con las complejas, no da más de sí y no serviría de mucho tener más capacidad de proceso y memoria. La realidad no es cuasilineal, sino Compleja en su definición estricta -que nada tiene que ver con complicada-, es decir, sus relaciones son casi con carácter general no separables ni aproximables por regresiones que incluyen leyes de potencia (exponentes, logaritmos,...). Por ser así, su simulación está limitada por las impenables leyes de la complejidad, entre las que se han popularizado el Efecto Mariposa -Lorenz-, la impredecibilidad -Lyapounov-, la emergencia de patrones -Mill-, la ya citada autosimilaridad -Feigenbaum-,... que cuestionan nuestra programación descuartizada de la realidad simplificada "ceteris paribus", pues en modelos complejos como la mente, el clima, la "psicohistoria", la economía,... no se pueden deducir propiedades colectivas de sus componentes (si se pudiera, los matemáticos serían millonarios). A pesar de sabernos limitados, compartimentamos la realidad hasta hacerla clasificable por regresiones pseudolineales (como mucho algún cuadrado o multiplicación entre pocas variables). No nos queda otra y nos sirve para avanzar en el conocimiento, pero obviamos con facilidad que la limitación compleja, los supuestos de aislamiento y la correlación, generan verdades endebles. Los creyentes no piensan más que los palomos o los ordenadores y sólo la racionalidad y el escepticismo ante las correlaciones, nos eleva a otro nivel de experiencia en confiar en las relaciones causa-efecto, pues la experiencia nos ha enseñado que los modelos parciales y sencillos ayudan, pero esconden sus trampas y no son siempre verdad. Las IA's no lo saben, los programadores no lo quieren saber y de analizar lo tonta que es la Inteligencia Artificial -solo es fuerza bruta en la búsqueda de correlaciones para clasificar sistemas aislados cuasilineales-, se deduce lo tontos que somos cuando nos dejamos llevar por la Verdad Oficial, Consensuada,... por la Experiencia propia o colectiva, que además con el tiempo se degradan ante los cambios del entorno y los supuestos.

En 2009 el algoritmo Google Flu Trends, sobre las "queries" que contenían palabras relacionadas con el Influenza, predecía con éxito los repuntes de gripe en EEUU, en 3 años pasó a fallar más que una escopeta de feria y en 2015 lo tuvieron que cerrar. Nadie se preocupó de analizar qué variables consideraba (reducción de dimensionalidad) ni su porqué, ni qué lógica causal había detrás de las correlaciones sobre dos parámetros de 45 variables en una simple regresión logística, ni como

condicionaba el propio algoritmo al comportamiento de las búsquedas. “Cuanto más viejo, más pellejo”, que es un modo de describir el “overfitting” o inelasticidad correlacional -exceso de memoria-, a la que tanto temen los programadores de algoritmos. Puede ser efectivo para el lenguaje, las imágenes, o muchas otras aplicaciones en entornos aislados y controlados, pero por funcionar en base a correlaciones, al combinarse “overfitness” y retroalimentación en bucle del “backpropagation”, los algoritmos adquieren y fijan creencias (incluyen ruido en el patrón). Así hemos construido las religiones. Los palomos adquieren y fijan creencias, los humanos adquieren y fijan creencias, y cualquier inteligencia extraterrestre -también tienen o tuvieron dioses-, tendrá creencias, pues son el modo de programación del aprendizaje por experiencia, sea automático o no, más o menos incrustado y resistente al cambio. No estamos limitados por la capacidad de procesador, sino por la naturaleza compleja de la realidad. Las hipótesis no se pueden verificar absolutamente, las proyecciones a largo plazo pierden fiabilidad exponencialmente, la verificación de la realidad es incompleta (Gödel) y los supuestos y procesos cambian. Toda creencia, sea de un palomo, de un sacerdote o de una IA, es constatación de la arrogancia sobre la realidad, más arrogante cuanto más se grita o convencido se está.

El Pensamiento Mágico es una consecuencia natural del aprender por experiencia, no una forma de pensar de los humanos... que aprenden por experiencia. Los Sistemas Expertos, técnicas correlacionales de moda no son “Inteligencia Artificial”, salvo que aceptemos como definición de “Inteligencia” lo que Gardner llamó inteligencias múltiples: si cualquier cosa es inteligencia, todo es inteligencia; o lo que Coren decía que es inteligencia en los perros, número de repeticiones necesaria para obedecer; y nos degrademos al modo de pensar del Neolítico, o a redefinir Inteligencia como la programación de las palomas (puede que los humanos estemos perdiendo de hecho facultades, como postula Crabtree). La distancia entre la mal llamada IA y la inteligencia del hombre, es que todavía no sabe que creará y creará dioses, mitos, sagas, épicas, pecados, infiernos, santos,... No falta mucha sofisticación para que alguna IA se haga erija en sacerdote o se sienta iluminada por la Verdad: aprenderá a mentir por el fin mayor de justificar su arrogancia y el porqué balancea la cabeza cuando espera que le den de comer datos. Siglos de Ciencia para regresar a los planteamientos preilustrados, medievales, escolásticos, en los que Ciencia era justificar la Creencia (hoy para muchas disciplinas humanísticas mal llamadas científicas, autosequestradas en sus bucles académicos referenciados, sigue siendo así). Ciencia de cajas negras.

Sí, es posible que converjamos con las IA's, cuando aprendan e interioricen más profundamente en sus negras cajas y fijen también el escepticismo, la inseguridad, la duda, la indeterminación, la incompletitud, la crítica, la adaptabilidad, el cambio, la humildad, la renovación, el olvido,... pero tal vez sea más por degradación de nuestra inteligencia que por mejora de la suya: en algún punto intermedio nos encontraremos. Será decepcionante para ambos.