

## LA INTELIGENCIA ARTIFICIAL: SU IDENTIDAD, SUS CONFLICTOS, SUS POTENCIALIDADES Y LIMITACIONES ¥

Francisco RUEDA F

---

### RESUMEN

Como dice Turkle, la Inteligencia Artificial es una "cultura con aspiraciones globales". Pero, ¿qué tan razonables son éstas?, ¿cuáles son las limitaciones que tiene esta disciplina para cumplir sus objetivos?, ¿qué metas están buscando quienes trabajan en ella?. En este artículo trataremos de contestar algunos de estos interrogantes. Para ello reflexionaremos primero sobre lo que es la inteligencia, más adelante sobre lo que es la IA y sus diferentes enfoques y finalmente sobre sus limitaciones.

### INTRODUCCIÓN

En el año de 1957 Herbert Simon, uno de los gestores de la idea de Inteligencia Artificial (IA) hizo varias predicciones con respecto a ella: que antes de que terminara el decenio un programa de computador sería campeón mundial de ajedrez, que, dentro del mismo lapso de diez años, un programa habría de componer música de valor estético serio, y descubrir y demostrar un importante teorema matemático, y que en un plazo de diez años los programas serían la formulación corriente de la teoría psicológica. El triunfalismo del científico contrasta con lo modesto de los resultados obtenidos al respecto. Sin embargo, la IA sigue realizando esfuerzos por conseguir metas parecidas a las planteadas por Simon, buscando quizás obtenerlas en un plazo más largo, sin que seamos por el momento capaces de dar fechas específicas al respecto. En este artículo nos proponemos analizar cuáles son las pretensiones de la IA y qué tan factible es que se logren.

La Inteligencia Artificial es uno de los pilares de la llamada ciencia cognitiva que pretende establecer una nueva forma de entender la mente humana. Y para tener una mejor idea de hasta dónde van sus pretensiones miremos lo que dice Schank al respecto: "Somos en gran medida filósofos de la época moderna. Abordamos las mismas cuestiones que abordó Aristóteles, y todos los que vinieron detrás. Tenemos un método distinto para hacerlo. Este método puede sintetizarse en una sola palabra: proceso. Lo que decimos es que

¥ Este artículo se basa en el estudio hecho para RIBIE-COL con auspicio del MEN-SISNIED "La inteligencia artificial: sus principios básicos y sus aplicaciones educativas" [<sup>0</sup>] y presentado en el *Congreso Iberoamericano de Informática Educativa* con auspicio de COLCIENCIAS.

el enfoque correcto para un análisis de los hechos del mundo es un enfoque de proceso: ver cuáles son los pasos, determinar cuáles son las entradas, y suministrar algoritmos para ir de un lugar a otro, del lugar A al lugar B" (Schank, citado en [<sup>1</sup>]).

Y, para complementar lo anterior, miremos un poco más en detalle cuál es el aporte de la IA, según Turkle: "Pregúntese a distintos teóricos de la IA cuáles son las teorías más

importantes en dicho campo, y se obtendrán diferentes respuestas. Pero lo común a todas ellas es el énfasis en una nueva forma de conocimiento. Dicho nuevo modo de conocer requiere que todo, en especial todos los aspectos de la mente, se piense en términos de programa y procesamiento de información. Los fenómenos freudianos no quedan excluidos. Existe incluso una teoría computacional de los que se conocen como *actos fallidos* freudianos".

"Al afirmar la primacía del programa, la IA realiza una importante aseveración, y se anuncia, al igual que el psicoanálisis y el marxismo, como una nueva manera de comprender todas las cosas. En cada caso, hay un concepto central que reestructura la comprensión en gran escala: para el freudiano, el inconsciente; para el marxista, la relación con los medios de producción. Hemos visto que para el fanático, la programación es algo que se valora en sí mismo, pero que para el investigador en IA, la idea de programa tiene un valor trascendental: se considera que es la clave, el término hasta ahora faltante, para desentrañar misterios intelectuales " [1].

De las afirmaciones anteriores podemos ver que la IA es, como dice Turkle, una "cultura con aspiraciones globales". Pero, ¿qué tan razonables son éstas?, ¿cuáles son las limitaciones que tiene esta disciplina para cumplir sus objetivos?, ¿qué metas están buscando quienes trabajan en ella?. En este artículo trataremos de contestar algunos de estos interrogantes. Para ello reflexionaremos primero sobre lo que es la inteligencia, más adelante sobre lo que es la IA y sus diferentes enfoques y finalmente sobre sus limitaciones.

### QUÉ ES LA INTELIGENCIA

Antes de poder contestar a la pregunta de qué es la IA, debemos primero clarificar qué es la inteligencia. En la mayoría de los casos se entiende que ésta se refiere al enriquecimiento del pensamiento lógico-matemático y no al de la inteligencia en un sentido amplio. No es sorprendente esta confusión en una cultura que establece que "la inteligencia se mide por la capacidad de resolver problemas complejos; de escribir, leer y computar a ciertos niveles; y de resolver rápidamente ecuaciones algebraicas... (y que) postula la educación formal y el conocimiento académico o la cultura como la verdadera medida de la realización personal...." ([<sup>ii</sup>] p.19). Aun los psicólogos tienen a veces a dar una definición muy limitada de la inteligencia, como la siguiente: "Término general que se refiere a las capacidades de una persona para un amplio grupo de tareas, que implican vocabulario, cálculo, resolución de problemas, conceptos, etc.. Cuando se mide con un test de inteligencia tipificado, supone ciertas capacidades específicas, especialmente verbales" ([<sup>iii</sup>], p.638). Sin embargo conviene recordar que "los hospitales psiquiátricos están atiborrados de pacientes que tienen todas las credenciales debidamente presentadas, como de muchos que no las tienen..."([3], p.19), y que puede haber diferentes opiniones, como la siguiente: "el verdadero barómetro de la inteligencia es una vida feliz y efectiva vivida cada día y en cada momento de cada día " ([3], p.19).

Podemos concluir, entonces, que la inteligencia es un término muy amplio, que se refiere a una gran cantidad de habilidades y potencialidades de la mente y que, si bien en la mayoría de los casos se asocia con los aspectos formales de ella, esto no quiere decir que ésta sea la única manera de entender el término y que en muchos casos hay manifestaciones

más importantes de esta cualidad. Por esta razón al usar la definición tradicional estamos ignorando aspectos substanciales de ésta, que es indudablemente la principal cualidad del "homo sapiens".

Pensando en la situación de nuestro país, por ejemplo, nos preguntamos si no sería conveniente que en nuestra definición de inteligencia le diéramos cabida a aspectos como la capacidad de los ciudadanos para vivir en paz, que sospechamos no tiene mucho que ver con la solución de ecuaciones ni con la construcción de programas de computador.

Otra pregunta que surge al reflexionar sobre el asunto es si existe un mecanismo general de la inteligencia, aplicable a todos los casos, o si hay diferentes tipos de inteligencia. Siguiendo una primera hipótesis, que nos sugiere el sentido común, podemos afirmar que la inteligencia requerida para resolver problemas en diferentes áreas y situaciones tiene distintas características dependiendo del ámbito (Ingeniería, Administración, Humanidades, Arte, problemas familiares), y que para su desarrollo se requieren actividades diferentes en cada caso. El solucionador general de problemas, que constituyó un objetivo de los investigadores en las primeras etapas de la IA, muestra cuál era en esa época la opinión de la comunidad de IA al respecto, aunque su estruendoso fracaso nos indica que el asunto debe ser abordado cautelosamente. Más adelante se idearon los llamados sistemas expertos que van en la dirección contraria, que consiste en especializar las habilidades de los sistemas de IA para obtener altos niveles de desempeño en un área específica. Hasta el momento no parece existir una respuesta al asunto, sino sólo hipótesis al respecto. Nuevamente aquí encontramos que estas dudas van a pasar intactas al campo de la IA.

Podríamos seguir haciéndonos preguntas sobre la inteligencia pero lo que vamos a encontrar es que hay más interrogantes que respuestas, lo cual seguramente va a ser válido por mucho tiempo. Por esta razón la IA también va a estar colmada de dudas y de contradicciones. Lo que los creyentes de la IA esperan es que con sus herramientas muchas de ellas puedan ser resueltas, tanto en lo referente a la inteligencia natural como a la artificial.

## QUÉ ES LA INTELIGENCIA ARTIFICIAL

Tratar de dilucidar lo que es la IA es una labor ardua y sin mucha esperanza. En primer lugar hay que aclarar que tenemos dos vías para resolver el problema: podemos especular sobre lo que debería ser o acudir a quienes han trabajado en ella y analizar qué tienen en mente cuando hablan de este término. En este artículo tomaremos el segundo camino.

Las discusiones sobre la inteligencia de las máquinas tienen una antigüedad muy respetable. En un artículo ya clásico [<sup>iv</sup>], que podría considerarse como uno de los precursores del debate sobre la IA, Turing analiza algunas de las objeciones que se hacen a la posibilidad de que las máquinas puedan ser inteligentes, como la objeción matemática que se basa en consideraciones como el teorema de Gödel, "que demuestra que pueden formularse, en un sistema lógico cualquiera, afirmaciones suficientemente fuertes, que no pueden ni demostrarse ni negarse dentro del sistema, a no ser que el mismo sea inconsistente" [ibid], o la del avestruz que se pueden expresar diciendo que "las consecuencias de que las máquinas pensarán serían demasiado terribles; esperemos y

creamos que no pueda ser" [ibid] . Pero quizás uno de los aspectos más conocidos de este autor es la llamada "prueba de Turing", que permite establecer si una máquina es inteligente, y consiste en colocar a un interrogador en un cuarto y decirle que formule distintas preguntas a dos sujetos ubicados en otra habitación (y a los cuales no puede por lo tanto ver ni identificar directamente) con el fin de averiguar sus identidades; si al cambiar a uno de los sujetos por una máquina el interrogador no puede mediante sus preguntas diferenciar entre el sujeto y la máquina, podremos decir que ésta última es inteligente.

La prueba de Turing nos conduce muy naturalmente a la definición clásica de la IA que fue enunciada por Minsky, uno de los más destacados investigadores en el área, quien afirma que "es la ciencia de hacer máquinas que hagan cosas que si fueran hechas por los hombres requerirían inteligencia" [v].

Esta forma de definir la IA tiene varios inconvenientes: por un lado los criterios para decidir si una máquina es inteligente varían con el tiempo, para demostrar lo cual basta con mencionar que una de las actividades que antiguamente se hubieran podido considerar como que requieren inteligencia es la de jugar ajedrez y sin embargo hoy en día tenemos computadores que son maestros internacionales en esta disciplina y sin embargo nos resistimos a aceptar que por eso sean inteligentes. Uno de los episodios más pintorescos sobre el tema se refiere a Dreyfus, un famoso investigador del MIT, que no creía en los sistemas de IA y afirmaba que un niño de diez años puede vencer a la máquina, pero a pesar de eso desafió a uno de ellos a una partida de ajedrez, con el resultado que nos podemos imaginar, después de lo cual la comunidad de IA se divirtió mucho elaborando el siguiente titular: "un niño de diez años puede vencer a la máquina, pero la máquina puede vencer a Dreyfus".

Por otro lado, si tomamos la emulación del comportamiento como la definición de la IA, y fijamos sus objetivos en consonancia, poco podremos averiguar del funcionamiento de la mente humana, además de que no podemos garantizar que los procesos seguidos por la máquina para tener ese comportamiento sean similares a los que sigue una persona para resolver el mismo problema.

El inconveniente anterior no preocupa a quienes opinan, desafiando nuestra soberbia humana, que la inteligencia no se refiere únicamente a los actos de los humanos sino que es concebible hablar de este concepto en abstracto, y para ilustrarnos esta afirmación sostienen que de la misma forma como se crearon artefactos para volar sin seguir los modelos naturales, podremos quizás encontrar métodos de desarrollar inteligencia sin imitar las conductas de las personas.

Una de las objeciones clásicas a la IA definida como emulación del comportamiento de los humanos proviene de Searle quien en un artículo ya clásico [vi] nos presenta sus argumentos para mostrar que el imitar las conductas inteligentes no necesariamente es síntoma de inteligencia. Para ello nos presenta el siguiente caso: supongamos que a una persona que no sabe chino le damos una reglas o métodos escritos en papeles, en su idioma nativo, que le permiten contestar a cualquier pregunta que se le haga en este idioma. Searle se pregunta si por hacer esto podemos decir que la persona entiende el chino y se apresura a contestar que no y que por lo tanto no se puede hablar de una conducta inteligente. Por esta

razón, concluye que si bien puede ocurrir que una máquina satisfaga la prueba de Turing eso no nos indica que piense, o para decirlo más simplemente, que la prueba de Turing no prueba nada.

Quienes seguramente estarían de acuerdo con la definición de Minsky son los conductistas quienes sostienen que para estudiar la mente de las personas debemos guiarnos por su conducta ya que es lo único sobre lo que tenemos evidencia. Pero para muchos esta es una enorme simplificación que no es admisible, pues no sólo es importante el "qué" sino también el "cómo".

La siguiente definición de Charniak ([<sup>vii</sup>], p.6) proporciona otro punto de vista sobre lo que debe ser la IA: "el estudio de las facultades mentales a través del uso de modelos computacionales". Este enfoque es interesante porque abre nuevas perspectivas para la investigación en psicología, y para el desarrollo de la ciencia cognitiva. Es indudable que una gran parte del trabajo de la IA se ha hecho teniendo en cuenta este punto de vista. Por ejemplo, muchos de los investigadores del área realizaron procesos de introspección para identificar cuáles eran los métodos que usaban para resolver problemas y trataron de reflejarlos en sus sistemas.

La idea de que las máquinas nos pueden proporcionar metáforas para pensar sobre nuestra mente la encontramos ya en algunos científicos del siglo XVII, pero ha tenido un gran florecimiento en los últimos años como consecuencia del vertiginoso desarrollo de los modelos computacionales.

Podría haber, sin embargo, dos objeciones a ella: por un lado, los "reduccionistas" tecnológicos alegan que su interés es valerse de la IA para producir artefactos útiles, imitando quizás, pero no necesariamente, a los humanos; por otra parte, algunos opinan que los modelos computacionales de la mente que se pueden construir son muy limitados, y que no vale la pena la investigación en el área pues es una lucha sin posibilidades de éxito, y quizás sin sentido. Más adelante veremos algunas de los argumentos que se dan para defender este punto de vista.

La opinión de Newell y Simon [<sup>viii</sup>] con respecto a la IA es muy distinta. Ellos opinan que existen unos principios básicos que ha ido estableciendo la ciencia y que le han permitido explicar algunos fenómenos del mundo real al aplicarlos a situaciones específicas. Es el caso del principio que establece que algunas enfermedades son causadas por la presencia y multiplicación en el cuerpo de pequeños microorganismos que pueden ser transferidos de una persona a otra produciendo un contagio, el cual ha sido la base para el diagnóstico de muchas enfermedades (las cuales corresponden a las situaciones específicas), que han podido ser tratadas al identificar cuál es el microorganismo específico que las produce. En forma similar, se puede establecer un principio general que define cuáles son las condiciones requeridas para una acción inteligente, que puede aplicarse a problemas específicos para resolverlos. Así, de la misma forma que en el caso de las enfermedades hay un principio general (que algunas enfermedades son causadas por microorganismos) y con base en él se realizan investigaciones particulares (para identificar cuál es el microorganismo que causa una enfermedad específica), en el caso de la inteligencia podemos establecer un principio general (que enunciaremos a continuación)

que puede ser aplicado a situaciones o problemas específicos que requieran de una acción inteligente para resolverlos. El principio que proponen es el siguiente: "un sistema de símbolos físicos tiene los medios suficientes y necesarios para la acción inteligente". Un ejemplo de un sistema de símbolos físicos es un conjunto de elementos que permiten establecer espacios de búsqueda y hacer recorridos en ellos. Desde esta perspectiva, la IA se reduciría, dado un problema específico que se quiere resolver, a encontrar el sistema de símbolos físicos que lo representa (de la misma forma que para encontrar las causas de una enfermedad basta con identificar cuál es el microorganismo que la produce).

Schank [<sup>ix</sup>], por su parte, se pregunta cuáles son los fundamentos de la IA: las matemáticas, la ingeniería de software, la lingüística o la psicología, para concluir que esto depende de los objetivos de los investigadores y de los métodos empleados para construirla. Afirma, además, que hay dos objetivos básicos que persigue la IA: construir máquinas inteligentes e investigar sobre la naturaleza de la inteligencia. Además, y para complementar los principios anteriores, establece las principales características que debe tener un sistema inteligente (y con esto define cuáles son algunos de los aspectos más importantes que debe investigar la IA): capacidad de comunicación, conocimiento interno, conocimiento del mundo, intencionalidad y creatividad, las cuales expresan cualidades deseables en un ente inteligente.

La capacidad de comunicación se puede entender sin dificultad al afirmar que "entre más fácilmente nos comuniquemos con una entidad más inteligente parece ésta"

La capacidad de conocimiento interno la podemos expresar diciendo que "esperamos que las entidades inteligentes tengan algún conocimiento sobre ellas mismas".

El conocimiento del mundo debe permitir que "las experiencias pasadas estén codificadas y puedan ser usadas como una guía para procesar nuevas experiencias".

La intencionalidad se manifiesta en saber "cuándo se quiere algo y en conocer un plan para conseguirlo".

La creatividad consiste en "ser capaz de adaptarse a los cambios del entorno y de aprender de la experiencia".

Es indudable que las ideas de Schank nos sirven, no sólo para entender mejor lo que es la IA sino también para identificar qué es lo que debemos buscar de los sistemas inteligentes.

Para complementar las ideas de Schank podemos enunciar los tres atributos que atribuye Sloman [<sup>x</sup>] a los sistemas inteligentes: Intencionalidad, flexibilidad y pereza productiva. Además de la primera, ya comentada, vale la pena resaltar la segunda, que es quizás una forma de expresar algunas de las propiedades de Schank, y la tercera, que se refiere al hecho de que en la búsqueda de la solución a un problema se utilicen métodos que eviten hacer un análisis exhaustivo de todas las posibilidades, pues en su concepto eso no es inteligente (además de que en la mayoría de los casos no conocemos de antemano todas las posibilidades y en otros éstas son tan grandes que aún en el más veloz de los computadores no podemos hacerlo).

Para otros, los sistemas que utilizan la IA son aquellos que tienen la capacidad de aprender. Para ilustrar lo anterior, algunos dicen que no por el hecho de que posea muchos conocimientos se puede decir que algo es inteligente, ilustrando su argumento con el ejemplo de una enciclopedia (¿cuánto convendría que los profesores reflexionáramos sobre esta afirmación !). Pero esto no parece tan evidente si tenemos en cuenta que la interacción entre memoria, conocimientos e inteligencia es muy estrecha. En cualquier caso, si aceptamos la hipótesis, esto nos lleva a la definición de aprendizaje que es quizás más compleja que la que buscábamos establecer, pero refleja una característica muy deseable de los sistemas inteligentes (aunque no sepamos definir sin ambigüedades lo que es aprender).

Una pregunta que surge al indagar sobre la identidad de la IA es si ésta es una ciencia o una técnica. Como se puede ver por algunas de las definiciones que hemos dado, los investigadores del área tienden a identificarse con la primera hipótesis. Sin embargo si indagamos un poco más en el asunto, siguiendo las ideas de Narayanan [xi], encontramos que esta afirmación es difícil de sostener. Una de las características más importantes de las teorías científicas es que estas pueden ser falseadas o verificadas experimentalmente, pero en el caso de la IA (y en general de la computación) ¿cómo se puede hacer la confrontación con la realidad ?. Algunos sostienen que ésta se puede asimilar a la implementación a través de programas de computador o, que los programas de computador producidos mediante la IA tienen implícita una teoría. La dificultad reside en el hecho de que la suposición fundamental de la IA es que los conceptos computacionales pueden ser usados para describir la mente humana y es natural por lo tanto pensar que las teorías de la IA sean posibles de implantar en el computador puesto que son construidas a partir de conceptos computacionales. Hay otros argumentos expuestos por Narayanan que nos conducen a pensar que las teorías de la IA no son falseables y no constituyen por lo tanto un saber científico.

Si la IA no es una ciencia (que se justifica por sí misma), podemos afirmar entonces que es una tecnología que utiliza sus técnicas y herramientas en el desarrollo de aplicaciones del mundo real. Esto tiene algunas implicaciones como el hecho de que no existan unos principios básicos de la IA compartidos por todos los sistemas, de que su evolución se rija por los parámetros que le indica el mercado y, lo que es peor, que no es lícito relacionar los principios de la IA con los de la mente humana.

Si, de acuerdo con los principios anteriores, la IA no es una ciencia sino una técnica, ¿cómo podemos saber de sus avances ?. En el caso de las teorías científicas podemos decir que una es más evolucionada que la otra cuando puede explicar más fenómenos, pero esto no es aplicable al caso de las técnicas y por consiguiente no tenemos instrumentos para decidir si un principio de IA es mejor que otro, y por consiguiente si se ha evolucionado. ¿Será que podemos tomar la rapidez, o el uso de sistemas distribuidos como formas de medir la evolución de la IA? Muy posiblemente no pues estos aspectos dependen de factores externos. Esto nos muestra que no es fácil identificar parámetros para medir los avances de la IA, lo cual puede explicar en parte lo enconado de los debates en torno a esta área.

Como puede verse de lo anterior, no hay un acuerdo común sobre lo que es la IA y ni siquiera existe un consenso sobre si es una ciencia o una técnica. Sin embargo, podemos

recoger algunas de las caracterizaciones dadas por los autores arriba discutidos para establecer algunas ideas básicas sobre lo que persigue la IA: crear máquinas que se comporten como los humanos, estudiar los procesos de la mente humana y crear artefactos que tengan algunas cualidades que consideramos esenciales en el ser inteligente como la capacidad de comunicación, el conocimiento interno, el conocimiento externo, la intencionalidad, la creatividad la capacidad de aprendizaje y la flexibilidad. Si bien no podemos dar una definición de lo que es la IA, si podemos enunciar algunos de los objetivos que persiguen los investigadores que trabajan en el tema.

En vista de las dificultades que existen para definir qué es la IA algunos investigadores no consideran necesario establecer su punto de vista sobre el asunto, ni esclarecer si ella es una disciplina o cuáles son sus límites, sino aplicar sus principios en la solución de problemas. El problema con este enfoque es que no permite siempre definir sus rumbos, ni encontrar sus limitaciones.

### LOS DOS ENFOQUES

Uno de los debates clásicos sobre la IA lo protagonizan los simbolistas y los conexionistas. Los primeros, siguiendo principios parecidos a los planteados por Newell y Simon, opinan que los procesos inteligentes deben poderse expresar mediante operaciones sobre símbolos. Los segundos, inspirados de lo que se sabe sobre el funcionamiento del cerebro, creen que la inteligencia también puede residir en las conexiones que se establecen entre unidades, como en el caso de las que existen entre las neuronas. Un ejemplo, tomado de uno de estos últimos [<sup>xiii</sup>] nos puede ayudar a entender mejor el problema.

Para tratar de explicar lo que ocurre cuando una persona ve una manzana y descubre que en ella hay un gusano y nos lo dice, podemos acudir a dos teorías. La simbolista dice que cuando la persona ve la manzana esto la lleva a producir en su mente ciertos símbolos, a procesarlos y a comunicar esto al sistema que maneja el lenguaje quien se encarga de elaborar las frases. A la conexionista le parece muy complicada esta explicación y opinan que lo que hay son dos redes de neuronas conectadas, una encargada de la visión y otra del habla y no ven necesaria la generación de símbolos.

En el plano tecnológico el debate ha sido zanjado admitiendo que los dos métodos de hacer IA son válidos y complementarios, y que dependiendo del tipo de aplicación debe emplearse el método más apropiado. Por ejemplo, se ha visto que el método conexionista, o de redes neuronales, funciona bastante bien con aplicaciones como el reconocimiento de patrones o el manejo de voz, caracterizadas por la existencia de ruido y por la dificultad de especificar el dominio, mientras el segundo ha resultado más apropiado cuando la información disponible sobre el dominio está mas formalizada.

La pregunta de cómo funciona la mente, si en forma simbólica o conexionista es mucho más difícil de responder, aunque en este caso cabe también admitir que los dos métodos pueden coexistir, el conexionista a los niveles más bajos y el simbólico en la representación de aspectos más generales.

Se han construido tanto modelos conexionistas como simbólicos de los procesos mentales, como lo ilustran los dos siguientes ejemplos.



Neches, Langley y Klahr [<sup>xiii</sup>] elaboraron un modelo de la resta basado en reglas de producción (este es un formalismo muy utilizado en IA para representar conocimientos) con el que se pudieron predecir en forma precisa los resultados de un grupo de estudiantes. Con el fin de ilustrar sobre la estructura de dichas reglas veamos una de ellas:

SI    Se está procesando la columna c  
      Y el dígito d1 está en la columna c y la fila f1  
      Y el dígito d2 está en la columna c y la fila f2  
      Y la fila f1 está sobre la fila f2  
      Y el dígito d1 es mayor o igual al dígito d2

ENTONCES

      Calcule la diferencia entre d1 y d2  
      Y escriba el resultado en la columna c.

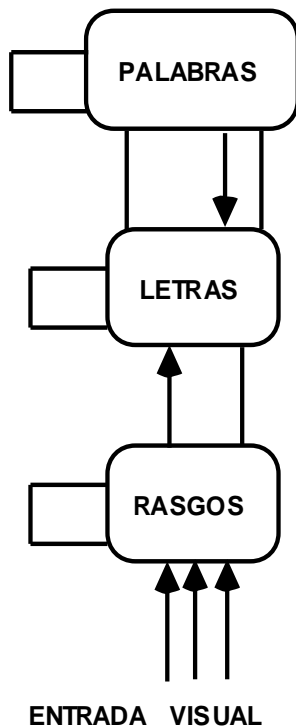
La regla anterior formaliza en reglas de producción el siguiente principio de la resta: "si al restar dos dígitos de una columna cualquiera el de la fila de arriba es mayor o igual que el de abajo, se puede hacer directamente la resta de los dígitos". En otras reglas se formalizan todas las demás situaciones que se pueden presentar en la resta (por ejemplo cuando el dígito de la fila de arriba es menor que el de la fila de abajo, en cuyo caso hay que "pedir prestado").

El modelo IAM ("Interactive Activation Model") busca representar el reconocimiento de letras y palabras, utilizando el esquema conexionista [<sup>xiv</sup>].

Según IAM, cuando una cadena de letras se presenta brevemente, los sujetos encuentran más fácil reconocer las letras si ellas forman una palabra que si forman una cadena sin sentido. Además, las cadenas de letras que no forman palabras pero que pueden ser pronunciadas son más fáciles de reconocer (que las que no se pueden pronunciar). Para dar cuenta de lo anterior se contruyó un modelo, en el que se supone que la percepción de una letra es más fácil si ésta se presenta en el contexto de una palabra, y que nuestro conocimiento de las palabras puede influenciar el proceso de percepción.

Según el modelo de IAM existen tres niveles de percepción: uno a nivel de rasgos, otro a nivel de letras y otro a nivel de palabras, que se refuerzan o inhiben mutuamente, como lo muestra la siguiente figura de la página siguiente.

Según el esquema mencionado se reciben rasgos visuales (los rasgos se refieren a las partes de las letras, como el I o el \_ ) los cuales determinan la percepción que se haga de ciertas letras. Los rasgos entonces refuerzan la percepción de ciertas letras, lo cual se representa mediante una flecha, e inhiben la percepción de otras, lo cual se representa en la figura con una raya sin flecha. Además ocurre algo similar entre las letras y las palabras.



Puede verse también que dentro de un mismo nivel hay inhibiciones (por ejemplo, la percepción de una determinada letra inhibe la percepción de otra).

Puede verse además que hay una doble relación entre letras y palabras: la percepción de una letra refuerza (o inhibe) la percepción de una palabra, pero también a la inversa, la percepción de una palabra refuerza (o inhibe) la percepción de una letra.

Figura 1. Modelo de SIAM sobre niveles de percepción

Una de las cosas que puede verse en los modelos anteriores es que quienes hacen modelos de la mente basados en IA no pretenden construir uno que pueda dar cuenta de su funcionamiento global, sino de aspectos parciales como el proceso de resta o el de percepción.

### **LAS LIMITACIONES Y POTENCIALIDADES DE LA IA**

Muchas son las objeciones que se le han hecho a la IA, hasta el punto de que ha caído en un cierto descrédito, que contrasta con el optimismo de sus primeros tiempos. Miremos cuáles son algunas de ellas.

Los formalistas dejaron de creer en ella cuando vieron que podía conducir a sistemas cuyas conclusiones no podían ser certificadas. Pero si analizamos las conductas inteligentes de los humanos vemos que estas también tienen esta propiedad. ¿Por qué entonces pedirle a las máquinas lo que no podemos exigir a las personas? Además, si en el intento de construir IA nos inspiramos de los métodos usados por los humanos, por qué no habrían de equivocarse nuestras máquinas.

Por otro lado, al reflexionar sobre los sistemas de IA, su falta de confiabilidad y sus contradicciones surge un cuestionamiento a la filosofía, muy bien expresado por Turkle: "De tal modo, la IA plantea una exigencia a la filosofía de la mente. La lógica tradicional no alcanza. Es necesario reemplazarla por una lógica que asuma la incoherencia y sepa cómo superarla. Algunos llegan a decir que esta nueva lógica de proceso debe trascender también las categorías de verdad y falsedad". [1]

Otra de las objeciones clásicas a la IA surge del debate "ideológico" entre los simbolistas y los conexionistas, pues los segundos identifican a la disciplina con los métodos simbólicos y, como no creen en éstos, son escépticos con respecto a los posibles resultados de la IA. Me parece que fuera del debate de si la mente funciona o no en forma simbólica, podemos resolver fácilmente el asunto diciendo que ambos métodos constituyen formas de trabajo de la IA, y que por lo tanto el dilema entre las dos tendencias no desvirtúa las posibilidades de la IA.

Otros afirman que si bien reconocen los avances logrados en la IA todavía identifican un gran abismo entre las potencialidades de las máquinas y las de los humanos, como se evidencia en el hecho de que un niño puede diferenciar fácilmente a un gato en un sofá mientras las máquinas tienen serias dificultades para hacerlo. La pregunta que surge entonces es ¿será que esto limita perpetuamente a las máquinas, o habrá alguna esperanza de que no?. Esto nos lleva a lo que podemos denominar el problema de la "escalabilidad" de las soluciones, muy conocido en computación, que se refiere a que no siempre el haber encontrado soluciones para casos triviales puede garantizarnos que podremos hacerlo para casos más complejos (por ejemplo el hecho de que existan tecnologías que permiten construir computadores con 16 procesadores no implica necesariamente que existan soluciones para el caso de 64000). Un caso típico al respecto ocurre con los llamados sistemas expertos construidos con las llamadas reglas de producción, pues algunos afirman que ya se ha demostrado que estos pueden realizar ciertas funciones con desempeños similares a los de las personas y que si en algunos casos fallan o son incapaces de resolver el problema, para solucionar el asunto sólo debemos "agregar más reglas". No parece razonable, sin embargo, pensar que esto sea así. Sin embargo tampoco es evidente que podamos por esto descalificar a la IA, primero porque las reglas son sólo uno de los métodos usados por la IA y segundo porque el no estar seguros de que con aumentar las reglas podremos resolver el problema no nos autoriza a afirmar que no lo vamos a lograr.

Pero en lo que hay quizás más acuerdo es en que si bien es posible que algún día se logre reconstruir artificialmente los comportamientos racionales de la mente es muy poco probable que podamos decir lo mismo para las emociones. Cuesta trabajo rebatir esta afirmación pero aun en este caso podría haber dudas razonables. Uno de los métodos más efectivos hoy en día en el tratamiento de los desórdenes mentales es la llamada terapia cognitiva que se basa en el principio de que todos los estados de ánimo son creados por las cogniciones o pensamientos, o más simplemente, que los sentimientos de las personas dependen de sus pensamientos [<sup>xv</sup>]. Al analizar en detalle esta teoría y sus aplicaciones podemos sin mucha dificultad imaginar modelos de la mente que simulen, por ejemplo la distorsiones cognitivas (como sobregeneralizaciones o razonamientos de todo o nada) que según la teoría cometen las personas y que las llevan a ciertos estados emocionales como los estados depresivos.

Otro ejemplo que podríamos dar para rebatir la hipótesis de que con el computador sólo podemos modelar los procesos racionales del pensamiento lo constituyen las teorías de Piaget, que permiten relacionar modelos formales del pensamiento con el análisis de los sentimientos y los juicios morales (al respecto se pueden consultar [<sup>xvi</sup>] y [<sup>xvii</sup>]).

En donde parece también haber un cierto acuerdo es en que procesos como la creatividad y la intuición son difícilmente capturables por un modelo computacional. Me parece también que en estos casos debemos ser cuidadosos en nuestras afirmaciones.

Existen sin duda muchos puntos de vista de lo que es la creatividad. En una primera aproximación podemos definirla como la capacidad de generar respuestas novedosas para resolver un problema, saliéndose de las soluciones tradicionales. Si tomamos provisionalmente esta como la caracterización de esta propiedad no es fácil encontrar argumentos para demostrar que un computador no puede hacerlo, y más si utilizamos la definición de la IA como una emulación de los comportamientos inteligentes del hombre, pues en ese caso no importa cómo lo haga este sino los resultados que produzca. Si seguimos esta vía, nos veríamos conducidos a admitir, por ejemplo, que un programa de computador que utiliza la estrategia de analizar un número grande de jugadas posible para elegir la mejor, podría ser creativo con la condición de que en una situación específica produzca una jugada "novedosa", aún utilizando este método. Y en el caso de temas como la creación musical no se puede a priori descartar la posibilidad de que un programa de computador pueda realizar una composición de algún valor estético, siempre y cuando logremos especificarle qué quiere decir esto. Sin embargo, si asociamos la creatividad con otros aspectos que tienen en cuenta cómo se genera esta (por un impulso artístico, por intuición,...), que es quizás lo que mucha gente piensa cuando habla de esta propiedad, tendríamos mucho más dificultad para imaginarnos a un computador creativo.

Con respecto a la posibilidad de que los computadores sean intuitivos la posición clásica de la comunidad de IA es que "lo que parece intuitivo puede ser formalizado y que, si se descubren las fórmulas correctas, es posible lograr que lo haga una máquina" [1].

## CONCLUSIÓN

De las reflexiones anteriores podemos concluir que en el ámbito de la IA hay más preguntas que respuestas. Pero también, que si bien los resultados obtenidos hasta el momento no han sido los esperados, por el triunfalismo de los primeros años, no tenemos argumentos suficientes para negar la importancia que puedan llegar a tener en el futuro los modelos computacionales de la mente, particularmente en un área en donde hay todavía muchos interrogantes debido a la complejidad del tema, y que por lo tanto podemos conservar un optimismo razonable de que éstos van a desempeñar un papel importante en el futuro.

## REFERENCIAS

- <sup>0</sup> RUEDA, F. (1992). La inteligencia artificial: sus principios básicos y sus aplicaciones educativa. RIBIE: *Primer Congreso Iberoamericano de Informática Educativa* (Santo Domingo, RD, Marzo 11 a 13 de 1992).
- <sup>i</sup> TURKLE Sherry, (1984). *El segundo yo, las computadoras y el espíritu humano*, Ediciones Galápagos.
- <sup>ii</sup> DYER W. W., (1978). *Tus zonas erróneas*. Barcelona: Ed. Grijalbo.
- <sup>iii</sup> MORGAN C., (1969). *Introducción a la Psicología*, Madrid: Ed. Aguilar.
- <sup>iv</sup> TURIN A. M. (1949) ¿ Puede pensar una máquina ?, en *SIGMA, El Mundo de las matemáticas*, James R. NEWMAN, (1969), Barcelona: Ediciones Grijalbo.
- <sup>v</sup> MINSKY M. L.,(1968) *Semantic information processing*, Cambridge, MA: MIT Press.
- <sup>vi</sup> SEARLE John. Minds, Brains and programs. En *The behavioral and brain sciences* **3**, pp. 417-424, 1980.
- <sup>vii</sup> CHARNIAK E., McDERMOTT D. (1985). *Introduction to Artificial Intelligence*, Ed. Addison Wesley.
- <sup>viii</sup> NEWELL, A. Y SIMON, H. (1976) "Computer Science as Empirical Research: Symbols and Search". Reproducido en J. L. GARFIELD (ed.) *Foundations of Cognitive Science: The Essential Readings*, Parangon House, New York, 1990.
- <sup>ix</sup> SCHANK R. C., What is AI anyway, *AI Magazine*, winter 1987.
- <sup>x</sup> SHARPLES M., HOGG D., HUTCHISON C., TORRENCE S., YOUNG D. (1989) *Computers and thought, a practical introduction to Artificial Intelligence*, Cambridge, MA: The MIT Press.
- <sup>xi</sup> NARAYANAN A., (1986) *Why AI cannot be wrong*, Exeter, UK: University of Exeter, Department of Computer Science.
- <sup>xii</sup> FELDMAN J. A., BALLARD D. H., Connectionist models and their properties, *Cognitive Science*, **6**, 205-254, 1982.
- <sup>xiii</sup> NECHES R., LANGLEY P., KLAHR D., *Learning, development, and production systems*, En KLAHR D., LANGLEY P., NECHES R., (editores, 1987) *Production systems models of learning and development*, Cambridge, Mass : MIT Press.

- xiv McCLELAND J. L., RUMELHART D. E., An interactive activation model of the effect of context in perception, partes I y II, *Psychological Review*, **88**, 5 y **89**,1, pp. 375-407 y 60-94, 1980, 1981.
- xv BURNS David (1980). Feeling good, the new mood therapy, *New American library*.
- xvi PIAGET J. (1957). *Le jugement moral chez l'enfant*, Presses Universitaires de France.
- xvii PIAGET J. , INHELDER B., *Psicología del niño*, Capítulo IV. Madrid: Ediciones Morata S. A.