

Las claves de la mente humana a través de la supercomputación

(Conferencia impartida el 22 de Enero de 2015 en la Facultad de Educación, Economía y Tecnología de Ceuta en el Acto Académico de Santo Tomás de Aquino)

Julio Ortega Lopera

Catedrático de Arquitectura y Tecnología de Computadores

Universidad de Granada

En primer lugar agradezco a la Facultad de Educación, Economía y Tecnología de Ceuta la invitación a dar esta conferencia cuya preparación me ha permitido reflexionar sobre algunas de las consecuencias de la supercomputación en el estudio de la mente humana, y plantearlas en un ámbito menos técnico que aquellos para los que uno suele preparar sus presentaciones y conferencias. Como indica el título (que también tengo que agradecer a la Profesora Prieto, Subdirectora de esta Facultad), pretendo poner de manifiesto en qué sentido las capacidades de cómputo cada vez mayores de los computadores (los supercomputador no son más que los computadores más potentes en un momento dado, y se trata por tanto de un concepto relativo en el tiempo) permiten desvelar aspectos de la mente humana, pudiendo incluso poseer capacidades mentales humanas. Para ello, he intentado centrarme en los aspectos que permiten entender las perspectivas de la evolución de las arquitecturas de cómputo, además de plantear algunas de las posibles consecuencias sociales de la mejora en las tecnologías de la información y las comunicaciones (TIC).

Desde bastante antes de los primeros computadores se había planteado la posibilidad de que existieran máquinas pensantes, y se ha especulado bastante acerca de pueda existir una mente que emerja de la materia de la que están hechos los computadores (en nuestro caso circuitos electrónicos) de la misma forma que la mente humana emerge del cerebro. El test de Turing, propuesto en un artículo de 1950 (apenas cinco años después del primer computador electrónico, el ENIAC) proponía una estrategia para determinar si un computador poseía o no inteligencia humana. Mi propósito en esta charla es mostrar en qué sentido la capacidad de cómputo de los computadores ha llevado a contemplar como algo plausible la posibilidad de disponer de computadores con capacidades mentales comparables a las humanas. De hecho, hay predicciones de que en 30 años la inteligencia artificial de los computadores superará a la humana. No parece mucho, pero en treinta años pueden producirse mejoras casi inimaginables en la capacidad de los computadores.

En este sentido, recuerdo que hace casi 30 años (en 1986) los programas de la Tesina que estaba realizando tardaban alrededor de 20 horas del procesamiento en un MV10000 de Data General, un computador de 1984, uno de los más potentes de la Universidad de Granada en ese momento, que proporcionaba hasta 2.8 MIPS (Millones de instrucciones por segundo) con una frecuencia de reloj de 7.1 MHz (Millones de ciclos de reloj por segundo). Teniendo en cuenta que debía compartirse con otros usuarios, los programas tardaron 20 días en completarse. Como consecuencia, leí la Tesina por la mañana y por la tarde tuve que tomar el tren a San Fernando para completar los seis meses de mis prácticas como alférez de complemento (precisamente en el C.I.R. donde hacían la correspondiente instrucción militar los reclutas destinados a Ceuta). Hoy en día, los programas de mi Tesina hubieran tardado menos de cinco minutos... Todos, de una u otra forma hemos sido testigos de cómo los ordenadores han dejado de ser máquinas enormes ubicadas en Centros de Cálculo para convertirse casi en un electrodoméstico más (los Computadores Personales) y cómo ha pasado a intervenir y controlar tareas en ámbitos muy diversos que afectan a casi todas las facetas de nuestras vidas. En nuestros bolsillos tenemos teléfonos móviles (*smartphones*) con más potencia de cálculo que un computador personal de hace pocos años. Algunos datos para dar una pequeña idea de la magnitud de estos cambios.

Los agentes de la evolución de los computadores

En 1946, el ENIAC (considerado el primer computador electrónico) disponía de un procesador capaz de terminar 5000 sumas/s, 357 multiplicaciones/s o 35 divisiones/s, funcionaba a una frecuencia de 100 KHz, consumiendo una potencia de 174 KW, alcanzando por tanto 0.03 operaciones/s/W (0.03 operaciones por segundo y watio). Como he comentado, actualmente se pueden encontrar computadores de características muy diversas en diferentes dispositivos. El supercomputador más potente que se incluye actualmente en el TOP500 (la lista de los 500 computadores más potentes que se puede consultar en www.top500.org) dispone de más de tres millones de procesadores (hoy se suelen denominar núcleos de procesamiento, o *cores* en inglés), funciona a 2.2 GHz (2.2 miles de millones de ciclos de reloj por segundo) es capaz de completar 35×10^{15} operaciones/s (35000 billones de operaciones por segundo), consumiendo una potencia de 17.81 watos. Esto supone 2×10^9 operaciones/s/W (2000 millones de operaciones por segundo y watio). Existen *smartphones* de cuatro núcleos capaces de terminar 5×10^9 operaciones/s a 1.2 GHz y consumiendo 5 W, con lo que proporcionan unas 10^9 operaciones/s/W. Un computador personal puede proporcionar unas 5.32×10^9 operaciones/s/W. En casi 70 años, se ha incrementado la velocidad de los circuitos (ciclos de reloj por segundo) en un factor de 10^4 (diez mil) el número de procesadores en un factor de 10^6 (un millón) y la capacidad de cálculo y la eficiencia energética en un factor de alrededor de 10^{12} (un billón de veces). Un pequeño paréntesis para poner de manifiesto la relevancia de la ingeniería de computadores, área en la que desarrollo mi labor docente e investigadora: la mejora en la capacidad de procesamiento no sólo se ha debido a la velocidad de los circuitos electrónicos que constituyen el computador (la frecuencia de reloj *solo* ha aumentado en un factor de 10000, sino también gracias a la estructura y organización de los computadores que ahora pueden aprovechar el procesamiento paralelo de varios núcleos o procesadores). De hecho, si se multiplica el factor de 10^4 de la mejora en velocidad de los circuitos y el factor de 10^6 del número de procesadores, todavía queda un factor de 100 que debe explicarse por factores relacionados con la organización del computador.

Para analizar las posibilidades de los computadores en un futuro no es suficiente conocer las mejoras que han proporcionado hasta hoy sino que hay que predecir sus capacidades futuras. Pero para un ingeniero tener una cierta capacidad de predicción es fundamental: lo que se diseña hoy se fabricará y aparecerá en el mercado transcurrido un tiempo desde su concepción y debe competir en ese momento futuro con los otros productos existentes en el mercado. El ingeniero debe tener un modelo plausible de lo que se puede alcanzar y para eso se deben tener en cuenta los factores que determinan la evolución de los computadores y analizar hasta que punto pueden mantener su influencia.

En realidad, en la evolución de los computadores concurren varios agentes que interactúan entre sí, influyéndose mutuamente: tecnología, economía, aplicaciones, y por supuesto la ingeniería de computadores [1]. La tecnología crea las posibilidades que permiten a los ingenieros diseñar computadores eficientes para las aplicaciones socialmente demandadas, y establece los límites para los sistemas viables desde el punto de vista económico. Por otro lado, las aplicaciones más demandadas contribuyen a seleccionar las características de las arquitecturas de computador en un momento dado. Estas interacciones entre los distintos agentes dan lugar a varios ciclos de realimentación que dificultan la predicción detallada de los efectos futuros. Sin embargo leyes como la ley de Moore (1965), la ley de Amdahl (1967) y la ley de Grosch (1953), relacionadas respectivamente con la tecnología, las aplicaciones, y la economía, y su interacción con el resto de agentes, han condicionado las características de los computadores en el tiempo.

Quizá la más conocida es la ley de Moore, que plantea que el número de transistores en un circuito integrado se dobla pasado un cierto tiempo. Esta ley se ha venido cumpliendo desde que se enunció aunque el tiempo tras el cual se dobla la cantidad de transistores en el circuito integrado ha ido cambiando, cuando se publicó la predicción en 1965 era un año, después fue año y medio, ahora parece más bien ser de dos años. En cierto modo se trata de una profecía auto-cumplida: si se intenta desarrollar una tecnología que no sea capaz de doblar el número de

transistores en un circuito integrado en el periodo de tiempo que establece la ley de Moore se podría dejar de ser competitivo (otros fabricantes intentarán conseguirlo, y podrían tener éxito), pero también es arriesgado ser más ambicioso puesto que seguro que será necesario disponer de una tecnología que podría no estar disponible en el momento en que se precisa (el ritmo aceptado por todos es el que marca la ley de Moore), o porque es difícil justificar la inversión a realizar ante los responsables de la empresa.

La ley de Moore también ha definido el paradigma que ha marcado el ritmo de mejora de los computadores basados en el uso de circuitos integrados. Así, es posible diseñar microprocesadores más complejos y más rápidos dado que se pueden utilizar más transistores más próximos entre sí y por lo tanto establece una predicción respecto a la mejora en las prestaciones de los procesadores futuros. También aquí se ha supuesto que la capacidad de los microprocesadores debe doblarse transcurrido un periodo de tiempo fijo que coincidía con el establecido para la ley de Moore. Incrementar la capacidad de algo por un factor constante después de un intervalo de tiempo fijo (lo que plantea la ley de Moore) implica un crecimiento exponencial. Por ejemplo, si cada dos años se doblan las prestaciones, pasados cuatro tenemos una mejora de cuatro, pero pasados 6 tenemos una mejora de 8 y pasados 8 una mejora de 16 (en un caso de mejora lineal, pasado un cierto periodo de tiempo, se produce un incremento de la misma cantidad). Tal y como se indica en [2] y se pone de manifiesto en la Figura 1 (elaborada con datos extraídos de [2]), la dinámica de cambio exponencial descrita y justificada en el caso de la ley de Moore, es característica del desarrollo de los computadores, y también se ha producido cuando se utilizaban otras tecnologías, como por ejemplo la electromecánica.

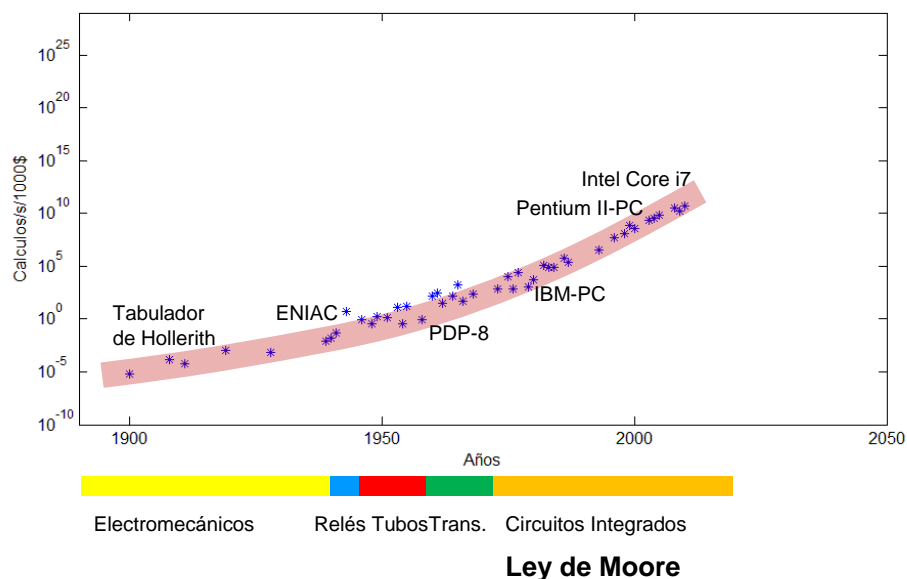


Figura 1. Evolución exponencial de la velocidad de cómputo (cálculos por segundo) por cada 1000 dólares de coste del sistema de cómputo (datos extraídos de [2]).

Los computadores actuales

A continuación resumiré algunas de las características principales del espacio de las arquitecturas de computador a las que hemos llegado. Éstas constituyen la base para entender las capacidades de los computadores en un futuro próximo.

Actualmente, los computadores se basan en procesadores, que ejecutan las instrucciones de los programas a través de los cuales se describe la aplicación que se procesa. Esas instrucciones, junto con los datos que utilizan, se encuentran en la memoria del computador. Existen elementos, denominados *buses*, que permiten conectar los procesadores y las memorias con los discos, u otros dispositivos de almacenamiento de datos e instrucciones que no se están procesando en un momento dado, y con otros circuitos que permiten adaptar las señales internas

con las que trabaja el computador a las señales externas a través de las que se recibe información de periféricos (teclado, pantalla táctil, un sensor de presión, un acelerómetro, etc.), se actúa sobre otros periféricos (pantallas, impresoras, válvulas, etc.), o se accede a redes a través de las que los computadores pueden comunicarse. Precisamente Internet no es más que una estructura de redes interconectadas que permite acceder, desde cualquier computador personal o dispositivo móvil, a servicios que se están ejecutando en computadores remotos con una arquitectura adecuada para ello (servidores de almacenamiento, servidores de cómputo, servidores web, etc.). Han surgido así dos paradigmas muy relevantes actualmente: Internet de las Cosas (dispositivos conectados entre sí y conectados a computadores a través de Internet que pueden encontrarse en electrodomésticos, ropa, sensores, etc.) y la computación en nube (*cloud computing*) que permite el acceso a servicios ubicados en computadores remotos (servidores), normalmente a través de páginas *Web*, pagando por el uso de los mismos. Un ejemplo bastante extendido de uso de *la nube* está en aplicaciones del tipo de *Dropbox*, a través de la que se tiene acceso a almacenamiento masivo, que se contempla como si fuera espacio en la estructura de almacenamiento (carpetas y ficheros) del propio computador. No obstante, también se puede acceder a una mayor capacidad de procesamiento. Este es el caso de SABLE, que permite que el usuario envíe una secuencia de aminoácidos de una proteína a través de una página web, y aproveche la capacidad de un servidor que ejecuta una aplicación para determinar cual es la estructura tridimensional de la proteína (información de gran utilidad para, por ejemplo, la síntesis de medicamentos). No solo se accede a una aplicación sino también a la potencia de cálculo del computador donde se ejecuta.

La velocidad de procesamiento de un computador depende de la velocidad con que los procesadores que lo constituyen ejecutan las instrucciones de los programas y del tiempo de acceso a esas instrucciones y a los datos que procesan. La velocidad del procesador depende el número de instrucciones que termina en un ciclo de reloj y de la frecuencia de reloj a la que funciona (ciclos de reloj por segundo). Si un procesador puede terminar de procesar hasta una instrucción por ciclo y funciona a un GHz (mil millones de ciclos por segundo) su velocidad máxima (también denominada velocidad pico) es de 1 instrucción/ciclo x 1 GHz = 1 GIPS (mil millones de instrucciones por segundo). Cuando se analiza la capacidad del procesador en aplicaciones que utilizan números reales (usualmente denominados datos en coma flotante) se considera el número de operaciones en coma flotante (FLOPs, del inglés FLoating POint operations) que el procesador es capaz de completar por ciclo. Así, si un procesador de 1 GHz termina una operación en coma flotante por ciclo, tendría una velocidad de 1 FLOP/ciclo x 1 GHz = 1 GFLOPS (mil millones de FLOPs por segundo). Hasta aproximadamente la primera década del siglo XXI, el aumento del número de transistores en los circuitos integrados se ha aprovechado para conseguir procesadores más rápidos fundamentalmente de dos formas. Por un lado se han diseñado procesadores con estructuras más complejas (dado que se dispone de más transistores en el circuito integrado) para poder procesar más instrucciones por ciclo. Por otra parte, se utilizaban frecuencias de reloj más elevadas (dado que al ser los transistores más pequeños eran más rápidos). Por lo tanto, las mejoras en los dos factores que determinaban la velocidad del procesador aumentaban y contribuían a mantener el ritmo de mejora exponencial en las prestaciones que marcaba la ley de Moore. Esta estrategia para mejorar la velocidad de los procesadores se mantuvo hasta que el incremento en el consumo energético de los circuitos integrados debido al aumento de la frecuencia de funcionamiento de los transistores y al incremento en el número de transistores en el circuito integrado la hizo inviable: un circuito integrado tendría que disipar tanta energía por unidad de superficie como el Sol. La alternativa ha consistido en aprovechar el incremento que proporciona la tecnología electrónica en cuanto al número de transistores en los circuitos integrados para incluir más procesadores (llamados ahora núcleos de procesamiento, o *cores* en inglés) en los microprocesadores y mantener la frecuencia de reloj. Así cada vez se dispone de microprocesadores con más núcleos (los microprocesadores multi-núcleo o multi-cores) con frecuencias que prácticamente no crecen. Por esta razón, actualmente la mejora de prestaciones de los computadores proviene de incluir más procesadores en el computador y aprovechar el trabajo paralelo de los mismos. Hay que tener en cuenta, que en estas arquitecturas paralelas la forma de distribuir los datos entre los

distintos procesadores y memorias y la necesidad de comunicación entre los procesadores a través de las redes que los interconectan es fundamental para conseguir buenos rendimientos en las aplicaciones y que el desarrollo de programas paralelos que aprovechen eficientemente las características de la máquina introduce una dificultad adicional y debe tenerse muy en cuenta. No obstante, no nos detendremos más en estas cuestiones. Nuestro objetivo simplemente ha sido poner de manifiesto que se verán computadores más potentes que basarán el aumento de su capacidad de procesamiento en el aumento del número de procesadores que utilizan.

Los supercomputadores son los computadores que en cada momento tienen las capacidades de procesamiento más elevadas (un supercomputador deja de serlo con el tiempo). Desde 1993 se viene publicando dos veces al año la lista TOP500, que incluye los 500 computadores más potentes del mundo (www.top500.es). Precisamente en el año 2007, el computador UGRGrid de la Universidad de Granada entró en la lista TOP500, en la posición 467. Disponía de 1264 procesadores (núcleos de procesamiento), con 3 TBytes (3 billones de Bytes, siendo un Byte igual a 8 bits) de memoria, 20 TBytes de almacenamiento y una velocidad de 4.2 TFLOPS (4.2 billones de operaciones en coma flotante, es decir operaciones con números reales, por segundo). Posteriormente, en el año 2013 la capacidad de este supercomputador se ha ampliado a través de Alambra, una máquina con 1808 núcleos de proceso, 4.28 TBytes de memoria, red de interconexión Infiniband QDR, y 40 TBytes de almacenamiento. Ofrece unas prestaciones de 31.6 TFLOPS y su coste fue de algo menos de 700.000 euros. Los investigadores de la UGR que trabajamos en estos ámbitos nos podemos sentir satisfechos, independientemente de las posibilidades de acceso a infraestructuras más potentes gracias a la integración en distintas redes.

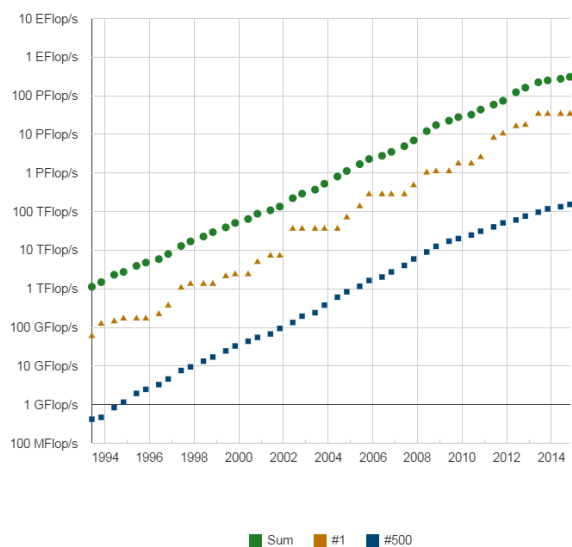


Figura 2. Evolución exponencial de la velocidad de cómputo (operaciones en coma flotante por segundo) en la lista TOP500 para el primer computador de la lista (#1), el último (#500) y la suma de los 500 computadores (Sum).

En la Figura 2 se muestra la evolución temporal de la velocidad de procesamiento de los computadores del TOP500. Actualmente existen computadores que pueden conseguir decenas de miles de billones de operaciones en coma flotante por segundo. Suponiendo que una persona fuese capaz de realizar una operación por segundo (lo que correspondería a una velocidad de cálculo considerable) tardaría más de mil millones de años en hacer lo que uno de estos supercomputadores hace en un segundo. Como se puede ver, la evolución ha sido exponencial, y se prevé que se mantenga este ritmo. El mismo comportamiento exponencial que se observaba en la Figura 1 para la velocidad de cómputo por unidad de coste.

Los computadores y la comprensión de la mente humana

El Cerebro humano constituye uno de los principales desafíos del conocimiento para el siglo XXI. Con una masa aproximada (y según los individuos) de kilo y medio, y en un volumen de un litro y medio (por lo tanto una densidad similar a la del agua), el cerebro humano incluye alrededor de 86×10^9 neuronas (casi 100.000 millones, de las que entre 20 y 30×10^9 están ubicadas en el córtex cerebral) interconectadas a través de 10^{14} - 10^{15} sinapsis, consume unos 10 W, y consigue procesar alrededor de 10 bits por segundo y almacenar unos 100 Tbits (100 billones de bits). A pesar de ser el sistema más complejo conocido, comprender el cerebro humano se ha convertido en un objetivo plausible gracias a la mejora de las técnicas de imagen digital, por un lado, y a la capacidad de cómputo de los supercomputadores actuales y la que se espera que se alcance en un futuro no muy lejano, por otro lado. Así, recientemente se han iniciado proyectos de investigación como el Human Brain Project en Europa (www.humanbrainproject.eu/es) y la BRAIN Initiative en los EE.UU. (<http://braininitiative.nih.gov/>), donde aparecen estimaciones que sitúan en este año 2015 la posibilidad de simular el cerebro de una rata, y la simulación del cerebro humano en 2013.

Para entender un sistema complejo se suele utilizar una aproximación basada en la superposición de distintos niveles de complejidad. El comportamiento detallado de los elementos que constituyen un nivel permite abstraer los principios de comportamiento de los elementos del nivel superior. El estudio de ese nivel y su aprovechamiento no necesita tener en cuenta los detalles de funcionamiento del nivel inferior. Un ejemplo de esta aproximación a la complejidad lo tenemos precisamente en el computador. El conocimiento de las ecuaciones según las que funcionan los transistores y los circuitos electrónicos, permite generar modelos de circuitos que se utilizan en el nivel de lógica digital para diseñar el hardware. En ese nivel las entradas y salidas de los distintos sistemas son ceros o unos y no se tienen en cuenta los detalles relativos a las corrientes que se tienen en cuenta en el nivel inferior de circuito electrónico. Igualmente, las secuencias de ceros y unos que esos los circuitos digitales generan al ejecutar un programa se encapsulan en instrucciones máquina. Así, los programas elaborados utilizando lenguajes de programación de alto nivel, como C, se traducen a través de un compilador, en las correspondientes secuencias de instrucciones máquina. Un programador no tendría que ser consciente de los detalles de la ejecución de las instrucciones máquina, salvo en los casos en los que esté interesado en obtener un comportamiento óptimo en cuanto a velocidad de procesamiento o consumo energético, por ejemplo, que el compilador no es capaz de alcanzar de forma automática. En el caso del cerebro humano, con cuyo estudio se pretende comprender la relación entre el comportamiento inteligente propio de la mente humana y el sustrato físico que proporciona el cerebro, igualmente se pueden distinguir distintos niveles de abstracción que van desde el nivel más alto que correspondería al de la mente humana hasta el nivel de genoma, pasando por el de regiones neuronales, microcircuitos neuronales, neuronas, sinapsis, etc. Según las capacidades de los computadores disponibles ha habido aproximaciones diferentes para abordar el conocimiento de la mente humana que ha incidido más sobre unos niveles u otros según los requisitos de computación que demandaban. A continuación describiremos algunas de estas aproximaciones, haciendo referencia a algunos de los hitos más relevantes, las capacidades de los computadores que las han hecho posible o que serán necesarias para hacerlas posibles en el futuro.

Un primer ejemplo de acercamiento de los computadores a las capacidades de la mente humana se pone de manifiesto a través de la capacidad de los computadores para desenvolverse de forma eficiente en juegos de carácter fundamentalmente lógico como el ajedrez. Así, en 1997, el computador Deep Blue de IBM derrotó al campeón del mundo G. Kasparov. El computador estaba basado en un RS/6000 con 32 procesadores y 512 coprocesadores de propósito específico, funcionaba a 135 MHz, y era capaz de analizar alrededor de 200 millones de posiciones de ajedrez por segundo. Dado que el análisis de una posición implicaba la ejecución de varios miles de instrucciones, el computador debía ser capaz de ejecutar cientos de miles de millones de instrucciones por segundo, es decir cientos de Giga-instrucciones por segundo

(GIPS). En ese año 1997 los computadores incluidos en el TOP500 proporcionaban entre 1068.0 y 7.7 GFLOPS (miles de millones de operaciones en coma flotante por segundo).

Sin embargo, para aquellos que defienden que las capacidades mentales humanas solo pueden emerger del cerebro humano, no extrañó mucho que un computador pudiera superar la capacidad humana en el ajedrez y la victoria de Deep Blue no se juzgó en estos entornos como algo sorprendente y realmente significativo en cuanto al conocimiento de la mente. Los ordenadores son básicamente máquinas lógicas y el ajedrez, a fin de cuentas, es un juego de lógica. Los críticos defendieron que los ordenadores no dominarían las sutilezas del lenguaje humano, con metáforas, juegos de palabras, chistes, dobles sentidos, etc. De hecho, la forma de trabajar del computador, analizando las consecuencias de las posiciones tras cada movimiento posible y hasta un determinado número de jugadas consecutivas, es bastante diferente de la forma en que un jugador experto actúa en una partida. Así, el propio Kasparov comentaba que apenas podía procesar más de una posición por segundo. Es la capacidad de reconocer patrones de la mente humana es lo que permite a un buen jugador distinguir las oportunidades que surgen tras un movimiento a partir de una distribución dada de las piezas. El hecho de que un experto en un campo dado llega a dominar alrededor 100.000 fragmentos de conocimiento [2], también pone de manifiesto que existe una relación entre la capacidad de reconocer patrones y la actividad inteligente. Así se estima que Karparov había aprendido unas 100.000 posiciones del tablero de ajedrez, y parece que en las obras de Shakespeare se incluían palabras con alrededor de 100.000 significados diferentes. Para entender esta relación hay que tener en cuenta el modelo, al que nos referiremos más adelante, que considera que la mente humana es un sistema jerárquico de reconocedores de patrones.

No obstante, en 2011 se produjo otro hito importante respecto a la implementación de comportamiento inteligente en los computadores. A pesar de que se había creído que el lenguaje humano estaría siempre fuera de la capacidad de los computadores, en ese año el computador Watson de IBM ganó a los dos mejores jugadores de un concurso de televisión denominado Jeopardy! (parecido al trivial aunque las respuestas se expresan en forma de preguntas). En dicho concurso, el computador recibía una pregunta del presentador (esta vez con sonido, nada de ficheros de texto o teclados) y respondía con sonido sintetizado en menos de tres segundos. El conocimiento necesario se extraía de estructuras de datos que el propio computador había generado a partir de 200 millones de páginas de documentos en lenguaje natural (entre ellos la Wikipedia): un total de 4×10^{12} Bytes (4 Terabytes, es decir 4 billones de Bytes) de conocimiento basado en el lenguaje. A continuación reproducimos algunos ejemplos de preguntas del concurso junto con las respuestas (preguntas en realidad) correctas generadas por el computador:

- En el tercer acto de una ópera de Verdi de 1846 este Azote de Dios es acuchillado hasta la muerte por su amante Odabella: ¿Quién es Atila el Huno?
- Wordsworth dijo de ellas que planean de aquí para allá pero que nunca vagan sin rumbo: ¿Qué son las alondras?
- Un discurso largo y tedioso escrito en el trivial aderezo de un pastel: ¿Qué es una arenga de merengue?

Resulta bastante evidente que las respuestas generadas por el computador no solo requieren el conocimiento de datos almacenados sino que también ponen de manifiesto una cierta habilidad para establecer relaciones que solo parece posible si se comprende correctamente el lenguaje. Esto es particularmente evidente en el tercer ejemplo. El IBM Watson tubo un coste de alrededor de 3 millones de dólares y estaba constituido por 90 servidores IBM Power750, con una memoria total de 16 Terabytes (16 billones de Bytes). Esos servidores contenían un total de 2880 núcleos de procesamiento (*cores* en inglés) POWER 7, con una frecuencia de reloj de 3.55 GHz, y utilizaban una red de 10 Gbps (capaz de comunicar diez mil millones de bits por segundo). El código que ejecutaba constaba de unas 1.000.000 líneas de código, unas 700.000 escritas en el lenguaje Java y 300.000 en C++. El primer computador del TOP500 en 2011 tenía

una velocidad de 10510.0 TFLOPS y el último de la lista proporcionaba 50.9 TFLOPS, el IBM Watson alcanzaba los 80 TFLOPS (80 billones de operaciones en coma flotante por segundo). Esta línea de trabajo, que aprovecha las capacidades cada vez mayores de los computadores para dotarles de capacidades que se habían considerado exclusivas de la mente humana, está dando lugar a programas en ámbitos muy diversos que muestran niveles de inteligencia comparables, o incluso superiores, a los del ser humano. Así por ejemplo recientemente se han mostrado programas capaces de jugar muy bien al póquer, juego en el que interviene el azar y hay que ser capaz de gestionar eficientemente la falta de información (www.sciencemag.org/content/347/6218/145.full), y también robots capaces de expresar sentimientos (www.youtube.com/watch?v=Ka5diu53aQE).

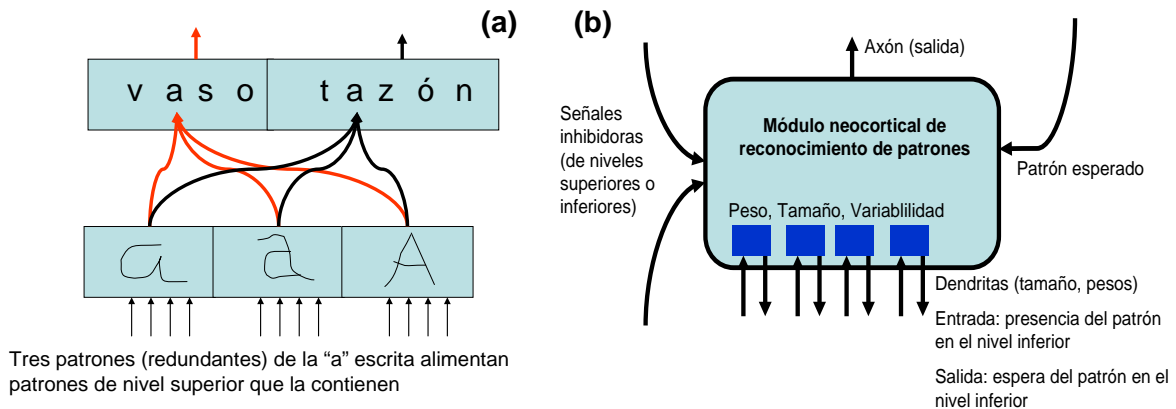


Figura 3. Modelo de sistema jerárquico de reconocedores de patrones: (a) Conexión de módulos reconocedores de patrones; (b) entradas y salidas de un módulo.

Ciertos resultados obtenidos recientemente por diversos neurocientíficos [3] están corroborando un modelo para el córtex cerebral según el cual las neuronas se organizan a partir de una serie de columnas corticales que constituyen módulos reconocedores de patrones. En la Figura 3.a se muestra un esquema simplificado de la interconexión entre distintos módulos, que se activarían al reconocer alguna de las alternativas para la letra "a" escrita, y que actuarían sobre módulos que incluyen dicha letra. Los módulos que se asocian a palabras en el ejemplo de la Figura 3.a se activarían cuando reciban un nivel suficiente de activaciones de las letras que los constituyen. La Figura 3.a también pone de manifiesto la redundancia de módulos en el cerebro (pueden existir varios módulos para un mismo patrón). La Figura 3.b muestra las entradas y salidas de uno de los módulos reconocedores. Estos módulos no solo reciben señales de los niveles inferiores (como se muestra en el ejemplo de la Figura 3.b) sino que también pueden recibir señales de niveles superiores que permiten activar un determinado reconocedor de patrones si en los niveles superiores ya se ha alcanzado un nivel de activación para un patrón que lo incluye. Así, por ejemplo, podría alcanzarse un nivel de activación suficiente para "vaso", si se han activado "v", "s", y "o", aunque no se haya activado una "a", si no se contemplan palabras alternativas por en el contexto (entrarían en juego reconocedores de patrones de nivel superior). Cada uno de los módulos reconocedores de patrones está constituido por alrededor de 100 neuronas, y por lo tanto en el córtex cerebral existirán del orden de 3×10^8 reconocedores de patrones (como se ha dicho anteriormente hay alrededor de 3×10^{10} neuronas en el córtex). Se estima que para simular el comportamiento del cerebro sería necesario realizar alrededor de 300.000 cálculos/s por cada reconocedor [2], lo que significa 9×10^{13} de cálculos por segundo. Es decir, alrededor de 100 billones (10^{14}) de cálculos/s. En cuanto a las necesidades de memoria, se estima que habría que utilizar unos 4 Bytes por entrada y en cada reconocedor habría un promedio de unas 8 entradas. También habría que sumar más un Byte más por entrada correspondiente al peso y 32 Bytes para conexiones *descendentes*. Esto haría un total de 72 Bytes por reconocedor, que teniendo en cuenta los 3×10^8 reconocedores, implica que la descripción del estado del córtex en términos de reconocedores de patrones precisa alrededor de 2×10^{10} Bytes, o sea, unos 20 GBytes. Los supercomputadores actuales superan las 10^{15}

operaciones/s, y tienen memoria con un tamaño bastante superior a los 20 GBytes. Por lo tanto, simular una mente a partir de un córtex basado en un sistema de reconocedores de patrones se encuentra dentro de la capacidad de los supercomputadores actuales.

En cualquier caso, como hemos indicado más arriba, un conocimiento completo del cerebro humano implica estudiar distintos niveles de descripción, y alcanzar un conocimiento de cada uno de ellos que permita generar modelos adecuados para entender los niveles superiores. En [4] analizan con detalle las cuestiones relacionadas con las escalas y los niveles de descripción necesarios para emular un cerebro humano (incluyendo la posibilidad de emular el cerebro específico de una persona con sus neuronas y conexiones propias). Parece improbable que un sistema complejo de tantos grados de libertad como el cerebro pueda modelarse correctamente a un nivel, encapsulando lo que ocurre en sus elementos. De hecho se ha observado que la estimulación de neuronas individuales puede afectar a las respuestas sensoriales y ocasionan modificaciones en los niveles superiores del comportamiento. En [5] se pone de manifiesto que en los sistemas cuya dinámica abarca escalas de tamaño, tiempo, espacio, etc., de diferentes órdenes de magnitud, es posible que exista un comportamiento “desacoplado” entre lo que sucede en escalas diferentes, y por tanto se puedan llevar a cabo simulaciones con elementos incluidos en un único nivel de descripción. Sin embargo, también es posible que no exista un total “desacoplamiento” de escalas (como por ejemplo ocurriría en las turbulencias de un fluido), y no sean posibles predicciones a una escala que no tengas en cuenta lo que ocurren en las escalas inferiores. La cuestión es, por tanto, determinar la situación que se observa en el cerebro humano: si en él existe algún nivel a partir del cual se pueda entender el comportamiento consciente e inteligente y cuál es dicho nivel. Todavía queda mucho que hacer en prácticamente todos los niveles, y de hecho es preciso desarrollar modelos lo suficientemente detallados de neuronas, y de estructuras de neuronas.

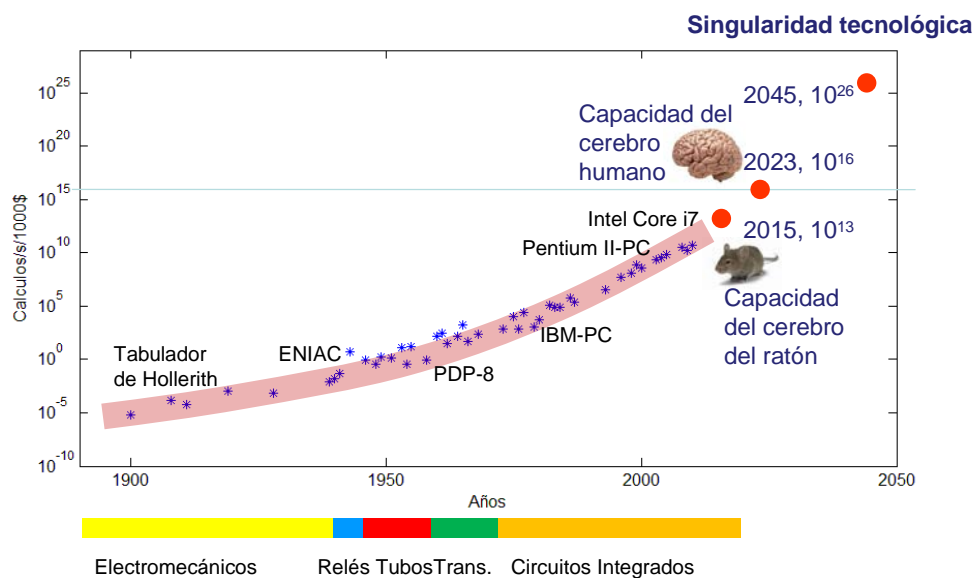


Figura 4. Evolución de la capacidad de procesamiento por dólar y predicciones sobre la capacidad de simulación de cerebros.

Sin embargo, hay que tener en cuenta que las capacidades de cómputo necesarias para simular los niveles más bajos (nivel de redes de neuronas, nivel electro-fisiológico, nivel de metaboloma, nivel de proteoma, etc.) son cada vez mayores, y el carácter de dichas simulaciones también cambia de forma que mientras que los niveles más bajos utilizan cantidades masivas de información sencilla (tipos de moléculas y sus posiciones, por ejemplo), los niveles más elevados usan estructuras de información mucho más complejas. De esta forma, la descripción de una región del cerebro, por ejemplo un módulo reconocedor de patrones del córtex que detecte un rostro, podría ser más sencilla que la de la dinámica de los

neurotransmisores y demás procesos que tienen lugar en una neurona. Así, la simulación detallada de una neurona que se realizó en 2005 necesitó entre 10^9 y 10^{10} FLOPS (entre mil y diez mil millones de operaciones en coma flotante por segundo) y entre 106 y 107 Bytes (uno y diez millones de bytes). Para este año 2015 se considera que se podrá simular el cerebro de una rata utilizando modelos detallados de neuronas. Serán necesarias velocidades del orden del PetaFLOPS (10^{15} FLOPS, es decir mil billones de operaciones en coma flotante por segundo) y entre diez y cien billones de bytes (10^{13} – 10^{14} Bytes). La simulación completa del número de neuronas de un cerebro humano necesitaría entre 10^{18} y 10^{19} FLOPS, y 10^{17} bytes de memoria. Es decir, se trataría de computadores con potencia de cálculo del orden del ExaFLOPS (10^{18} operaciones en coma flotante por segundo, o un trillón de operaciones en coma flotante por segundo). Precisamente este es el objetivo de la investigación actual en ingeniería de computadores. Algunas predicciones sitúan la posibilidad de simular el cerebro humano completo en el año 2023. En la Figura 4 se ilustra la evolución de la capacidad de cómputo por cada dólar junto con las predicciones del momento en que se alcanzarán las capacidades necesarias para simular el cerebro humano.

Evidentemente, el conocimiento de la mente humana acarrea unos beneficios importantísimos. Por ejemplo en cuanto a la cura de enfermedades mentales se podría avanzar en los diagnósticos basados en las causas biológicas en lugar de en síntomas (generando nuevos tratamientos, nuevos procedimientos para la detección precoz, etc.). También se podría avanzar en la teoría de la mente (con el desarrollo una visión multinivel del cerebro) que mejoraría el conocimiento sobre el lenguaje, las emociones, el control corporal, y la cadena de eventos desde los genes a la mente. Y por supuesto también se avanzaría en el desarrollo de supercomputadores y software interactivo para supercomputación multiescala, en la gestión masiva de datos, y en el desarrollo de computadores inteligentes (*human-like intelligence*). Así se reconoce tanto en la UE como en los EE.UU. donde se están aportando cantidades importantes para financiar proyectos relacionados con la investigación en el cerebro humano. Así tenemos el “Human Brain Project” en la UE, financiado con 1000 millones de euros, y la “BRAIN Initiative” financiada con 300 millones de dólares por año durante 10 años.

Conclusión

No se puede concluir una conferencia en una Facultad de Educación, Economía y Tecnología sin reflexionar sobre las consecuencias sociales que se derivarían de computadores con capacidades equiparables a la inteligencia humana.

En 1983, Vernor Vinge introduce el término *singularidad tecnológica* para designar el momento en el que la tecnología evolucionaría más rápidamente que la capacidad de comprensión de la misma por parte de los humanos, debido a que los sistemas de cómputo (cognitivos) serían “más inteligentes”. Por otro lado, hay previsiones de que los computadores piensen por sí mismos (pasen el test de Turing) en 2029, y de que, tal y como se ilustra en la Figura 4, dicha singularidad tecnológica se alcance alrededor de 2045 (Ray Kurzweil, director de ingeniería en Google desde 2012). Por otra parte, el cuarto paradigma de la ciencia (tras los paradigmas de ciencia teórica, experimental, y computacional) pone de manifiesto que la actividad científica no será posible sin el uso de ordenadores debido a la gran cantidad de datos que se están generando y se manifiesta cada vez con más fuerza a través de las aplicaciones que van poniéndose en marcha dentro del denominado Big Data. Precisamente este cuarto paradigma es considerado por muchos como una evidencia de la evolución hacia la singularidad tecnológica. Teniendo en cuenta la situación actual, la singularidad tecnológica todavía puede parecer más propia de la ficción que de la ciencia, pero si recordamos las tres leyes del futuro de Arthur C. Clarke,

1. Cuando un científico viejo y distinguido afirma que algo es posible, es casi seguro que acierte, y cuando afirma que algo es imposible, es muy probable que esté equivocado.
2. La única manera de descubrir los límites de lo posible es aventurarse más allá de ellos.
3. Cualquier tecnología suficientemente avanzada no se diferencia de la magia

no podemos estar seguros de que no se darán situaciones que se deriven de unas tecnologías cada vez más próximas a la singularidad tecnológica. Tenemos que tener en cuenta que, si bien los retos tecnológicos son importantes, los retos sociales que se plantearían serían todavía mayores. No obstante, dado que, de alguna manera, no serían más que una manifestación más intensa de las situaciones que ya pueden observar hoy en día como resultado de la implantación de las tecnologías de la información y las comunicaciones, se deberían de considerar muy seriamente. Así, habría que tener en cuenta muy seriamente los aspectos relativos a los empleos que se requerirían en un futuro (<http://www.pewinternet.org/2014/08/06/future-of-jobs/>), donde gran parte de las tareas (incluso tareas muy especializadas como diagnósticos médicos por ejemplo) podrán ser realizados por los computadores. Es cierto que la sociedad siempre se ha adaptado creando nuevos tipos de trabajos y que siempre ha existido una controversia sobre las consecuencias sociales de los avances tecnológicos (<http://www.hipeac.net/system/files/hipeac-vision-2015-draft.pdf>), pero no son muy alentadores algunos datos sobre la evolución del mercado laboral. Por ejemplo, los trabajos creados tras la recesión en USA no solo han sido menos que los que se perdieron sino que están peor pagados.

En este contexto, la formación que se debería procurar a las nuevas generaciones para afrontar ese futuro es algo que debería ser objeto de una reflexión muy cuidadosa, para que sea la sociedad la que dirija la evolución tecnológica hacia el bienestar de todos sus miembros. Muchas gracias por la atención.

Referencias

- [1] J. Ortega, M. Anguita, A. Prieto: “Arquitectura de Computadores”. Ed. Thomson-Paraninfo, 2005.
- [2] R. Kurzweil: “How to create a mind: the secret of human thought revealed”. Lola Books GbR, Berlin, 2013.
- [3] V.B. Mountcastle: “An organizing principle for cerebral function: the unit model and the distributed system”. (1978) En M. Edelman, V.B. Mountcastle (Ed.), *The mindful brain: cortical organization and the group-selective theory of higher brain function*, Cambridge, MIT Press, 1982
- [4] A. Sandberg, N. Bostrom: “Whole brain emulation: a roadmap”. Technical Report #2008-3, Future of Humanity Institute, Oxford University, 2008. www.fhi.ox.ac.uk/2008-3.pdf
- [5] R. Hillerbrand: “Scale separation as a condition for quantitative modelling. Why mathematics works for some problems and fails for others”. Synthese, 2008.