

Tendencias del procesamiento computacional: bioinformática y cómputo ubicuo

Trends in computational processing: bioinformatics and ubiquitous computing

LUIS CARLOS GONZÁLEZ-GURROLA^{1,2} Y FERNANDO MARTÍNEZ-REYES¹

Resumen

Desde sus inicios, la computación ha sido una herramienta eficaz y popular para la resolución de problemas de distinta índole y niveles de complejidad. Los usos que la computadora ha tenido van desde operación de editores de texto y hojas de cálculo hasta procesamiento intensivo a través de computadoras interconectadas que buscan dar solución a problemas con un gran número de variables. Las tendencias del procesamiento computacional han estado ligadas a los avances en diversas áreas de la ciencia, creando una simbiosis que motiva diseños más eficientes de cómputo y a la vez hace posible extender los límites del conocimiento. Dos áreas que han explotado el procesamiento computacional buscando elevar la calidad de vida de la sociedad son la bioinformática y el cómputo ubicuo. Las aportaciones de cada una de estas líneas son evidentes, aún así creemos que lo mejor de ambas está por venir.

Palabras clave: procesamiento de información, biocomputación, cómputo consciente del contexto, interacción humano-máquina.

Abstract

Since its beginnings, the computing has been an efficient and popular tool for solving problems of different kinds and levels of complexity. The uses of computers range from text editors and spreadsheet programs to intensive processing via computer clusters that seek to tackle problems with a big number of variables. The trends in the computational processing have been linked to advances in the sciences, creating a symbiosis that motivates more efficient computing designs as well as to increase the knowledge. Two main areas that have exploited the computational processing to achieve a better life quality in our society are the bioinformatics and ubiquitous computing. The contributions made by each of these areas are evident, even though the best of both is yet to come.

Keywords: information processing, computational biology, context awareness, human-computer interaction.

Introducción

El poder de procesamiento de las computadoras ha abierto la posibilidad de abordar nuevos problemas que antes no se alcanzaban a visualizar o incluso eran pospuestos por la falta de técnicas y procedimientos adecuados para lograr su solución. Hoy sabemos que es raro aquel descubrimiento científico que no conlleve el uso de una computadora. Y es precisamente esta la justificación del uso del procesamiento computacional: extender nuestros alcances, validar hipótesis, realizar experimentos, ejecutar simulaciones y muchas otras tareas que nos permiten extender los horizontes de nuestro entendimiento. Conforme este conocimiento se hace más palpable, nuevas áreas de aplicación inmediata surgen, este es el caso de la bioinformática y el cómputo ubicuo.

¹ Universidad Autónoma de Chihuahua, Facultad de Ingeniería. Campus II, Chihuahua, Chih. México. Tel: (614) 442 9500.

² Dirección electrónica del autor de correspondencia: lgonzalez@uach.mx.

A finales de la década de 1990, una de las grandes esperanzas de la humanidad se centraba en la decodificación del genoma humano, es decir, desentrañar los misterios mismos de nuestro organismo. Muy pocos alcanzaron a visualizar que ese logro no concluía la búsqueda, sino que apenas la iniciaba. Hoy, a 10 años de la publicación del genoma, mucha información se ha generado, pero también muchas preguntas siguen en el aire. Este es precisamente uno de los objetivos de la bioinformática, darle sentido a los enormes bancos de datos de experimentos biológicos que se generan día con día.

Durante esa misma década, Mark Weiser presentó al mundo un nuevo concepto que exitosamente describía la forma en la que la computadora empezaba a cambiar la forma en como vivíamos. La tecnología invisible nos rodea y muchas veces está "atenta" a nuestras necesidades. Esta es una de las premisas del cómputo ubicuo. Con dispositivos cada vez más sensibles, pequeños y sustentables, la tecnología nos facilita la vida. El abaratamiento de sensores, teléfonos celulares, tabletas y computadoras ha incrementado el número de usuarios de servicios relacionados a estos dispositivos. Esto requiere un manejo eficiente del procesamiento computacional donde los servicios de cómputo deben estar siempre presentes y disponibles.

Estas dos áreas tecnológicas requieren, pues, de un procesamiento computacional eficiente. El objetivo principal de este artículo es realizar una introducción de estas tendencias del procesamiento computacional, presentaremos hechos significativos, retos y problemas actuales, buscando en todo momento impulsar el desarrollo de trabajo en estas áreas.

Bioinformática

Padecimientos comunes en la población mexicana como diabetes mellitus o hipertensión arterial pueden ser abordados desde el análisis de la predisposición genética. Existe una necesidad genuina de contar con mecanismos eficaces que permitan, primero, un conocimiento detallado de nuestra carga genética como población (codificada en nuestro genoma), para después desarrollar e implementar estrategias y tecnología que hagan posible interpretar

estos datos. La adquisición de este conocimiento permitirá tomar medidas preventivas y correctivas para elevar la calidad de nuestro nivel de vida, siendo ésta precisamente una de las acepciones de la medicina genómica. Los primeros pasos para conocer el "genoma mexicano" ya se han dado (Silva-Zolezzi *et al.*, 2009), sin embargo, es necesario completar esta información y lograr su análisis e interpretación.

Los centros nacionales de salud de Estados Unidos definen la bioinformática como la investigación, desarrollo o aplicación de herramientas computacionales para ampliar el uso de información biológica, médica, conductual y de salud. Desde la digitalización de las primeras secuencias de ADN en 1977, se ha registrado un crecimiento constante en la información biológica almacenada en repositorios digitales. Para tener una idea de la magnitud de información nueva, basta observar que un solo secuenciador de ADN de última generación es capaz de producir más de 20 gigabytes de información por semana. Podemos observar pues, que generar datos biológicos es cada vez más fácil y barato, el reto se encuentra en analizar e interpretar estos datos con el objetivo de transformarlos en información útil cuya aplicación incida en la salud pública.

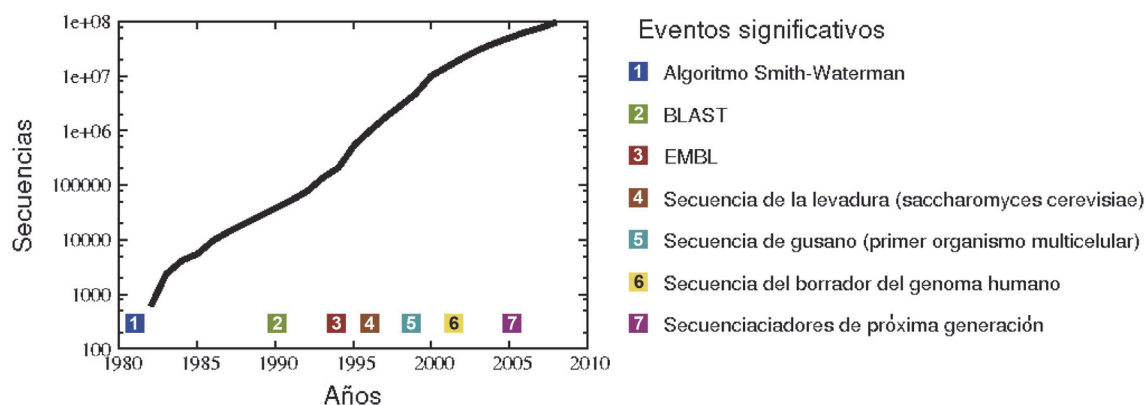
Una de las principales bases de datos (BD) de secuencias de ácidos nucleicos en el mundo es el GenBank; su crecimiento es impresionante, de hecho, el número de secuencias digitalizadas ha crecido exponencialmente. La Figura 1 muestra el acumulado de secuencias de ADN almacenadas en un periodo de 25 años. Esta BD duplica su capacidad cada 1.4 años, incluso superando la tasa de crecimiento de los transistores en circuitos integrados descritos por la popular ley de Moore (un procesador duplica su número de transistores cada dos años). Mucha de la información que almacena esta BD todavía se encuentra en espera de ser analizada, esto nos indica que aún estamos lejos de mantenernos al día respecto al estudio y entendimiento de nueva información.

Uno de los hitos más populares de la bioinformática (ver el apartado de Eventos significativos en la Figura 1) ha sido la consecución del primer borrador del genoma humano, publicado simultáneamente por las revistas Science y Nature en el año 2003. Este logro, que inicialmente provocó

grandes expectativas sobre el advenimiento de una nueva era de medicina personalizada basada en el genoma, dio inicio a una etapa de trabajo intenso y desarrollo de procedimientos con el firme objetivo de darle interpretación a la secuencia de 3,000,000,000 de letras que forman nuestro genoma (cada letra representa una base, A para adenina, C para citosina, T para timina y G para guanina). Hoy en día, ya

computadora de escritorio con características similares a las requeridas en la instalación de cualquier procesador de textos. Proyectos de reingeniería de proteínas (Frey *et al.*, 2010) hacen factible el diseño de fármacos que se adapten a las mutaciones en bacterias, dejando atrás protocolos médicos clásicos, y atacando con una sola dosis de medicamento la bacteria y sus mecanismos de supervivencia.

Figura 1. Acumulado del número de secuencias de ADN disponibles en el GenBank en los últimos 25 años. También se muestran algunos eventos significativos en el campo de la bioinformática.



disponemos de un mayor número de genomas, tanto de animales como de plantas (recientemente se conoció la secuencia del genoma del frijol), pero es claro que esto es sólo el principio de una avalancha de preguntas de investigación, retos, hipótesis e intuiciones que se empiezan a generar respecto a toda esta información genómica.

El campo de la bioinformática ha permitido darle un nuevo enfoque a algoritmos tradicionales, así como motivar el diseño de nuevos procedimientos. Uno de los algoritmos más representativos de esta área es el *Smith-Waterman*, que permite comparar un par de secuencias e identificar las regiones de mayor similitud entre éstas. Otra herramienta que ha sido de gran ayuda para encontrar relaciones de homología entre grupos de secuencias es *BLAST*, llegando a consolidarse como la herramienta de mayor uso entre la comunidad bioinformática. Los logros en materia algorítmica de hace un par de años a la fecha han sido sobresalientes. El programa *Bowtie* (Langmed *et al.*, 2009) permite alinear el genoma humano completo a otro de referencia, utilizando sólo una

Con esta explosión de información, nuevos y fascinantes retos se hacen presentes: reducción del error en la secuenciación de ADN, secuenciar y caracterizar genomas de comunidades enteras (ecosistemas, metagenómica), identificación de la estructura tridimensional del ARN, entender el rol que juega el ambiente en las diferencias a nivel fenotípico (epigenética), identificación de la estructura tridimensional de las proteínas, extracción de patrones y análisis estadístico de grupos de secuencias, entendimiento del procedimiento de plegado de proteínas (dinámica molecular), reingeniería de proteínas, reducción de incertidumbre y ensamble de árboles filogenéticos y medicina genómica, son parte de un pequeño grupo de tareas que todavía quedan pendientes en esta área.

La omnipresencia de la computación

Si volteamos a nuestro alrededor seguramente observaremos la existencia, además de nuestra computadora de escritorio, iPad y teléfono inteligente, de una variedad de dispositivos, objetos y artefactos que cuentan con algún tipo de computadora integrada.

Existe tecnología computacional presente en el refrigerador, la estufa, el horno de microondas, la caja digital de televisión, la televisión, las consolas de video juegos, los reclinables con masaje, las alarmas para el hogar, el sistema de calefacción, entre otros dispositivos. Varios de estos artefactos ya se conectan a la Internet, pero más aún, dentro de poco tiempo veremos a estos artefactos interconectados e intercambiando información para sugerir recetas de cocina basados, por ejemplo, en la disponibilidad de la despensa. Pero, ¿qué podríamos esperar de la computación en el futuro?

La computación ubicua (Weiser, 1999), computación del siglo XXI, considera la integración de redes de computadoras de diferentes tamaños y capacidades de cómputo a través de las cuales se identifican las actividades del ser humano, se aprende de ellas y se pone a su disposición información o servicios en el espacio físico en el cual desarrollan sus actividades, como se aprecia en el siguiente escenario de aplicación:

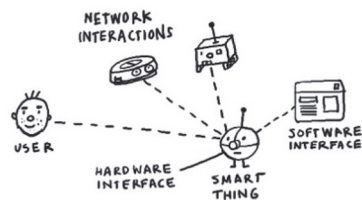
Una persona llega al restaurant y en su dispositivo móvil recibe la información de la mesa que se le ha reservado. La 'mesa digital interactiva' inicia su diálogo con el cliente, sugiriendo el platillo que el cliente acostumbra degustar en esa temporada y a esa hora del día, además de las recomendaciones del chef. En el menú digital también se informa acerca de los tiempos de espera para cada platillo de acuerdo con la cantidad de clientes presentes. Mientras espera a que el platillo esté listo, el cliente, a través de la

mesa digital, navega por diferentes niveles de información relacionados con el origen del restaurante y opiniones de los clientes acerca de diferentes platillos. Puede leer alguna revista o viajar hasta los orígenes de la receta, revisar el correo, utilizar la banca electrónica, descargar un documento desde la computadora de su oficina o incluso utilizar alguna opción de juego de mesa. Mientras degusta el platillo, la mesa solicita permiso para poder reproducir la música que va acorde con la ocasión. Al finalizar, el cliente coloca la tarjeta de crédito sobre la mesa, verifica el recibo con los consumos, agrega propina y selecciona si desea su recibo impreso o desea enviar la copia electrónica a su correo.

Para implementar este escenario de cómputo ubicuo se requiere contar con redes de objetos computacionales (WSN – Wireless Sensor Networks), Figura 2a, y redes inalámbricas portadas por el humano (WBSN – Wireless Body Sensor Networks), Figura 2b. Estas redes comparten información con redes de área local y de área amplia (LAN y WAN) para ofrecer información o servicios al usuario. La computación ubicua considera la implementación de ambientes inteligentes, en los cuales la interacción del usuario con la computadora se da a través de eventos biológicos o a través de voz e incluso gestos; la realidad virtual y aumentada son comunes en estos espacios sociales. En el escenario de ejemplo, estas redes intercambiaron información para identificar el estado de ánimo de la persona y con ello reproducir la música adecuada.

Figura 2. Dos componentes tecnológicos en la computación ubicua a) redes de objetos «inteligentes» y computadoras (WSN) y b) redes del cuerpo humano (WBSN)

ANATOMY of a SMART THING



"Objects with limited processing power can exhibit interesting behaviors with just a little networking"
~ @mikekunavsky

a) (<http://communicationnation.blogspot.com/>)



b) (<http://blog.memisc.com/2010/11/wireless-body-sensor-networks.html>)

La computación ubicua promueve la generación de ambientes inteligentes en los cuales se brinda soporte de forma confiable y segura. Recopilar y procesar información, e «inteligentemente» anticiparse a las necesidades del ser humano no es trivial. Se puede automatizar el encendido de la luz del cuarto cuando existe cierto nivel de oscuridad, pero controlar el encendido de la luz en relación al nivel de cansancio de la vista de una persona es mucho más complejo. Actualmente, la tecnología funciona para automatizar procesos repetitivos, pero a pesar de que el ser humano utiliza «rutinas» diarias (Crabtree & Rodden, 2004) estas pueden ser alteradas en cualquier momento, lo que podría afectar la calidad de servicio ofrecida por la computadora. Por otro lado, interactuar con una gran cantidad de artefactos y redes computacionales implica dejar rastros de datos privados en varias computadoras. La computación tiene un reto importante en este sentido.


Es posible identificar que la computación ubicua tiene retos tecnológicos y sociales por atender (Edwards & Grinter, 2001), sin embargo, también es posible observar algunos de los beneficios potenciales que la computación del siglo XXI podría ofrecer al ser humano. Considerando que en el futuro cercano seremos una población mundial en edad adulta, es posible que la tecnología del restaurante inteligente se implemente en otros espacios sociales, en los cuales la computadora apoya en las «rutinas» desarrolladas por adultos mayores, que monitorea y cuida de su salud y que, incluso, le ofrece un acompañamiento digital (Heerink *et al.*, 2006).

Conclusiones

La bioinformática y la computación ubicua son dos tendencias de la computación con grandes aportaciones científicas en los últimos años. Ambas disciplinas buscan generar conocimiento para

preservar o mejorar las condiciones de vida del ser humano. Resultados en bioinformática nos ayudan a entender a detalle el ciclo vital de un organismo, lo cual directamente se traduce en beneficios en la salud pública. La computación ubicua, por otro lado, identifica y está alerta de cambios emocionales y de salud de las personas con el objetivo de anticipar necesidades en seguridad y confort que las personas esperan del lugar en el cual desarrollan sus actividades. Un reto común para ambas disciplinas es el procesamiento de grandes volúmenes de información. Se requiere experimentar con técnicas y metodologías novedosas para poder encontrar patrones de comportamiento de células cancerígenas o para que un sistema computacional identifique los estados de ánimo de las personas. En el futuro cercano seguiremos observando la influencia de estas dos áreas de investigación en nuestras vidas.

Literatura citada

- SILVA-ZOLEZZI I, Hidalgo-Miranda A, Estrada-Gil J, Fernandez-Lopez JC, Uribe-Figueroa L, Contreras A, Balam-Ortiz E, del Bosque-Plata L, Velazquez-Fernandez D, Lara C, Goya R, Hernandez-Lemus E, Davila C, Barrientos E, March S and Jimenez-Sanchez G. 2009. Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico. *Proc. Natl. Acad. Sci. USA* 106:8611–8616.
- LANGMEAD, B., Trapnell C, Pop M and Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25.
- FREY, KM, Georgiev I, Donald BR and Anderson, A.C. 2010. Predicting resistance mutations using protein design algorithms. *Proc. Natl. Acad. Sci. USA* 107:13707-13712.
- WEISER, M. 2004. The computer for the 21st century. *Scientific American*, 94–104.
- CRABTREE, A., and Rodden T. 2004. Domestic routines and design for the home, *CSCW*, 13, 2, 191-220.
- EDWARDS, WK, and Grinter RE. 2001. At Home with Ubiquitous Computing: Seven Challenges, *Ubicomp: Ubiquitous Computing*, Lecture Notes in Computer Science, vol. 2201, Springer-Verlag, Berlin, pp. 256–272.
- HEERINK, M., Kröse B, Evers V and Wielinga B. 2006. Studying the acceptance of a robotic agent by elderly users. *International Journal of Assistive Robotics and Mechatronics*, 7(3): p. 33-43. 

Este artículo es citado así:

González-Gurrola, L. C., F. Martínez-Reyes. 2013: *Tendencias del procesamiento computacional: bioinformática y cómputo ubicuo*. *TECNOCIENCIA Chihuahua* 7(1): 1-6.

Resúmenes curriculares de autor y coautores

LUIS CARLOS GONZÁLEZ-GURROLA. Finaliza sus estudios en Ingeniería en Sistemas Computacionales en el Instituto Tecnológico de Durango en 2002. Obtiene el grado de Maestro en Ciencias de la Computación por parte del Centro de Investigación Científica y de Educación Superior de Ensenada en 2004. Fue catedrático de tiempo completo en el Instituto Tecnológico Superior de Santiago Papasquiaro (Durango) por 3 años. En 2011 finaliza sus estudios doctorales en el área de Tecnologías de la Información en la Universidad de Carolina del Norte en Charlotte, EE.UU. Actualmente se desempeña como profesor/investigador de la Facultad de Ingeniería de la Universidad Autónoma de Chihuahua. Entre sus intereses actuales se encuentran la bioinformática, problemas de optimización combinatoria y diseño de algoritmos.

FERNANDO MARTÍNEZ-REYES. Terminó su licenciatura en 1991, año en que le fue otorgado el título de Ingeniero industrial en electrónica por el Instituto Tecnológico de Veracruz (ITV). Posee el grado de Maestro en Ciencias en el área de Ingeniería Electrónica, otorgado por el Instituto Tecnológico de Chihuahua en 1997. Realizó sus estudios de Doctorado y obtuvo el grado de Doctor en Ciencias de la Computación en la Universidad de Nottingham, Inglaterra, en 2009. Desde 1994 labora en la Facultad de Ingeniería de la Universidad Autónoma de Chihuahua (UACH). Su área de especialización es la computación basada en localización. Las áreas de interés en investigación incluyen el cómputo móvil, interacción humano-computadora, computación consciente del contexto y cómputo ubicuo.