



Computadores moleculares: una tecnología prometedora

Jorge Eduardo Ortiz Triviño, Instructor Asociado, Departamento Ingeniería de Sistemas, Universidad Nacional de Colombia.
E-mail: jorgeo@ingenieria.ingsala.unal.edu.co

RESUMEN

En este artículo se explican algunos detalles básicos para la realización de computadores moleculares basados en ADN; describe el método general para realizar programas en un computador molecular. Se concluye, que la labor de construcción e implementación de esos computadores no es tarea exclusiva del Ingeniero de sistemas sino, y con una participación importante, de profesionales en genética, medicina, biología y ciencias relacionadas. De otro lado, aunque es una tecnología que se encuentra en su primera infancia, puede constituirse en una buena alternativa en los próximos años, si se tiene en cuenta que apoyaría en la solución de problemas, en distintas áreas, cuya solución mediante computación electrónica convencional pueda ser difícil, sino imposible.

INTRODUCCIÓN

Un computador es un dispositivo (natural y/o artificial) capaz de almacenar información y transformarla de alguna manera. En computación no todo está hecho. Los investigadores en esa área intentan construir nuevos dispositivos que permitan solucionar problemas extremadamente complejos, que los actuales computadores no han podido resolver, bien sea porque teóricamente

es imposible o bien porque, aunque tecnológicamente sea viable, los costos son tan altos que ni siquiera valga la pena intentarlo.

Sin duda, los computadores electrónicos de hoy permiten resolver un alto número de problemas, de diversa naturaleza, mediante el empleo de sofisticados dispositivos electrónicos, que admiten la implementación de modelos que pueden resolver distintos problemas de la realidad, y que han hecho avances significativos en muchas áreas de la actividad humana. Esos avances van desde juegos simples, programas procesadores de palabra, programas de dibujo, hasta sofisticados programas de comunicación al estilo Internet y aplicaciones complejas en distintas áreas como biología, genética, medicina, etc. por citar solamente algunas de ellas.

Con todo, sin embargo, existen una amplia variedad de problemas que aún siguen siendo irresolubles o cuya solución alcanza un nivel de complejidad que puede resultar inmanejable y/o de elevado costo la búsqueda de la solución. Cabe destacar, por ejemplo, aquellas situaciones que involucran un número elevado de variables o que intentan alcanzar comportamientos considerados inteligentes.

No obstante, este tipo de problemas,

en un número elevado de casos, ya han sido abordados y resueltos satisfactoriamente por la naturaleza a lo largo del proceso evolutivo. Lo anterior ha despertado la curiosidad en la comunidad científica, principalmente en el área de las ciencias de la computación, quienes intentan copiar y/o aprovechar los métodos de procesamiento biológico con el fin de resolver problemas complejos mediante artefactos naturales. En otras palabras: aprovechar la Ingeniería natural para el avance y desarrollo de la Ingeniería humana.

En este sentido algunos investigadores en ciencias de las computación están trabajando, básicamente, en dos direcciones. En primer lugar, la simulación de características de los sistemas naturales que son llevadas luego a computadores para que el procesamiento de la información se haga de un modo similar a como ocurre en la realidad; aquí se engloban técnicas como las redes neuronales, los autómatas celulares, la programación evolutiva y el hardware evolutivo, las cuales ya llevan varias décadas de trabajo.

Como segunda instancia, y es el caso que nos ocupa aquí, el empleo directo de estructuras biológicas en la recepción de señales o en el manejo de información; entre estas técnicas, mucho más recientes que las anteriores,

se encuentran los biochips y la computación con ADN.

En este sentido, se ha mostrado que ésta última es un camino hacia la producción de sistemas que permitan resolver diversos tipos de problemas mediante computación masivamente paralela. La ventaja de ello es evidente: de un lado, se pueden resolver problemas de naturaleza NP, se reduce la complejidad en tiempo y, de otro lado, los costos pueden llegar a ser bajos con relación a la computación electrónica convencional, específicamente es posible optimizar el almacenamiento de información, construcción del hardware, espacio y consumo de energía.

COMPUTACIÓN NATURAL

La naturaleza ha desarrollado sistemas orgánicos complejos dotados con mecanismos biológicos que ejecutan un sinnúmero de tareas eficientemente y se integran de una manera casi perfecta a su entorno. Uno de los elementos más representativos de esta clase de estructuras funcionales se tiene en los ácidos nucleicos: ácido desoxirribonucleico (ADN) y ácido ribonucleico (ARN). Estas dos biomoléculas están presentes en los núcleos de las células de todos los organismos vivientes.

Los ácidos nucleicos son polímeros largos no ramificados, consistentes en una serie de unidades mantenidas por enlaces fosfodiéster, responsables del almacenamiento, transporte, codificación y decodificación de la información genética.

El ADN está construido de subunidades llamadas nucleótidos constituidos por una parte constante, un azúcar de cinco carbonos llamado desoxirribosa y un grupo fosfato, unidos a una parte variable que puede ser una de cuatro bases nitrogenadas : Las purinas, Adenina (A) y Guanina (G), o las pirimidinas, Timina (T) y Citosina (C)

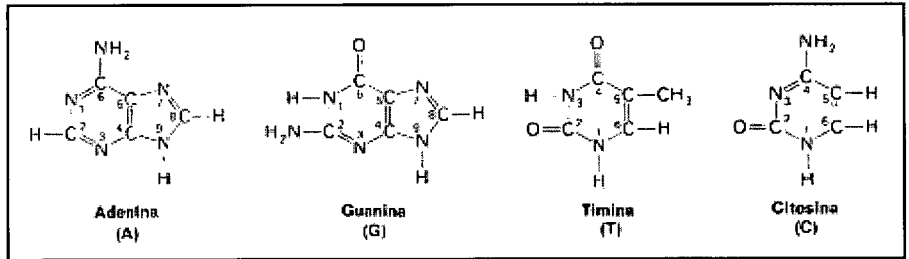


Figura 1. Bases nitrogenadas del ADN.

(ver figura 1).

La clave de los computadores moleculares está aquí: es posible almacenar información usando esos cuatro elementos variables (A,T,G,C), es más, esa información se puede manipular con técnicas de laboratorio. De esta manera, por ejemplo, es posible codificar nombres de ciudades. Específicamente, por convención, puede interpretarse la secuencia AGTTTC como Bogotá mientras que la secuencia GGTC CC como Medellín (la secuencia empleada para cada ciudad es arbitraria). Como se dijo en la introducción, para tener un computador, además del almacenamiento de información, debe poderse transformar. Este último aspecto se desarrolla en la sección que sigue.

OPERACIONES CON ADN

Uno de los aspectos más importantes en computación molecular, está en el hecho que la información almacenada en el ADN se puede manipular en el laboratorio. Esa serie de pasos necesarios para transformar la codificación del problema inicial en la respuesta que se desea se denomina programa molecular. Se enuncia a continuación la metodología que se puede emplear para construir programas moleculares, denominado Modelo No Restringido o Modelo ADN y que consiste en un conjunto de operaciones que permite realizar cómputo mediante ADN. Para ello se supone que se cuenta con una serie de tubos de ensayo en los cuales se pueden incluir cadenas de ADN.

1. Separación: Es el aislamiento de un

subconjunto de moléculas de ADN de una muestra particular incluida en un tubo de ensayo determinado. Por ejemplo es posible que se desee aislar todas las cadenas que contiene Timina.

2. Mezcla : Es la unión física del contenido de dos recipientes, como pueden ser dos tubos de ensayo.

3. Detección : Es la confirmación de la existencia o ausencia de un subconjunto de moléculas de ADN que cumplen alguna característica específica en el contenido de un recipiente.

4. Amplificar: Se trata de una operación que permite duplicar el contenido de un tubo de ensayo.

Las técnicas bioquímicas específicas empleadas para implementar cada una de las operaciones mencionadas es un aspecto que depende directamente de la manera como el problema se haya enfrentado. En la figura 2 se muestra, a grandes rasgos, el procedimiento bioquímico empleando las operaciones básicas mencionadas, que se refieren por los números encerrados en óvalos. Para una mayor información sobre el procedimiento se puede consultar la referencia (5)

A un nivel global, se pueden explorar métodos basados en enzimas de restricción y/o escogencia de información del ARN.

COMPUTACIÓN ELECTRÓNICA VS. MOLECULAR

La resolución de problemas complejos mediante el uso de computación

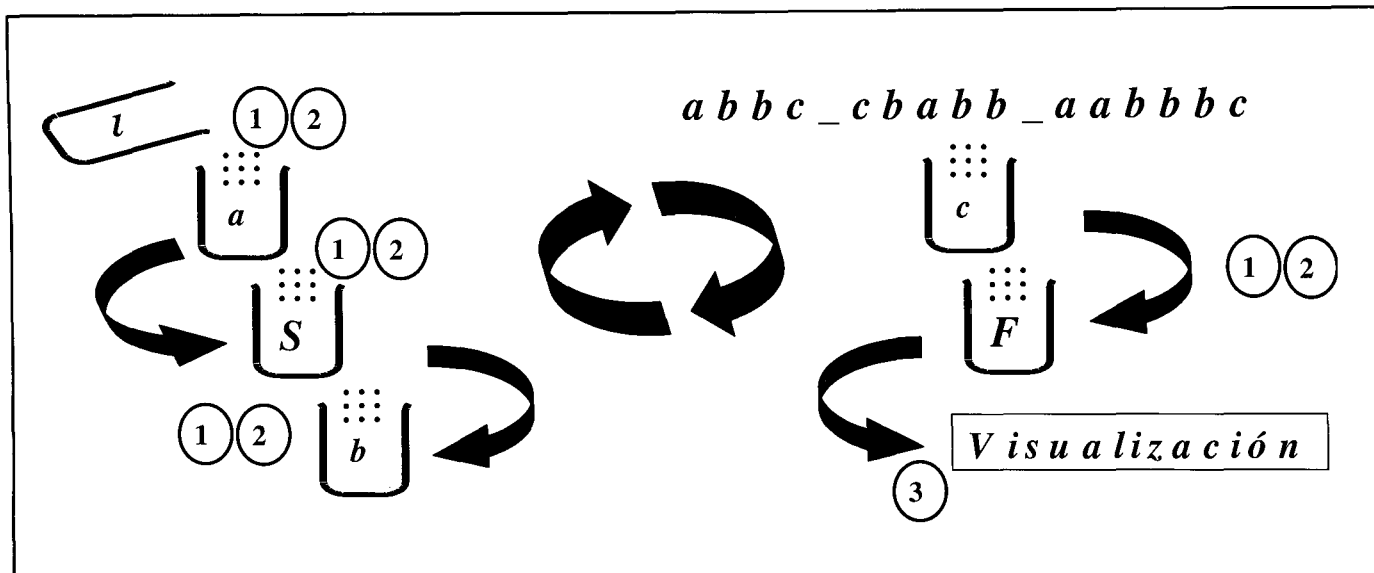


Figura 2. Método no restringido.

molecular es particularmente atractiva al comparar sus ventajas y debilidades frente a la computación electrónica tradicional. En lo que resta de esta sección se discuten algunos de estos factores:

Capacidad de codificación y almacenamiento: El genoma humano, por ejemplo, posee 2900 millones de pares de bases que representan cerca de 750 MB de información, de la cual solo cerca de un 30% codifica genes verdaderamente. En otras palabras, la información necesaria para "construir" un ser humano equivale a la capacidad de almacenamiento de un modesto disco duro convencional. Adicionalmente, la masa de esta información sólo llega al picogramo.

PROCESAMIENTO

Si tomamos cada uno de los 10 billones de células del cuerpo humano como un procesador que opera sobre los datos almacenados en el ADN, el nivel de paralelismo que resulta es realmente masivo. Con la tecnología actual, y algún esfuerzo adicional, se podría alcanzar, apenas, cientos de procesadores en un computador electrónico.

Confiabilidad: La tasa de error durante la replicación del ADN es (in vivo) de

uno entre 10000 millones, que lo hace un proceso prácticamente determinístico; es decir la probabilidad de error es muy cercana a cero. Además, existen controles de fallo a nivel superior sobre la información del ADN ya decodificada.

Consumo y manejo de energía: Es tan óptimo, en comparación con la tecnología electrónica actual, que, como observa Henry Backer, de Thimble Systems, California, si la replicación del ADN desperdiciara tanto calor por bit como una CPU moderna "haría que un pollito en desarrollo se cocinara dentro de su propio cascarón".

Contaminación ambiental: al emplear procesos únicamente naturales, fruto de millones de años de evolución y, por lo tanto, completamente integrados al medio ambiente, se asegura que el nivel de polución generado por un sistema molecular sea prácticamente nulo, al contrario de lo que ocurre con los dispositivos artificiales.

Lo anterior hace evidente que en cuanto a capacidades, el ADN presenta todas aquellas que se desean para un sistema de manejo de información y a un nivel muy superior a cualquiera que se haya creado hasta ahora.

DISCUSIÓN Y CONCLUSIONES

La implementación de herramientas de cómputo basadas en el ADN se puede llevar a cabo en nuestro país ; tales aplicaciones cubren un rango de posibilidades desde el diseño de autómatas finitos 4 y 5 hasta la construcción de redes de autómatas adaptativos, pasando por máquinas de Turing 3, autómatas celulares y redes neuronales. Otras aplicaciones, en ámbitos completamente distintos, como memorias asociativas o descryptadores, se han llevado a cabo.

Por supuesto, como toda tecnología incipiente, existen puntos débiles que deben ser tomados en consideración. Un obstáculo inicial es el costo de la ejecución de los procedimientos de laboratorio y de ingeniería. También debe tenerse en cuenta que en ambiente de laboratorio parte de las cualidades naturales del ADN pueden verse afectadas.

Debido a lo anterior, es importante que se adelanten trabajos tendientes a lograr la optimización en el manejo del ADN, ya sea con el mejoramiento de las técnicas existentes o con la utilización de procesos naturales basados en

estructuras biológicas.

De otra parte, la humanidad avanza hacia la era de la biología; en consecuencia, todas las áreas del saber, en particular la ingeniería de sistemas, se verán influidas por esta disciplina. Los computadores, por ejemplo, tendrán componentes vivos. Por lo tanto, es ne-

cesario en los actuales planes de estudio incluir asignaturas que puedan absorber el reto que se presenta para las próximas décadas.

AGRADECIMIENTOS

El autor desea agradecer a las siguientes personas quienes de una u otra for-

ma colaboraron en la realización de este artículo. En particular a los profesores Luz Gloria Torres del Departamento de Matemáticas y Estadística de la Facultad de Ciencias y J.J. Martínez del Departamento de Ingeniería de Sistemas de la Universidad Nacional de Colombia.

REFERENCIAS BIBLIOGRAFICAS.

1. **Stryer Lubert.** Bioquímica. W.H. Freeman and Co. 1988.
2. **Adleman L.** Molecular computation of solutions to combinatorial problems. *Science* 1994; 266,1021-1024.
3. **Alfonseca M, Sando J, Martínez M.** Teoría de lenguajes, gramáticas y autómatas. Ediciones Universidad y cultura. Madrid, España, 1987.
4. **Baum E, Lipton R.** (Editores). DNA Based Computers. American Mathematical Society. E.E.U.U., 1996.
5. **Cardona C, Ortíz J, Rojas C.** Diseño de un reconocedor universal de lenguajes regulares mediante computación con ADN. Memorias II Congreso Colombiano de Neurocomputación. Universidad Nacional de Colombia. Santafé de Bogotá, Noviembre de 1997.
6. **Cardona C, Rojas C.** Implementación de un autómata finito mediante computación con ADN. Memorias VII Congreso de Estudiantes de Ingeniería de Sistemas. Universidad EAFIT. Medellín, Marzo de 1998.
7. **Brookshera G.** Teoría de la computación, lenguajes formales, autómatas y complejidad. Addison-Wesley. E.E.U.U., 1993.
8. **Devlin T.** Bioquímica. Editorial Reverté. España, 1992.
9. **Grimaldi R.** Matemáticas, discreta y combinatoria. Addison-Wesley Iberoamericana. Washington, E.U.A., 1989.
10. **Langton C.** (editor). Artificial Life. MIT Press. Massachusetts, U.S.A., 1995.
11. **Lehninger A. et al.** Bioquímica. Worth Publishers Inc. 1990
12. **Murray, et al.** Bioquímica de Harper. Editorial El Manual Moderno. Ciudad de México, México, 1992.
13. **Vitaliano F.** La próxima marea en computación. En Status, Octubre 1994. Págs 12-16.