

Los algoritmos genéticos y los métodos de Monte Carlo

Algunas de las nuevas formas de cálculo para la medicina moderna permiten diseñar programas para la detección automática de lesiones en sistemas de diagnóstico computarizado.

Arnulfo Martínez Dávalos

En 1917 un matemático alemán prácticamente desconocido, J. Radon, publicó en una oscura revista especializada un artículo titulado “Sobre la determinación de funciones a partir de sus integrales a lo largo de ciertas variedades diferenciables” (Radon, 1917). En él describió de manera detallada la teoría matemática de la reconstrucción de imágenes tomográficas (imágenes de cortes del interior del cuerpo) a partir de sus proyecciones. En 1972 G. Hounsfield dio a conocer el primer tomógrafo de rayos X para uso clínico, trabajo por el cual recibiría en 1979 el Premio Nobel de medicina. El trabajo de Hounsfield detonó el desarrollo de otras modalidades de obtención de imágenes tomográficas (tomografía por emisión de positrones, imagenología por resonancia magnética), las cuales han revolu-

cionado el diagnóstico médico por imagen.

Los más de 50 años que tuvieron que pasar entre el trabajo de Radon y el desarrollo tecnológico de Hounsfield se debieron esencialmente a la falta de un elemento fundamental y absolutamente indispensable para llevar a la práctica las complejas operaciones matemáticas involucradas en el proceso

de reconstrucción tomográfica: la computadora. Hoy no deja de asombrarnos el avance abrumador del poder de cómputo puesto a nuestra disposición: computadoras de escritorio capaces de realizar el trabajo que hace sólo dos décadas estaba reservado a los grandes centros de investigación y a las universidades.

Este desarrollo ha favorecido la aparición de nuevos métodos de cálculo basados en viejas ideas que, de alguna manera, están generando una segunda revolución en la física médica, y cuya influencia será esta vez más obvia en las técnicas de tratamiento. Este artículo presenta una breve introducción a los dos métodos de la física computacional que a mi juicio tendrán influencia considerable en la medicina: los algoritmos genéticos y los métodos de Monte Carlo.

ALGORITMOS GENÉTICOS

La idea de aplicar los principios biológicos de la evolución natural a sistemas artificiales surgió hace más de cuarenta años. Durante las décadas de 1960 y 1970, John Holland estableció

los principios básicos de lo que hoy conocemos como algoritmos genéticos y cómputo evolutivo (Goldberg, 1989; Holland, 1992). La parte medular de estos métodos es la idea de resolver problemas haciendo evolucionar una población —generada inicialmente al azar— de soluciones posibles mediante el uso de “operadores genéticos”, de manera que gradualmente emerjan mejores soluciones.

Los algoritmos evolutivos se han aplicado con éxito en numerosos problemas de búsqueda y optimización en diversos campos, incluyendo la ingeniería, la programación automática, el aprendizaje de máquina y la economía; aquí examinaremos sus aplicaciones a la física médica.

ENCONTRANDO LA AGUJA EN EL PAJAR

Hay muchos problemas en los que es necesario encontrar un conjunto de valores óptimos para una serie de parámetros desconocidos. Normalmente la complejidad del problema está asociada directamente al número de variables que en él aparecen cuando es planteado. Los métodos que se han desarrollado para atacar los problemas requieren el uso de computadoras para manejar una gran cantidad de datos y efectuar muchas operaciones en el menor tiempo posible. En general, estos métodos se clasifican en tres grandes grupos: algebraicos, enumerativos y aleatorios (Coley, 1996).

Los métodos algebraicos pretenden identificar los valores extremos de una función matemática, ya sea mediante la resolución de un sistema de ecuaciones, o empleando la hipótesis de que la pendiente de la gráfica que representa a la ecuación es nula en los extremos. En los métodos enumerativos, la función se evalúa punto por punto hasta encontrar el extremo buscado. Aunque estos métodos pueden (al menos en principio) localizar directamente los valores máximos y mínimos de una función, éstos no necesariamente son extremos globales. Por ejemplo, la Figura 1 muestra un espacio de búsqueda hipotético para una función $z = f(x,y)$ que depende de dos parámetros, x e y , con varios extremos locales. Dependiendo de las condiciones iniciales, los métodos tradicionales de búsqueda pueden quedar atrapados fácilmente en alguno de estos extre-

Los algoritmos evolutivos se han aplicado con éxito en numerosos problemas de búsqueda y optimización en diversos campos

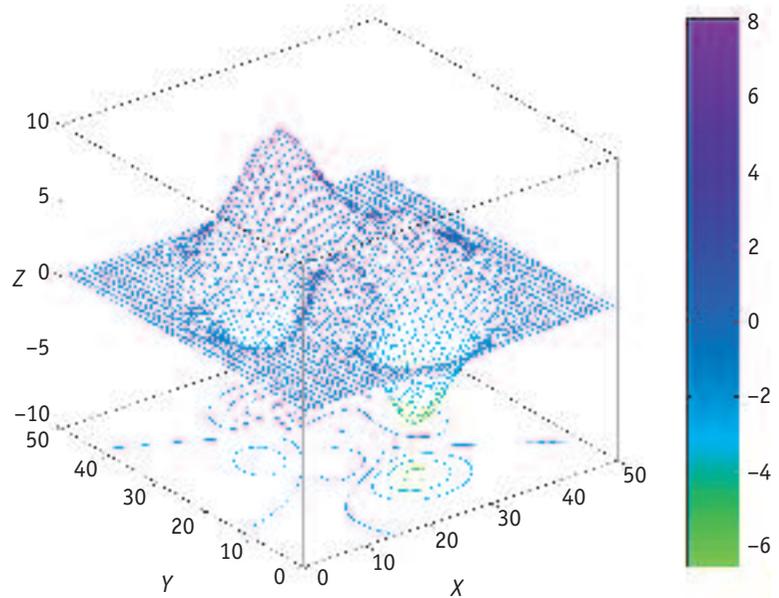


Figura 1. Ejemplo de una superficie de potencial $z = f(x,y)$, con máximos y mínimos locales. El objetivo de los algoritmos de optimización es encontrar el máximo o el mínimo absolutos, sin quedar atrapados en los extremos locales.

mos, y reportar erróneamente que se ha alcanzado el objetivo del procedimiento.

Los métodos evolutivos son algoritmos aleatorios mejorados, que hacen una búsqueda eficiente en el espacio de soluciones posibles; entre ellos se encuentran el método llamado *recocido simulado* (*simulated annealing*) y los algoritmos genéticos. Si comparamos los algoritmos genéticos con las técnicas descritas anteriormente, las principales diferencias que hallamos son:

En los métodos enumerativos, la función se evalúa punto por punto hasta encontrar el extremo buscado

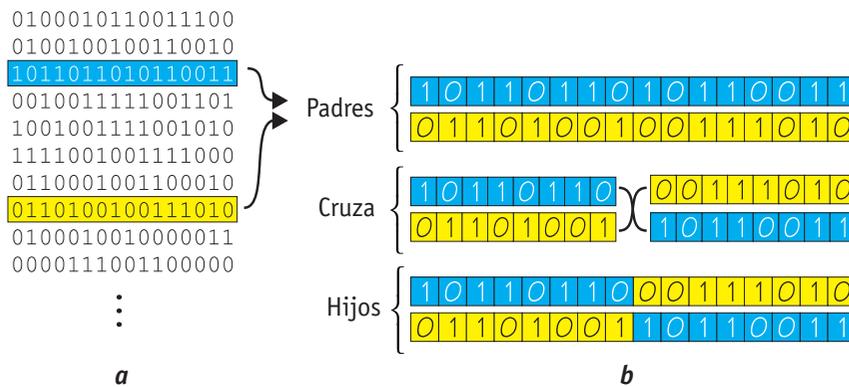


Figura 2. En el proceso de reproducción: en *a* se selecciona a los elementos más aptos para pasar a la etapa de cruce, en *b*. Algunos de los individuos de la nueva generación pueden sufrir mutaciones.

a) Trabajan con una representación (o codificación) del conjunto de parámetros, no con los parámetros mismos.

b) Buscan a partir de un conjunto de soluciones posibles (población), no a partir de un solo punto.

c) Usan una función de “costo” (también llamada *función objetivo*) como criterio de búsqueda, y no sus derivadas.

d) Usan reglas de transición aleatorias, no deterministas.

OPERACIONES BÁSICAS

En general, los algoritmos genéticos funcionan de acuerdo con dos tipos de procesos: el primero de ellos es genético, y tiene que ver con las operaciones de reproducción, cruce y mutación; el segundo es evolutivo, y está relacionado con la operación de selección. El proceso inicia con una población de individuos, en la que cada uno de ellos representa un conjunto de valores generado al azar para los parámetros en cuestión. Con estos parámetros se evalúa la función objetivo, y el resultado representa la calidad de la solución correspondiente a cada individuo. Esta información se usa para seleccionar de manera aleatoria a parejas de individuos que son aptas para la reproducción y cruce (Figura 2). El proceso se repite sobre la nueva generación, hasta que el resultado cae dentro del intervalo de precisión deseado.

La *reproducción* consiste en copiar sin modificación alguna cierto número de pobladores, usando el criterio de selección, de manera que los mejores individuos de una generación pasen a la siguiente. Así se garantiza que la carga genética de los individuos más aptos (aquellos que representan las mejores soluciones) participe en el proceso de cruce. El proceso de *cruza* constituye el operador genético principal. Como se observa en la Figura 2, la descendencia se forma escogiendo un punto de cruce aleatorio en la cadena binaria que representa a los padres, e intercambiando el material genético para formar dos nuevos individuos. Algunos de los nuevos individuos pueden sufrir mutaciones espontáneas (paso de 0 a 1, o viceversa, en alguna posición), lo cual ayuda a remplazar los genes perdidos o a introducir genes que no se encontraban en la población inicial. De esta manera, lo que hacen los algoritmos genéticos es realizar un rastreo en paralelo sobre el espacio de soluciones, disminuyendo así la posibilidad de que el algoritmo se estanque en un mínimo o en un máximo local.

APLICACIONES DE LOS ALGORITMOS GENÉTICOS

Algunas de las aplicaciones más recientes de los algoritmos genéticos en física médica incluyen tanto el diagnóstico como la radioterapia. Uno de los usos más interesantes consiste en dise-

ñar programas para la detección automática de microcalcificaciones (pequeñísimas lesiones asociadas al cáncer de mama) en sistemas de diagnóstico computarizado para mastografía digital (Anastasio y cols., 1998; Sahiner y cols., 2000). El objetivo final es que estos programas se conviertan en una herramienta equivalente a un “segundo observador”, que ayude al radiólogo a reducir el número de diagnósticos equivocados a niveles equivalentes a los que obtendrían dos observadores humanos especializados.

Otra aplicación interesante es en la planificación de tratamientos tanto de radioterapia conformal (Wu y cols., 2000), usando haces de radiación externos, como en braquiterapia de alta y baja tasa de dosis (Herrera y Martínez, 2000; Yang y cols., 1998), en la que fuentes radiactivas encapsuladas se insertan directamente en el tejido tumoral. En este tipo de tratamientos se requiere que la distribución espacial de dosis sea uniforme, que se adapte lo mejor posible a la forma y volumen del tejido que se desea tratar (que sea *conformal*), y que se minimice la dosis a órganos y tejidos radiosensibles que se encuentren cerca de la zona de tratamiento. En esta clase de problemas no existe un objetivo único, y por lo tanto se debe alcanzar un balance entre los diferentes requerimientos y restricciones que lo definen. La Figura 3 muestra un mapa de la distribución espacial de dosis obtenida después de aproximadamente 100 ciclos de cálculo sobre una población de 50 individuos en la planificación de un tratamiento para cáncer de próstata usando fuentes radiactivas de iridio-192 (Martínez-Dávalos, 2000). Se puede observar que la uretra (círculo central) queda protegida, al mismo tiempo que la uniformidad y la conformidad son razonablemente buenas. Una gran ventaja es que estos métodos permitirán en un futuro cercano planear tratamientos complejos de manera más rápida y precisa.

LOS MÉTODOS DE MONTE CARLO

Los métodos de Monte Carlo son herramientas matemáticas que han sido utilizadas extensamente en muchas áreas de la física desde hace 50 años, y consisten en la simulación numérica de fenómenos que ocurren en la naturaleza. Es bien sabido que muchos de los sucesos naturales son aleatorios, es decir, son eventos en los que no es posible predecir con certeza lo que va a ocurrir con ellos en algún momento determinado. Un buen ejemplo de este tipo de fenómenos es el clima: aunque los cien-

tíficos pueden tener información muy precisa respecto a la velocidad del viento, la humedad, la presión atmosférica, la latitud y la longitud de una región específica del planeta, no es posible predecir con absoluta certeza el desarrollo, por ejemplo, de algún huracán.

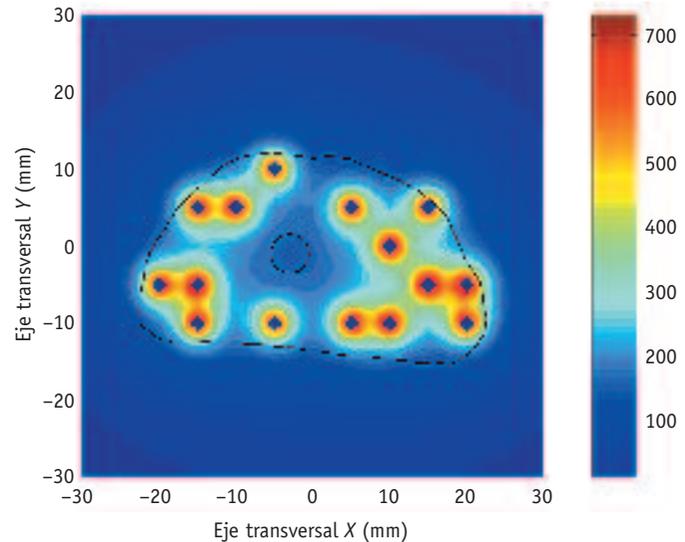


Figura 3. Distribución de dosis para un tratamiento de braquiterapia de próstata, calculada mediante un algoritmo genético. Los puntos representan las posiciones de las fuentes radiactivas, y los números corresponden al valor relativo de la dosis a lo largo de las curvas para cada color.

Los algoritmos genéticos funcionan de acuerdo con dos tipos de procesos: el primero es genético, y el segundo es evolutivo



El primer informe científico de una aplicación del método de Monte Carlo fue hecho en el siglo XVIII por Buffon, quien propuso un procedimiento para calcular π

El primer informe científico sobre una aplicación simple del método Monte Carlo fue hecho en el siglo XVIII por G.-L. Leclerc de Buffon, quien propuso un método numérico para encontrar el valor de π ($\pi = 3.1415926\dots$). El procedimiento es muy ingenioso y consiste en contar cuántas veces una aguja de longitud L , la cual se deja caer de manera aleatoria sobre un tablero, cruza un conjunto de líneas paralelas equidistantes (separadas también una distancia L) dibujadas sobre dicho tablero. Se puede demostrar matemáticamente que si se repitiera este procedimiento un número infinito de veces, entonces se cumpliría la siguiente relación: el número de veces que se deja caer la aguja dividido por el número de veces que cruza alguna de las líneas es igual a $\pi/2$. En la práctica, no es posible repetir el experimento un número infinito de veces, y lo que sucede normalmente es que se repite un número suficientemente grande de veces de tal manera que el cociente anterior se aproxima, o tiende, al valor de $\pi/2$.

El experimento se podría llevar a cabo físicamente tirando muchas veces la aguja sobre el tablero y registrando todos los eventos, o bien se puede realizar utilizando una computadora, en lo que se denomina *simulación del experimento*. En este caso estamos realizando genuinamente una simulación Monte Carlo tal como la conocemos hoy.

TRANSPORTE DE RADIACIÓN EN LA MATERIA

Las técnicas de Monte Carlo han sido particularmente útiles en la física médica por su facultad de simular el transporte de radiación ionizante en la materia, pues la mayor parte de las técnicas para el diagnóstico y tratamiento de enfermedades se basan en el uso de este tipo de radiación, es decir, la que cuenta con suficiente energía para sacar electrones atómicos de sus órbitas.

Los elementos básicos de cualquier simulación de Monte Carlo son los siguientes:

- 1) Un conjunto de funciones matemáticas que representen la probabilidad de que ocurra un fenómeno determinado.
- 2) Una lista (o conjunto) muy grande de números generados al azar.
- 3) Técnicas matemáticas que permitan obtener muestras de las funciones de probabilidad.

4) Una computadora rápida y eficaz para el procesamiento y para el acceso de datos.

Para las aplicaciones que nos interesan se requiere conocer básicamente las probabilidades de interacción de dos tipos de radiaciones: fotones y electrones. Éstos pueden interactuar con la materia de dos formas: ser absorbidos o ser dispersados. Además, la mayoría de las interacciones hacen que las partículas incidentes pierdan energía. La absorción implica la desaparición completa de la partícula original y la transferencia de su energía a los electrones del medio, o bien a la creación de nuevas partículas (por ejemplo, la producción de un par formado por un electrón y un positrón). En la dispersión, una parte de la energía se transfiere a los electrones del medio, y el resto ocasiona un cambio de dirección de la partícula incidente. Cualquiera de estos procesos hace que algunos electrones del medio se pongan en movimiento.

Los diferentes tipos de interacciones dependen de la energía de las partículas y de la composición química del material irradiado. Dentro de las interacciones que pueden sufrir los fotones se encuentran la absorción fotoeléctrica, la dispersión tipo Compton y la producción de pares. Por su parte, los electrones sufren colisiones eléctricas que se pueden clasificar como *elásticas* (cambian su dirección pero no su energía), *inelásticas* (cambian su dirección y su energía) y *radiativas* (la energía del movimiento de los electrones se convierte en radiación electromagnética).

En la Figura 4a se presenta un ejemplo de la interacción de un fotón con un átomo, que resulta en la absorción del fotón y la ionización de átomo. La Figura 4b muestra esquemáticamente la dispersión inelástica de un fotón (dispersión tipo Compton), en la que un fotón incidente interactúa con un electrón libre del medio, cambia de dirección y pone en movimiento al electrón.

CASCADAS ELECTROMAGNÉTICAS

Tomando en cuenta lo anterior, podemos pensar en la creación de un proceso de cascada en el que se puede llevar a cabo alguna de las siguientes posibilidades, o todas ellas:

- Los electrones pueden crear fotones y poner en movimiento a los electrones del medio, que a su vez pueden poner en movimiento a otros electrones.
- Los fotones pueden crear electrones y positrones.
- Los fotones, los electrones y los positrones se dispersan en el medio.

Las interacciones de las partículas ionizantes dependen de su energía y de la composición química del material irradiado

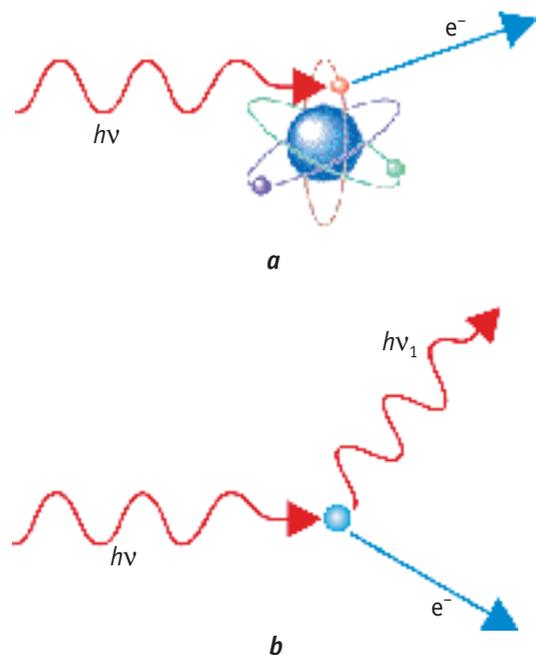


Figura 4. Ejemplos de interacciones de fotones con la materia: *a*, absorción fotoeléctrica; *b*, dispersión tipo Compton. ($h\nu$ = fotón; e^- = electrón.)

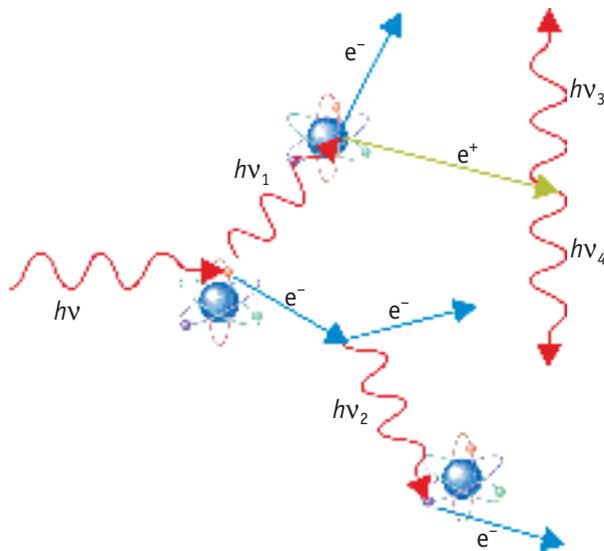


Figura 5. Diagrama de una cascada electromagnética iniciada por un fotón.

- Los positrones se aniquilan con los electrones, y crean fotones.

A cada sucesión particular de eventos (o “historia”) que un electrón o un fotón producen al viajar en un medio se le llama, junto con todos sus descendientes, *cascada electromagnética*. La Figura 5 muestra un ejemplo hipotético muy sencillo de una cascada electromagnética originada por un fotón viajando en un medio. Las líneas rojas representan fotones, las líneas azules electrones, y la verde, un positrón.

Dado que se conocen los procesos físicos que rigen a cada una de las interacciones mencionadas, y debido a que no se pueden resolver en forma precisa las ecuaciones para el transporte de electrones y fotones en un medio, la técnica de Monte Carlo es la herramienta ideal para estudiar los problemas del transporte de radiación en la materia.

Por ejemplo, la Figura 6 muestra la simulación de un haz de electrones de alta energía que inciden sobre un cubo de agua. Las líneas azules representan electrones, y las amarillas, fotones. La complejidad de la estructura de la cascada depende mucho de la energía de la radiación incidente; mientras mayor sea la energía, más compleja será la cascada.

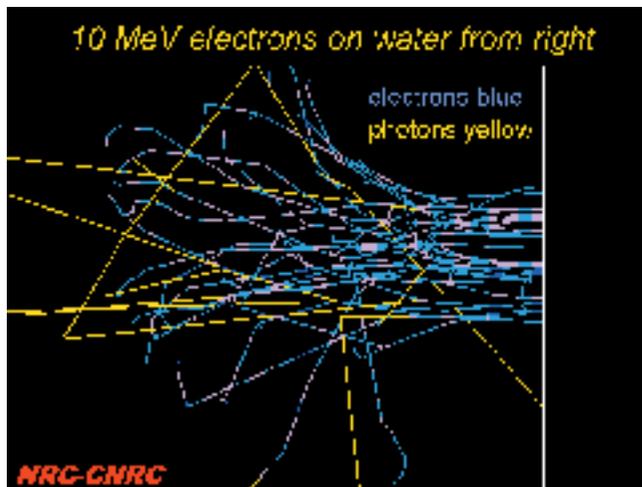
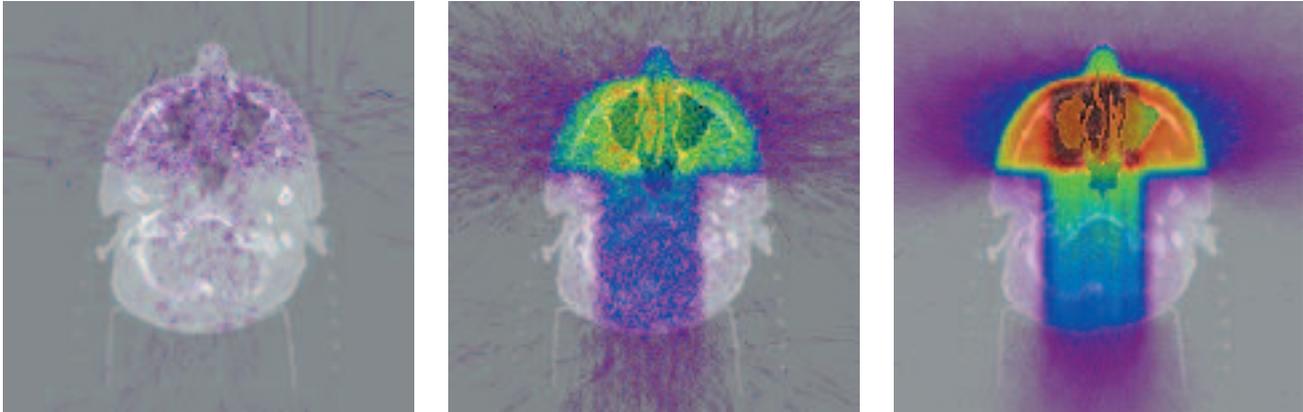


Figura 6. Cascada electromagnética producida en agua por un haz de electrones de 10 MeV de energía en agua. Las trayectorias de los electrones aparecen con líneas azules, y las de los fotones con líneas amarillas. (http://www.irs.inms.nrc.ca/inms/irs/BEAM/egs_windows/sld018.html)

Actualmente, los cálculos más precisos para determinar las distribuciones de la dosis en los tratamientos de cáncer que usan radioterapia se llevan a cabo mediante técnicas de Monte Carlo. La Figura 7 muestra una secuencia de tres imágenes en las que se observa la distribución de la dosis calculada (en colores) superpuesta a una imagen de tomografía axial computarizada de la cabeza de un paciente. El color rojo significa mayores dosis, mientras que la azul corresponde a menores dosis. En este cálculo se pretende irradiar al paciente utilizando tres haces de fotones, centrados en la región de los senos paranasales. La diferencia entre las tres imágenes es la cantidad de fotones que se usaron para el cálculo (66 mil, 1 millón 200 mil y 68 millones, respectivamente).

Si bien en la actualidad aún no es posible realizar este tipo de cálculos con la suficiente rapidez para usarlos rutinariamente en aplicaciones clínicas, se espera que el desarrollo del cómputo en paralelo permitirá en un futuro no muy lejano llevar estas técnicas a una mayor población.



BIBLIOGRAFÍA

- Anastasio, M. A., H. Yoshida, R. Nagel y cols. (1998), "A genetic algorithm-based method for optimizing the performance of a computer-aided diagnosis scheme for detection of microcalcifications in mammograms", *Med. Phys.* 25(9):1613-1620.
- Coley, D. A. (1996), "Genetic algorithms", *Contemporary Physics*, 37(2):145-154.
- Goldberg, D. E. (1989), *Genetic Algorithms in Search, Optimization and Machine Learning*, Reading, Addison-Wesley.
- Herrera-Rodríguez, M. R., y A. Martínez-Dávalos (2000), "Brachytherapy treatment planning algorithm applied to prostate cancer", *AIP Conf. Proc.*, 538:249-254.
- Holland, J. H. (1992), *Adaptation in Natural and Artificial Systems*, Cambridge, MIT Press.
- Martínez-Dávalos, A. (2000), "Monte Carlo dosimetry in HDR brachytherapy", *AIP Conf. Proc.*, 538:128-133.
- Radon, J. (1917), "Über die Bestimmung von Functionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten", *Ber. Verhändl. Sächs. Akad. Wiss. Leipzig, Math.-Phys. Kl.*, 69:262-339.
- Sahiner, B., H.-P. Chan, N. Petrick y cols. (2000), "Design of a high-sensitivity classifier based on a genetic-algorithm: Application to computer-aided diagnosis", *Phys. Med. Biol.*, 43:2853-2871.
- Wu, X., Y. Zhu, J. Dai y cols. (2000), "Selection and determination of beam weights based on genetic algorithms for conformal radiotherapy treatment planning", *Phys. Med. Biol.*, 45:2547-2558.

Figura 7. Simulación de Monte Carlo para el cálculo de dosis en un tratamiento de radioterapia con un haz de fotones. Las distribuciones de dosis (a colores) están superpuestas a una imagen tomográfica (tonos de gris) del paciente.
(<http://www.llnl.gov/peregrine/images.html>)

- Yang, G., S. Reinstein, S. Pai, Z. Xu y D. L. Carroll (1998), "A new genetic algorithm technique in optimization of permanent ^{125}I prostate implants", *Med. Phys.*, 25(12):2308-2315.

Arnulfo Martínez Dávalos es doctor en física de radiaciones por el University College de Londres, Inglaterra. Actualmente, como investigador asociado en el Instituto de Física de la UNAM, realiza análisis sobre detectores de radiación ionizante y sus aplicaciones en física básica y médica, métodos de Monte Carlo para el transporte de radiación en materia, y física computacional. Es profesor de la maestría en Física Médica de la misma universidad.