

Introducción a la Computación Evolutiva

Carlos A. Coello Coello

carlos.coellocoello@ccinvestav.mx

CINVESTAV-IPN

Evolutionary Computation Group (EVOCINV)

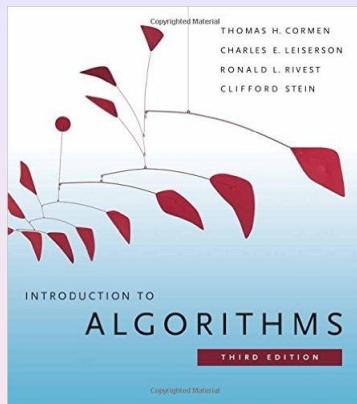
Departamento de Computación

Av. IPN No. 2508, Col. San Pedro Zacatenco

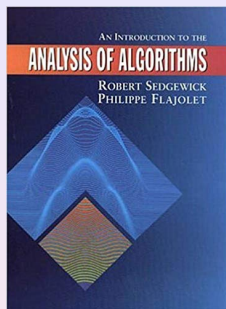
México, D.F. 07360, MEXICO

Clase 1

Conceptos Básicos de Análisis de Algoritmos

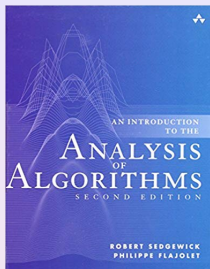


- Análisis *a priori* de algoritmos
- Orden de magnitud de un algoritmo



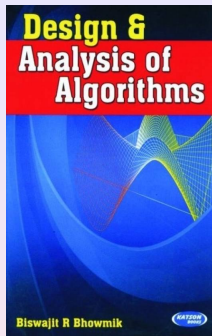
Análisis *a priori* de algoritmos

Se ignoran los detalles que sean dependientes de la arquitectura de una computadora o de un lenguaje de programación y se analiza el orden de magnitud de la frecuencia de ejecución de las instrucciones de un algoritmo.



Orden de magnitud de un algoritmo

- Suele usarse la notación “O” (big-O)
- Si un algoritmo tiene complejidad $O(g(n))$ significa que al ejecutarlo en una computadora con los mismos datos, pero valores incrementales de n , los tiempos resultantes de ejecución serán siempre menores que $|g(n)|$



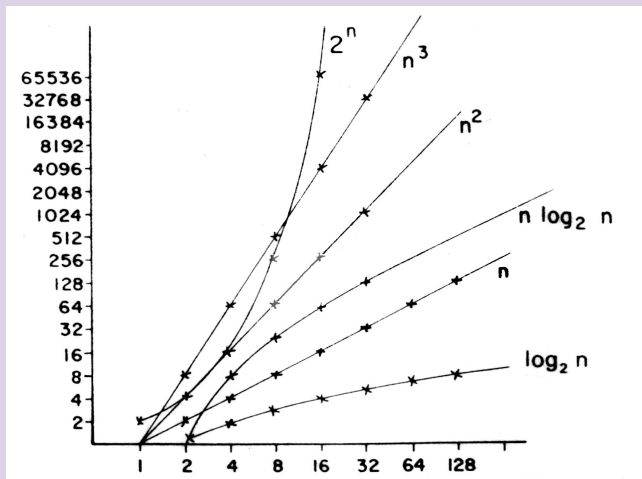
Orden de magnitud de un algoritmo

Los órdenes de magnitud más comunes de los algoritmos son:

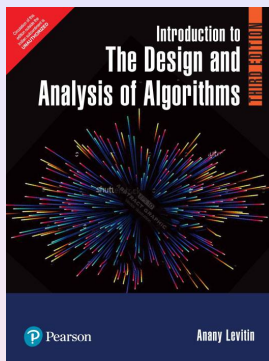
$$O(1) < O(\log n) < O(n) < O(n \log n) < O(n^2) < O(n^3) < O(2^n)$$

Conceptos Básicos de Análisis de Algoritmos

Orden de magnitud de un algoritmo



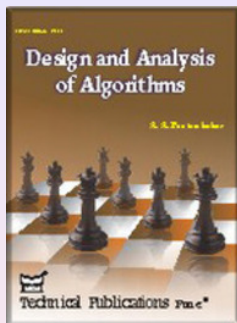
Conceptos Básicos de Análisis de Algoritmos



Clase **P**

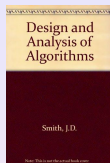
Un problema pertenece a la clase **P** si puede ser resuelto en tiempo polinomial en una computadora determinística.

Ejemplos: Quicksort, búsqueda binaria, multiplicación matricial.



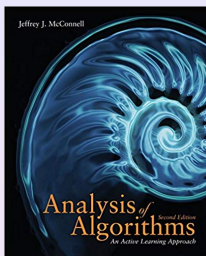
Clase **NP**

Un problema pertenece a la clase **NP** si puede ser resuelto en tiempo polinomial pero usando una computadora no determinística.



P vs. NP

- La clase **P** contiene problemas que pueden resolverse rápidamente.
- La clase **NP** contiene problemas cuya solución puede verificarse rápidamente.
- En 1971 se planteó la pregunta: ¿Es **P = NP**? Desde entonces, sigue siendo una pregunta abierta para los teóricos.
- Se cree que **P ≠ NP**

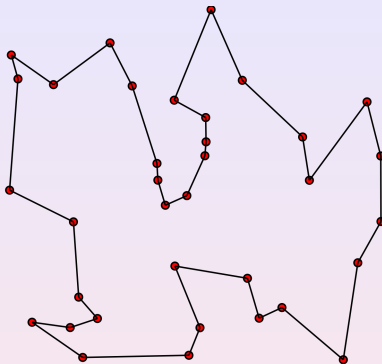


Problemas NP Completos

Todos los algoritmos requeridos para resolverlos requieren tiempo exponencial en el peor caso.

Es decir, estos problemas son sumamente difíciles de resolver.

Ejemplo: el problema del viajero, $O(n^2 2^n)$



El Problema del Viajero

Encontrar una permutación que represente el recorrido de una serie de ciudades de tal forma que todas sean visitadas minimizando la distancia total viajada.



El Problema del Viajero

Si consideramos n ciudades:

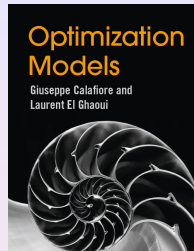
- El tamaño del espacio de búsqueda es: $(n - 1)!/2$
- Para $n=10$, hay unas 181,000 soluciones posibles.
- Para $n=20$ hay unas 10,000,000,000,000,000 soluciones posibles.



El Problema del Viajero

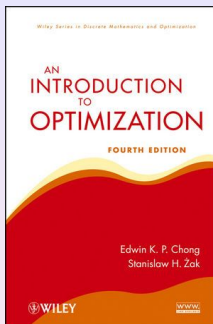
Sólo hay 1,000,000,000,000,000,000,000 litros de agua en el planeta

Técnicas Clásicas de Búsqueda y Optimización



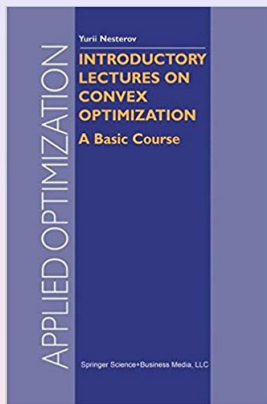
Existen muchas técnicas clásicas para resolver problemas con ciertas características específicas.

Es importante saber al menos de la existencia de estas técnicas, pues cuando el problema por resolverse se adecúa a ellas, no tiene ningún sentido usar heurísticas.

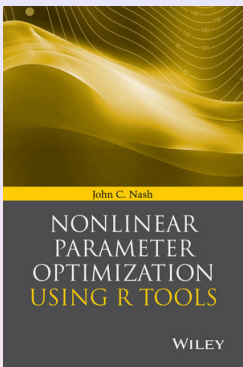


Para optimización lineal, el método Simplex sigue siendo la opción más viable.

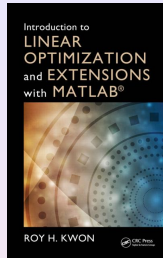
Para optimización no lineal, hay métodos directos (p. ej. la búsqueda aleatoria) y métodos no directos (p. ej. el método del gradiente conjugado).



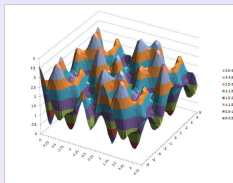
Existen también técnicas que construyen parcialmente una solución a un problema. Por ejemplo, la programación dinámica y el método de ramificación y búsqueda (*branch & bound*).



Cuando enfrentamos un cierto problema de optimización, si la función a optimizarse se encuentra en forma algebraica, es importante intentar resolverla primero con técnicas clásicas, antes de utilizar cualquier heurística.



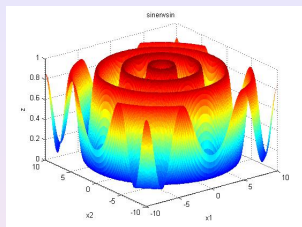
- Existen problemas que no pueden resolverse usando un algoritmo que requiere tiempo polinomial.
- De hecho, en muchas aplicaciones prácticas, no podemos siquiera decir si existe una solución eficiente.
- Hay muchos problemas para los cuales el mejor algoritmo que se conoce requiere tiempo exponencial.



Optimización Global

El objetivo principal de cualquier técnica de optimización es encontrar el óptimo (o los óptimos) globales de cualquier problema.

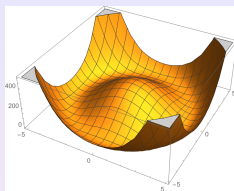
En matemáticas, existe un área que se ocupa de desarrollar los formalismos que nos permitan garantizar la convergencia de un método hacia el óptimo global de un problema.



Optimización Global

Desgraciadamente, sólo en algunos casos limitados, puede garantizarse convergencia hacia el óptimo global.

Por ejemplo, para problemas con espacios de búsqueda convexos, las condiciones de Kuhn-Tucker son necesarias y suficientes para garantizar optimalidad global de un punto.



Optimización Global

En problemas de optimización no lineal, las condiciones de Kuhn-Tucker no son suficientes para garantizar optimalidad global.

De hecho, todas las técnicas usadas para optimización no lineal pueden localizar cuando mucho óptimos locales, pero no puede garantizarse convergencia al óptimo global a menos que se usen técnicas exhaustivas o que se consideren tiempos infinitos de convergencia.

Optimización Numérica

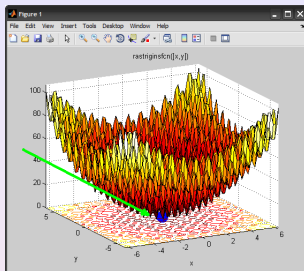
Existen muchos tipos de problemas de optimización, pero los que nos interesan más para los fines de este curso, son de los de optimización numérica, que pueden definirse de la siguiente manera:

$$\text{Minimizar } f(\vec{x})$$

sujeta a:

$$g_i(\vec{x}) \leq 0 \quad i = 1, \dots, p$$

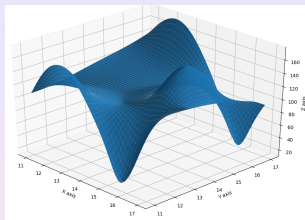
$$h_j(\vec{x}) = 0 \quad j = 1, \dots, n$$



Optimización Numérica

Llamaremos a (\vec{x}) las variables de decisión del problema, $g_i(\vec{x})$ son las restricciones de desigualdad, y $h_j(\vec{x})$ son las restricciones de igualdad.

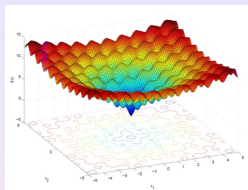
Asimismo, $f(\vec{x})$ es la función objetivo del problema (la que queremos optimizar).



Optimización Numérica

A las restricciones de igualdad y desigualdad expresadas algebraicamente, se les denomina “restricciones explícitas”.

En algunos problemas, existen también “restricciones implícitas”, relacionadas sobre todo con las características del problema.



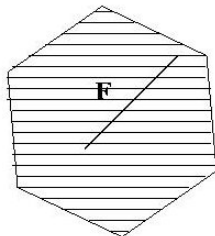
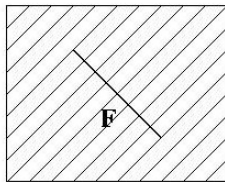
Optimización Numérica

Por ejemplo, si decimos:

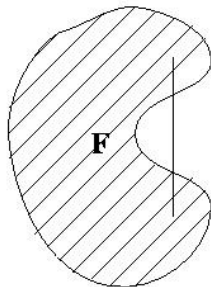
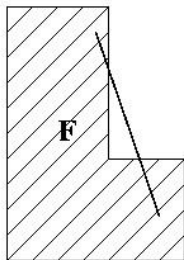
$$10 \leq x_1 \leq 20$$

estamos definiendo que el rango de una variable de decisión debe estar contenido dentro de un cierto intervalo. De tal forma, estamos “restringiendo” el tipo de soluciones que se considerarán como válidas.

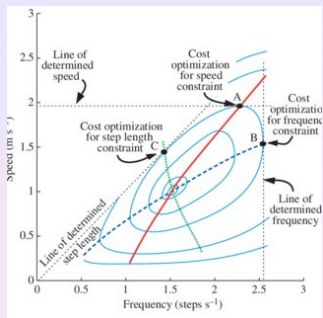
Ejemplos de espacios de búsqueda convexos



Ejemplos de espacios de búsqueda no convexos



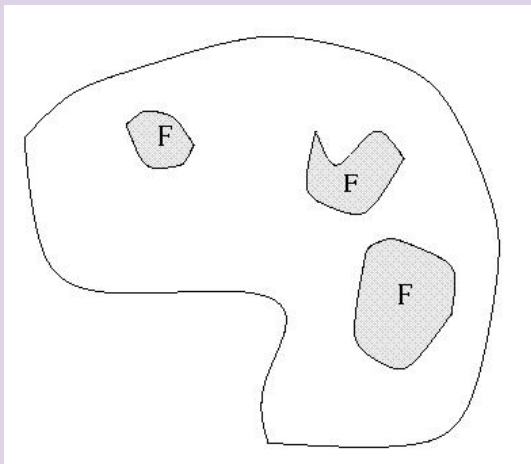
Técnicas Clásicas de Búsqueda y Optimización

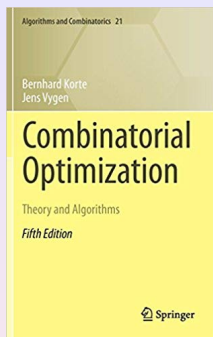


Zona factible y no factible

Todas las soluciones a un problema que satisfagan las restricciones existentes (de cualquier tipo), se consideran ubicadas dentro de la zona factible. De tal forma, podemos decir que el espacio de búsqueda de un problema se divide en la región (o zona) factible y la no factible.

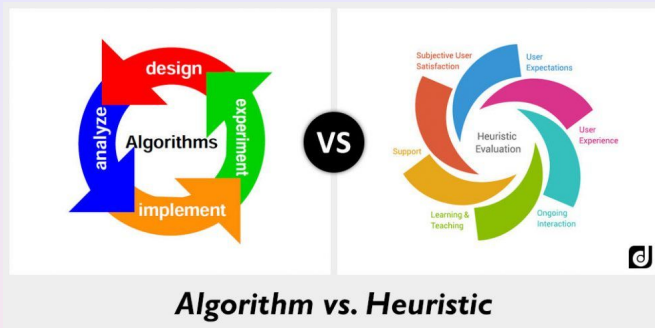
Zona factible y no factible





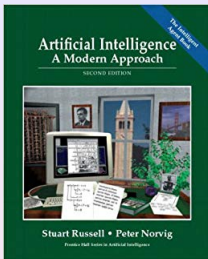
Optimización Combinatoria

Existe una clase especial de problemas que también serán de interés para este curso, en los cuales las variables de decisión son discretas y las soluciones suelen presentarse en la forma de permutaciones. A estos problemas se les denomina de **optimización combinatoria** (p. ej. el problema del viajero).



La palabra “heurística” se deriva del griego heuriskein, que significa “encontrar” o “descubrir”. El significado del término ha variado históricamente. Algunos han usado el término como un antónimo de “algorítmico”. Por ejemplo, Newell et al. dicen: “a un proceso que puede resolver un cierto problema, pero que no ofrece ninguna garantía de lograrlo, se le denomina una “heurística” para ese problema”

¿Qué es una heurística?



Las heurísticas fueron un área predominante en los orígenes de la Inteligencia Artificial. Actualmente, el término suele usarse como un adjetivo, refiriéndose a cualquier técnica que mejore el desempeño en promedio de la solución de un problema, aunque no mejore necesariamente el desempeño en el peor caso [Russell & Norvig, 1995].

¿Qué es una heurística?



Una definición más precisa y adecuada para los fines de este curso es la proporcionada por Reeves (1993):

Una heurística es una técnica que busca soluciones buenas (es decir, casi óptimas) a un costo computacional razonable, aunque sin garantizar factibilidad u optimalidad de las mismas. En algunos casos, ni siquiera puede determinar qué tan cerca del óptimo se encuentra una solución factible en particular.

¿Qué es una metaheurística?



El término **metaheurística** lo acuñó Fred Glover en su artículo seminal sobre búsqueda tabú de 1986.

Definición

Las metaheurísticas son procedimientos de búsqueda de alto nivel que aplican alguna regla o conjunto de reglas que se basa(n) en alguna fuente de conocimiento, a fin de explorar el espacio de búsqueda de manera más eficiente.

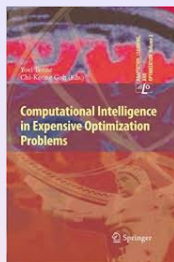
Ventajas y Desventajas de las Metaheurísticas



En optimización, una metaheurística puede verse como un método numérico que utiliza cierta información del problema para hacer una búsqueda relativamente eficiente, sin garantizar que se obtendrá el óptimo global.

Pese a sus aparentes limitantes, las metaheurísticas se han vuelto muy populares en los últimos 25 años, porque los métodos exactos sólo pueden aplicarse a ciertas clases de problemas de optimización y los diseñados para problemas discretos se basan en el uso de búsquedas exhaustivas o semi-exhaustivas.

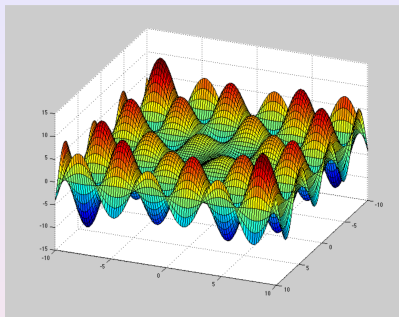
Ventajas y Desventajas de las Metaheurísticas



Las metaheurísticas se han vuelto muy populares por su flexibilidad y facilidad de uso, además de constituir la última línea de defensa en optimización.

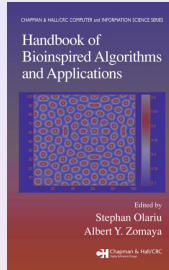
Sin embargo, además de no poder garantizar que encontrarán la solución óptima a un problema, su costo computacional hace prohibitivo el uso de las metaheurísticas en algunos problemas.

¿Cuándo usar las metaheurísticas?



Cuando enfrentamos espacios de búsqueda muy grandes y accidentados, y para los cuales, los algoritmos más eficientes que existen para resolverlos requieren tiempo exponencial, resulta obvio que las técnicas clásicas de búsqueda y optimización son insuficientes. Ese es precisamente el dominio de las metaheurísticas.

Metaheurísticas bio-inspiradas



Existe una clase particular de metaheurísticas que se ha vuelto muy popular en los últimos años, en la cual las reglas para elegir soluciones se basan en conceptos biológicos tales como la evolución natural, los movimientos de las hormigas, los patrones de vuelo de las aves, etc.

A este tipo de técnicas se les conoce como **metaheurísticas bio-inspiradas**. Los algoritmos evolutivos son un tipo de metaheurística bio-inspirada.

Teorías sobre el origen de las especies



Durante muchos años, la tesis más aceptada sobre el origen de las especies fue el **creacionismo**: Dios creó a todas las especies del planeta de forma separada.

Teorías sobre el origen de las especies



Además, según el creacionismo, las especies estaban jerarquizadas por Dios de tal manera que el hombre ocupaba el rango superior, al lado del creador.

Teorías sobre el origen de las especies



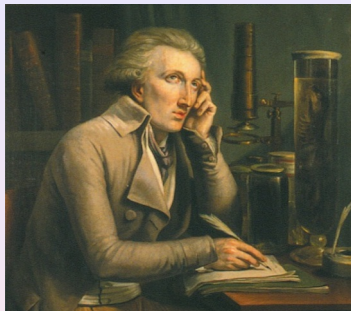
Georges Louis Leclerc (Conde de Buffon) fue tal vez el primero en especular (100 años antes que Darwin) que las especies se originaron entre sí, e incluso especuló sobre la posible existencia de un ancestro común entre el hombre y los simios, aunque después, él mismo refutó esta hipótesis. Varias de sus ideas fueron, sin embargo, revolucionarias para su época.

Teorías sobre el origen de las especies



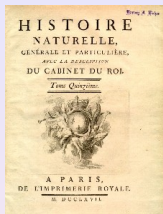
Leclerc sostenía que los cambios que ocurren en los seres vivos se deben a la acción de factores externos. Postuló la denominada **Teoría de las transformaciones infinitas**, la cual establece que los cambios de los seres vivos se deben al tiempo, al medio físico, a la forma de nutrición y a la domesticación.

Teorías sobre el origen de las especies



Aunque acarició la idea de una teoría evolutiva, Leclerc no llegó a romper del todo con el creacionismo. Sin embargo, reconoció que las especies experimentan variaciones y que algunas, incluso llegaron a desaparecer. Estas ideas tendrían gran influencia en naturalistas posteriores como Lamarck y Darwin.

Teorías sobre el origen de las especies



Leclerc sugirió que las especies pudieron haberse “mejorado” y “degenerado” después de haberse dispersado a partir de un eje central de la creación. En el volumen 14 de su *Histoire naturelle, générale et particulière*, argumenta que todos los cuadrúpedos del mundo se desarrollaron a partir de un conjunto original de sólo 38 cuadrúpedos. Es por ello que algunos lo consideran un “transformista” y precursor de las ideas de Darwin.

Teorías sobre el origen de las especies



Leclerc también indicó que el cambio climático pudo haber facilitado la dispersión de las especies. La interpretación correcta de sus ideas es, sin embargo, muy difícil, dado que las retoma varias veces en su extenso trabajo, cambiando en muchas ocasiones su punto de vista al respecto.

Teorías sobre el origen de las especies



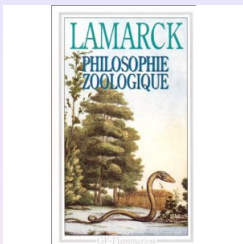
El biólogo francés Jean-Baptiste Lamarck enunció la que se considera como la primera teoría evolutiva coherente de la historia (en 1808).

Teorías sobre el origen de las especies



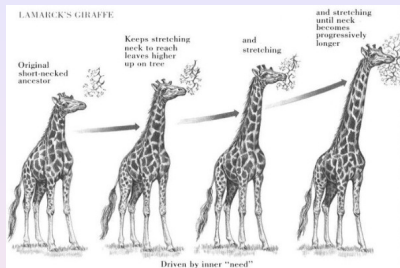
Lamarck indicó correctamente que el ambiente da pie a los cambios en los animales. Esto lo ilustró con ejemplos tales como la ceguera de los topos, la presencia de dientes en los mamíferos y la ausencia de dientes en las aves que para él constituían evidencia de esta teoría.

Teorías sobre el origen de las especies



En sus trabajos, Lamarck señaló que existían dos fuerzas principales que conformaban la evolución: una que forzaba los cambios en los animales, pasándolos de formas simples a otras más complejas, y una segunda que adaptaba a los animales a sus ambientes locales y que los diferenciaba entre sí. Lamarck creía que estas fuerzas debían ser explicadas como una consecuencia necesaria de principios físicos básicos.

Teorías sobre el origen de las especies



Los aspectos más importantes a tener en cuenta sobre la teoría evolutiva de Lamarck son los siguientes:

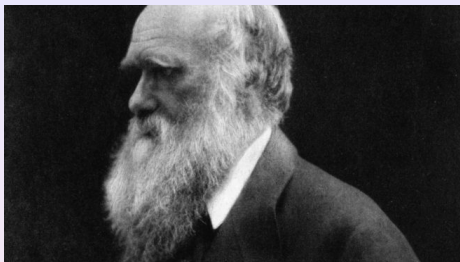
1. Su teoría se centra únicamente en la evolución de los organismos y no en su origen ya que, en aquel entonces se aceptaba que los organismos surgían espontáneamente en sus formas más simples.
2. Propuso que los cambios que sufren los organismos para adaptarse eran heredables. Años después se demostró que esto era incorrecto.

Teorías sobre el origen de las especies



3. La teoría evolutiva de Lamarck constituía una clara oposición a la creencia de la época de que las especies permanecían inmutables desde su creación.
4. Curiosamente, durante el siglo XX han existido evolucionistas que han defendido el llamado **Lamarckismo**, a través de las voces de varios biólogos y evolucionistas que han buscado reivindicar el trabajo de Lamarck.

Teorías sobre el origen de las especies



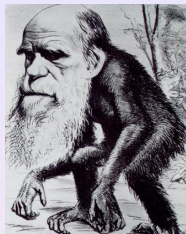
El naturalista inglés Charles Darwin presentó en 1858 los primeros bosquejos de su (ahora famosa) teoría sobre el origen de las especies. Su libro, titulado **On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life**, se publicó el 24 de noviembre de 1859 y se considera como una de las obras científicas más importantes de todos los tiempos.

Teorías sobre el origen de las especies



Darwin entendió que toda población consiste de individuos ligeramente distintos entre sí y que estas pequeñas variaciones hacen que cada uno tenga distintas capacidades para adaptarse a su medio ambiente, así como para reproducirse y para transmitir sus rasgos a sus descendientes.

Teorías sobre el origen de las especies



Con el paso del tiempo (o generaciones), los rasgos de los individuos que mejor se adaptaron a las condiciones del medio ambiente, se vuelven más comunes, haciendo que la población, en su conjunto, evolucione. Darwin llamó a este proceso “descendencia con modificación”. Del mismo modo, la naturaleza selecciona las especies mejor adaptadas para sobrevivir y reproducirse. A este proceso, Darwin lo denominó **selección natural**.