

# Algoritmia para problemas difíciles

Elvira Mayordomo

Universidad de Zaragoza

16 de septiembre de 2022

# Profesores de la asignatura

- Elvira Mayordomo Cámara
- Ubaldo Ramón Júlvez

# Calendario de la asignatura

- moodle

# Material de la asignatura

- Material sensible (claves, enlaces, esta presentación): en moodle
- Material no sensible (transparencias, guiones, etc) en <http://webdiis.unizar.es/asignaturas/APD> y en moodle

**Asymptotic Divergences and Strong Dichotomy.** X. Huang, J.H. Lutz, E. Mayordomo, and D. Stull. IEEE Transactions on Information Theory (2021)

[...] finite-state gamblers that bet on infinite sequences  $S$  [...]

Assume an underlying distribution  $\alpha$  on the symbols

Our main theorem is a *strong dichotomy theorem* that *quantifies* the exponential rates of winning and losing [...] one of the following two things holds for all prefixes  $w$  of  $S$ .

(1) The infinitely-often exponential rate of winning in 1 is  $2^{Divg(S||\alpha)|w|}$ .

(2) The exponential rate of loss in 2 is  $2^{-Risk_G(w)}$ .

**Computing Absolutely Normal Numbers in Nearly Linear Time.** Jack H. Lutz, E. Mayordomo. Information & Computation (2021)

A real number  $x$  is *absolutely normal* if, for every base  $b \geq 2$ , every two equally long strings of digits appear with equal asymptotic frequency in the base- $b$  expansion of  $x$ . This paper presents an explicit algorithm that generates the binary expansion of an absolutely normal number  $x$ , with the  $n$ th bit of  $x$  appearing after  $n \text{polylog}(n)$  computation steps. (The fastest previous algorithm takes  $n^2 \text{polylog}(n)$  steps.) [...]

## **Machine learning classifier for identification of damaging missense mutations exclusive to human mitochondrial DNA-encoded polypeptides.**

A. Martín-Navarro, A. Gaudio-Simón, J. Álvarez-Jarreta, J. Montoya, E. Mayordomo, and E. Ruiz-Pesini. BMC Bioinformatics (2017)

**BACKGROUND:** Several methods have been developed to predict the pathogenicity of missense mutations but none has been specifically designed for classification of variants in mtDNA-encoded polypeptides. [...] Because mtDNA sequencing of patients suffering mitochondrial diseases is revealing many missense mutations, it is needed to prioritize candidate substitutions for further confirmation. [...]

**RESULTS:** We have developed a SVM classifier (Mitoclass.1) specific for mtDNA missense variants. [...] Our classifier has performed better than other web-available tested predictors. [...] Our method has an improved sensitivity and better specificity in relation to PolyPhen-2. [...]

## **LDDMM meets GANs: Generative Adversarial Networks for diffeomorphic registration.**

U. Ramon, M. Hernandez, and E. Mayordomo, Workshop in Biomedical Image Registration (2022)

The purpose of this work is to contribute to the state of the art of deep-learning methods for diffeomorphic registration. We propose an adversarial learning LDDMM method for pairs of 3D mono-modal images based on Generative Adversarial Networks. The method is inspired by the recent literature for deformable image registration with adversarial learning. We combine the best performing generative, discriminative, and adversarial ingredients from the state of the art within the LDDMM paradigm. We have successfully implemented two models with the stationary and the EPDiff-constrained non-stationary parameterizations of diffeomorphisms.



# ¿Por qué estudiar algoritmia?

- El mejor algoritmo es mucho mejor que el mejor computador (Rico ignorante vs pobre experto en algoritmos)
- Los algoritmos son una poderosa lente para ver la informática:
  - Los grandes cambios en los estándares de Internet se pueden ver como debates sobre deficiencias de un algoritmo de camino más corto y ventajas relativas de otro.
  - Los grandes avances de empresas líder como google son algorítmicos (ranking, maps, self-driving car ...)
- No es exclusivo para informáticos:
  - Muchas definiciones novedosas en biología molecular/genética son algorítmicas: similitudes entre genes o genomas
  - En economía: la viabilidad de subastas combinatorias está ligada a la intratabilidad de problemas relacionados ...
- Es divertido ...
- Las empresas buscan estudiantes que sepan algorítmica

# No es exclusivo para informáticos: Elecciones

- ¿Cuál es el **mejor sistema** para elegir la opción más popular?
  - pluralidad, múltiples rondas, opciones 2 a 2, ...
- ¿cómo evitamos la **manipulación** o los sobornos?
  - manipular: mentir en las preferencias para conseguir un resultado más deseable
- ¿cómo conseguimos propiedades como unanimidad o independencia de opciones irrelevantes?
  - unanimidad: si todo el mundo prefiere A a B entonces en el resultado A debe estar por delante de B
  - independencia de opciones irrelevantes: si las opciones ganadoras son A y B, cualquier cambio de opinión en el resto de las opciones no afecta al resultado
- Las soluciones son todas **algorítmicas**
  - algoritmos eficientes para sistemas restringidos con buen comportamiento
  - intratable saber cómo de cerca estamos de uno de esos sistemas restringidos
- *The Handbook of Computational Social Choice*, Brandt, Conitzer, Endriss, Lang and Procaccia. Cambridge U. P. 2016

# Algoritmia para problemas difíciles

- En *Algoritmia Básica* hemos estudiado las principales técnicas para diseñar algoritmos
- ¿Qué otras técnicas quedan?
- ¿Qué problemas difíciles?

# Problemas difíciles

- Algunos ejemplos de problemas tratados en Algoritmia Básica:
  - Coloreado de grafos
  - El viajante de comercio
  - Ciclos hamiltonianos
  - Mochila 0-1
  - ...
- ¿Qué tienen en común estos problemas?
- Los intentos que hicimos (algoritmo voraz, programación dinámica, búsqueda con retroceso, ...) no eran satisfactorios
- O bien los algoritmos **no eran eficientes** o bien encontrar una solución (o una solución óptima) **no estaba garantizado**

# Problemas difíciles

- En esta asignatura trataremos sobre todo problemas:
  - NP-completos
  - NP-difíciles
- Son problemas que se consideran intratables porque nadie ha conseguido encontrar algoritmos **eficientes** y que **resuelvan completamente el problema**
- ¿Qué podemos hacer con uno de estos problemas?
- Primero convencernos de que hay pruebas claras para considerarlo **intratable**
- Segundo utilizar técnicas **imaginativas** que nos acerquen a la solución eficiente (aunque no lo consigan completamente)

# Técnicas imaginativas para problemas difíciles

- **Algoritmos aproximados:** Nos dan una solución cercana al óptimo con una estimación del error cometido
- **Heurísticas:** Son ideas felices que funcionan bien en algunos casos, sin garantías
- **Algoritmos probabilistas:** Algoritmos que resuelven el problema pero necesitan utilizar el azar
- **Estructuras de datos adhoc:** Resolver casos grandes de problemas tratables **Si hay suficiente tiempo**
- **Teoría de la información:** Los algoritmos de compresión de datos influyen en la resolución de problemas generales

# Ideas felices versus garantías

- ¿Qué diferencia hay entre una heurística y un algoritmo aproximado?
- Es la misma diferencia que entre una cuenta corriente y una cartera de acciones
- Una **heurística** es como en una cartera de acciones, puedes ganar mucho dinero pero no hay garantías
- Un **algoritmo aproximado** es como una cuenta corriente, el interés que consigues se conoce a priori
- ¿Qué es mejor? Depende del objetivo, del riesgo ...
- Según otras opiniones, una heurística deja de serlo cuando se convierte en un algoritmo eficiente (conseguimos análisis de eficiencia) o en un algoritmo aproximado (conseguimos análisis de prestaciones)

# Contenidos de la asignatura

- 1 Introducción. Los problemas NP-difíciles.
- 2 Algoritmos aproximados. Concepto. Diseño de algoritmos. Garantías y límites.
- 3 Algoritmos probabilistas: Las Vegas y Montecarlo. Análisis. Generadores pseudoaleatorios.
- 4 Árboles de sufijos (Estructuras de datos avanzadas I).
- 5 Algoritmos de compresión de datos.

**Nota:** Las heurísticas ya se han visto en la asignatura de Inteligencia Artificial.



# Bibliografía Básica

- S.S. Skiena. The algorithm Design Manual, Springer, 1997
- J. Hromkovic. Algorithms for Hard Problems, Springer, 2001  
*(En la biblioteca)*
- T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein.  
Introduction to Algorithms (3rd edition), The MIT Press,  
2009 *(En la biblioteca)*
- G. Brassard, P. Bratley. Fundamentos de Algoritmia, Prentice  
Hall, 1997 *(En la biblioteca)*

# Bibliografía Complementaria

- J. Kleinberg, E. Tardos. Algorithm Design, Addison-Wesley, 2005 (*En la biblioteca*)
- S. Dasgupta, C. Papadimitriou, U. Vazirani. Algorithms, McGraw-Hill, 2008 (*En la biblioteca*)
- V. V. Vazirani. Approximation Algorithms, Springer, 2001 *En la biblioteca y versión electrónica*

# Material adicional

- Información en <http://webdiis.unizar.es/asignaturas/APD>
- Twitter [@APDunizar](https://twitter.com/APDunizar) <https://twitter.com/APDunizar>
- Youtube <https://www.youtube.com/channel/UCPTrK09i344MdFPt1zN-SNA>
- Apuntes, ejercicios y guiones de prácticas

## Evaluación: opción recomendada

- Prácticas de laboratorio (en parejas) **durante el cuatrimestre** : 20 %
- Ejercicios de trabajo individual **durante el cuatrimestre** (3 ejercicios): 20 %  
*Importante: Hay que reservar un ejercicio a través de moodle*
- Prueba escrita intermedia: 30 %
- Examen final (realizar sólo una parte): 30 %

# Evaluación exclusivamente por exámenes finales

- Examen práctico: 20 %<sup>1</sup>
- Examen final completo: 80 %

---

<sup>1</sup>El examen práctico requerirá conocer en detalle las prácticas propuestas durante el curso.

## Dudas frecuentes

- Si un alumno no realiza **todas las prácticas (con nota  $> 0$ )** tiene que presentarse al examen práctico y examen final completo
- Si un alumno no realiza **3 ejercicios (con nota  $> 0$ )** tiene que presentarse al examen práctico y examen final completo
- Si un alumno no se presenta a la **prueba intermedia (con nota  $> 0$ )** tiene que presentarse al examen práctico y examen final completo

# Fecha de la prueba intermedia

- Viernes 18 de noviembre

# Fechas de los exámenes finales

- Viernes 13 de enero (tarde)
- Lunes 3 de julio (tarde)



## Nota:

- jueves 13/10/2022 horario de **martes**
- viernes 4/11/2022 horario de **martes**

Importante:

No hay clase el viernes 16 de diciembre

# Grupos de prácticas y problemas

- Hay que apuntarse a un grupo de prácticas (tiene que ser el mismo grupo al que se apunta la pareja de prácticas)
- Hay que apuntarse a un grupo de problemas (cada viernes diré si nos dividimos en dos el siguiente lunes, **si no nos dividimos la clase será el lunes de 12 a 12:50h**)
- Podéis apuntaros a través de moodle a partir de hoy viernes a las 12:30