Geometry in the Computer:

From polygons and museums to molecules and massive data

Ioannis Emiris

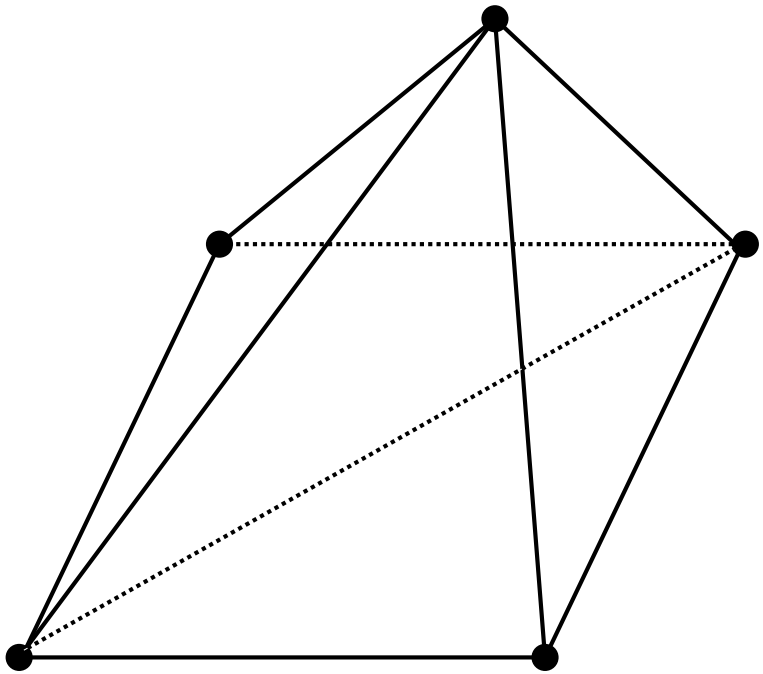
http://erga.di.uoa.gr

Dept Informatics & Telecoms, U.Athens

November 7, 2012

Overview

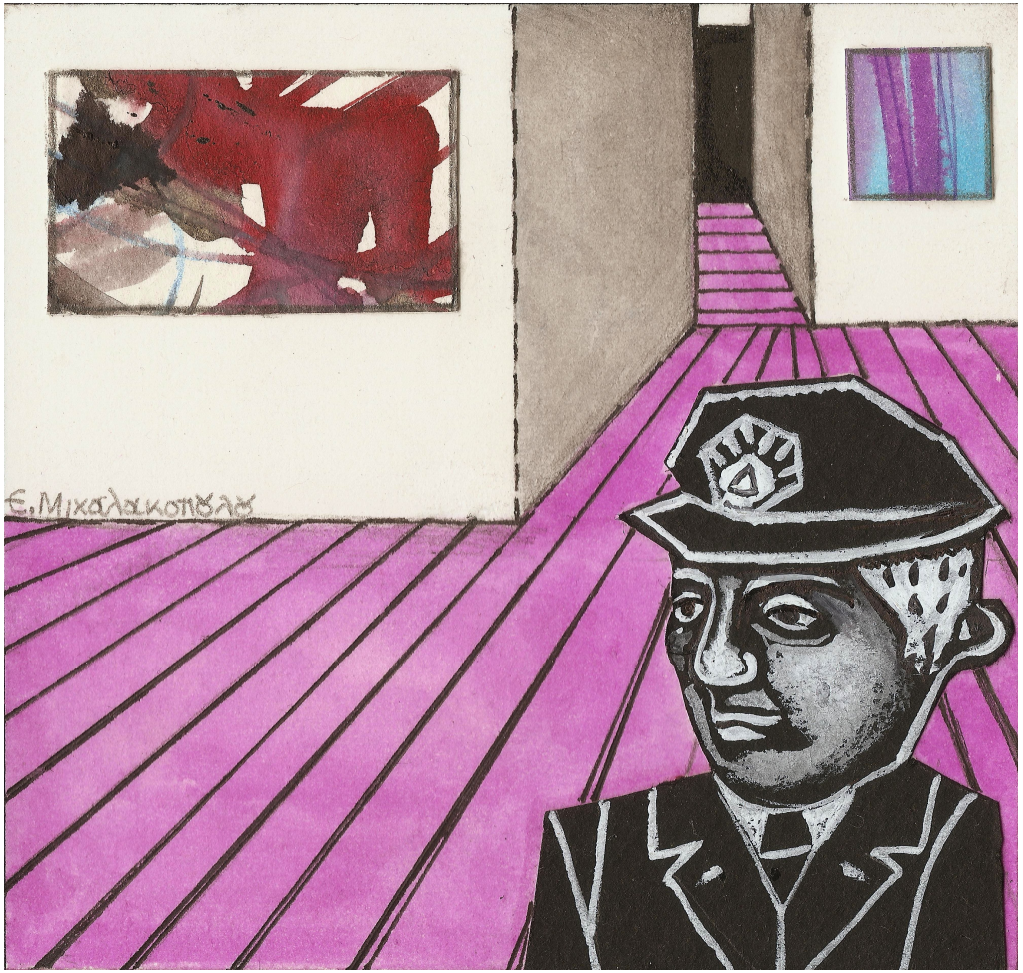
Antiquity



Continuous vs Discrete

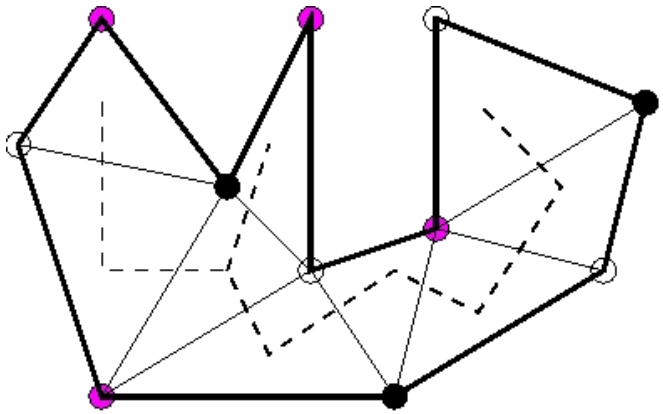
Outline

Museum Guards



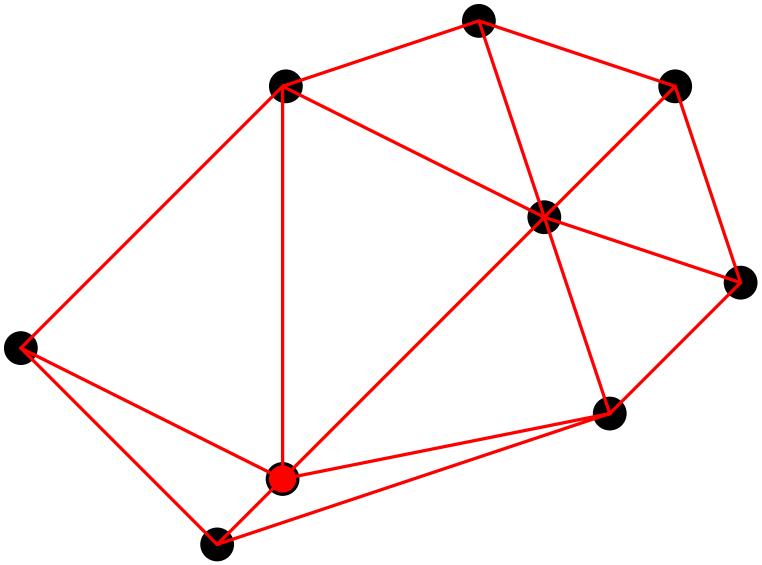
Art gallery Theorems: *a classic*

- Minimizing the number of point-guards to cover a simple polygon is NP-hard.
- In general, $\lfloor n/3 \rfloor$ guards cover a n-gon.
- The main algorithmic step is triangulation. It is in $O(n)$: deterministic [Chazelle], or simpler randomized algorithm.



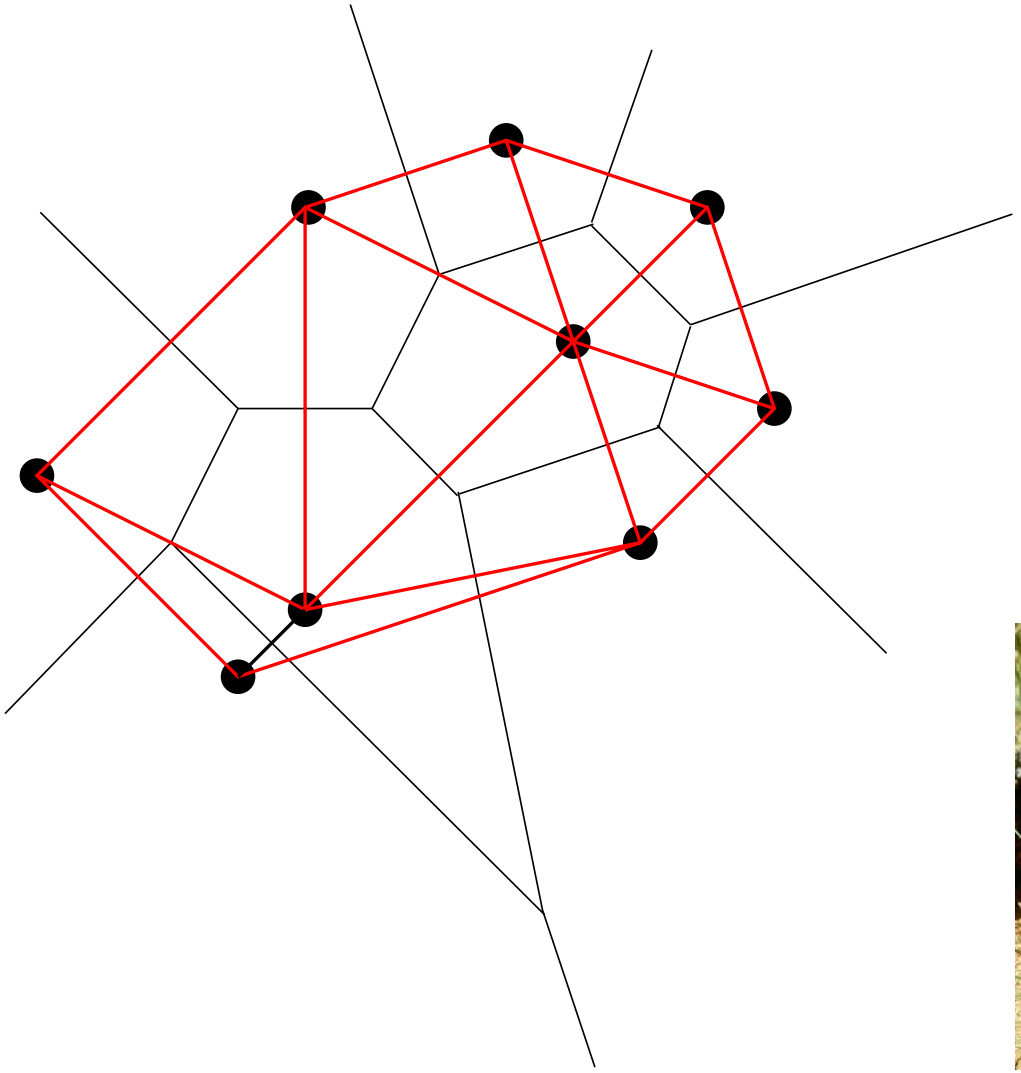
- Applications: visibility, telecom networks, extension to 3D.

The best triangulation ...



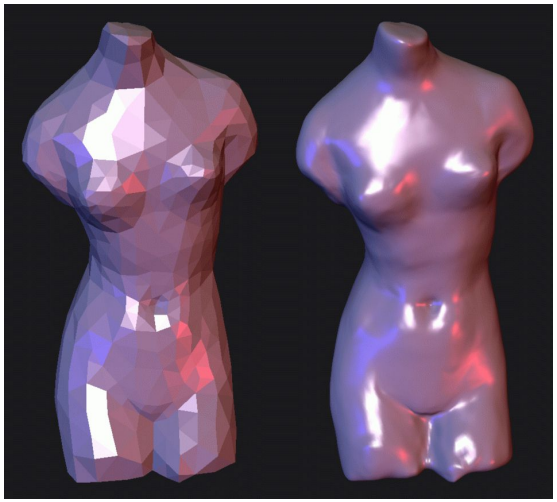
Boris Delaunay [1934]

... equivalent to a distance partition

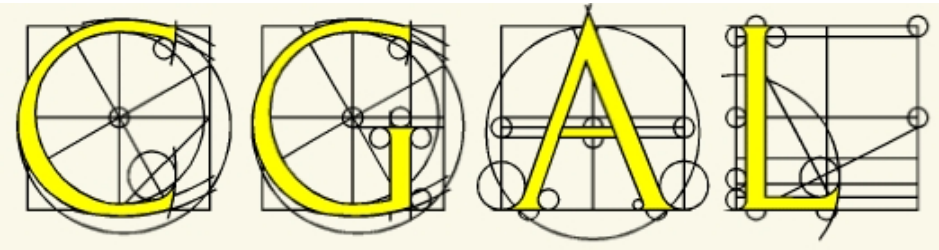


Georgy Voronoi [1908]

Graphics

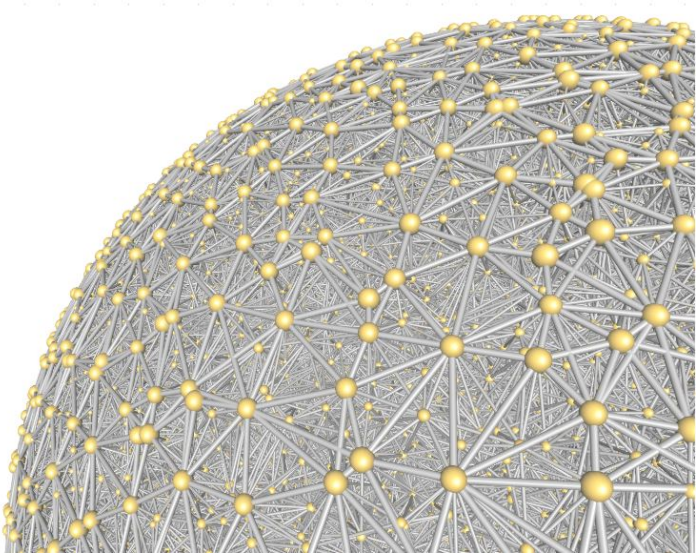


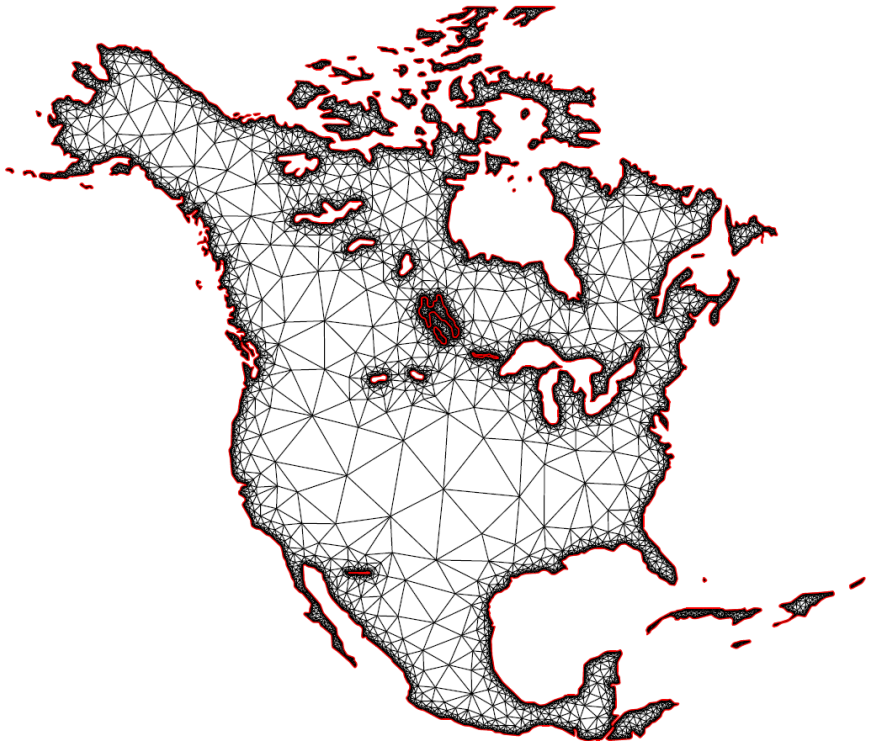
Computational Geometry Algorithms Library



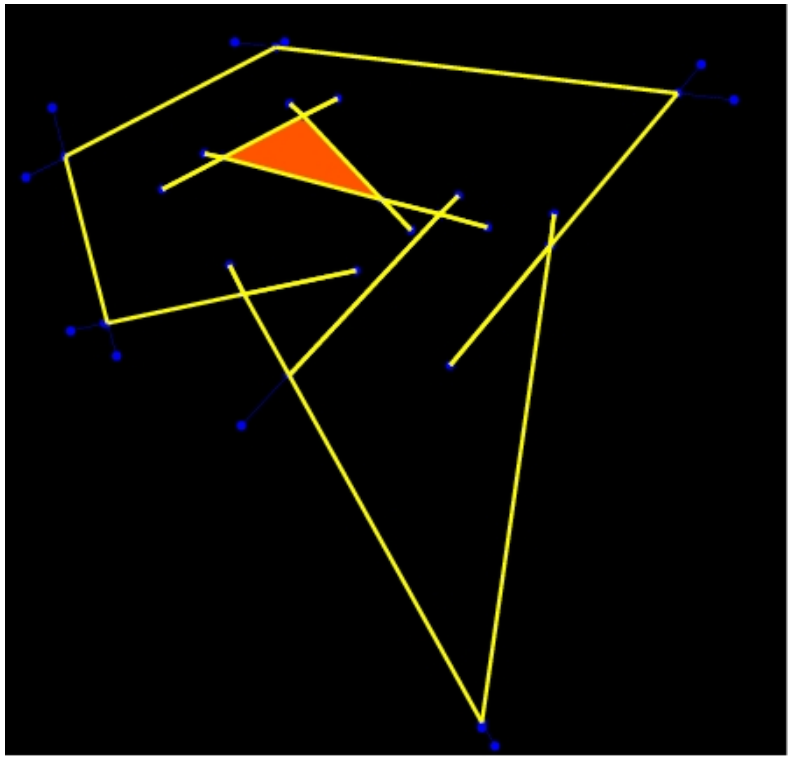
- Efficient algorithms
- Exact computation
- C++ / STL library
- Open project, mostly European

CGAL triangulation / mesh





CGAL convex hull / arrangement

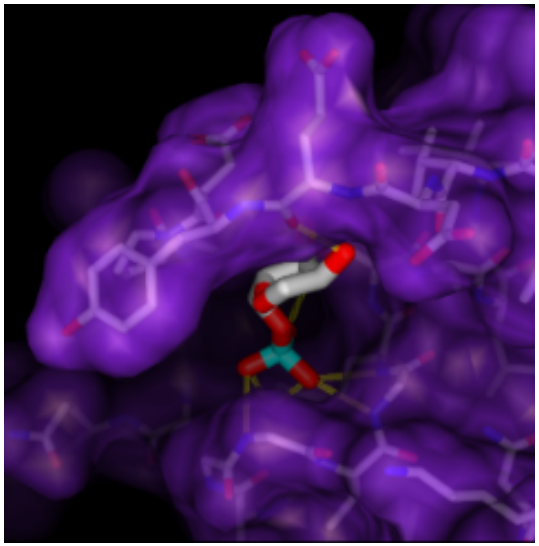


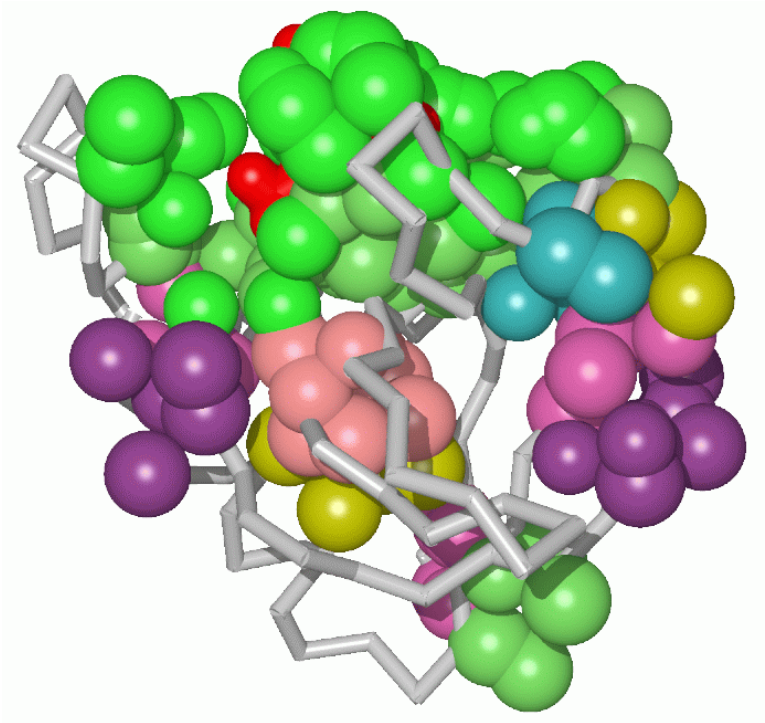
Outline

Dogma

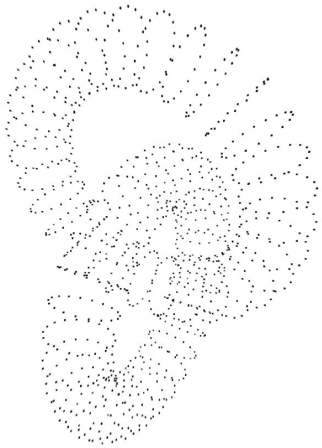
aminoacid sequence $\longrightarrow$ *folding* **3-dim structure** $\longrightarrow$ *"surface"* **chemical function**

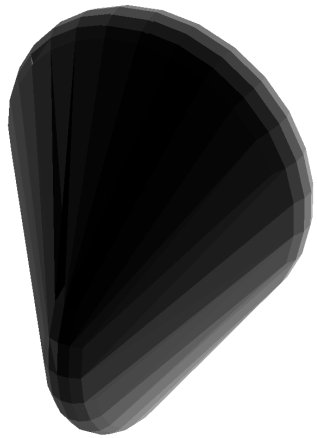
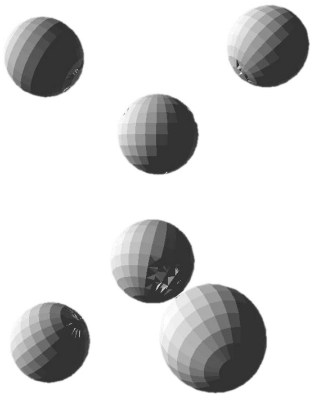
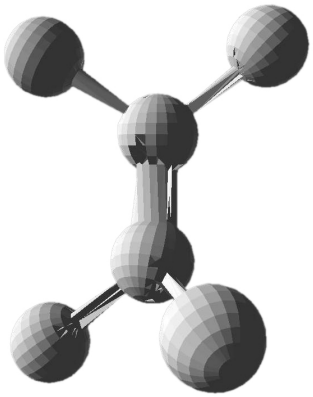
Docking



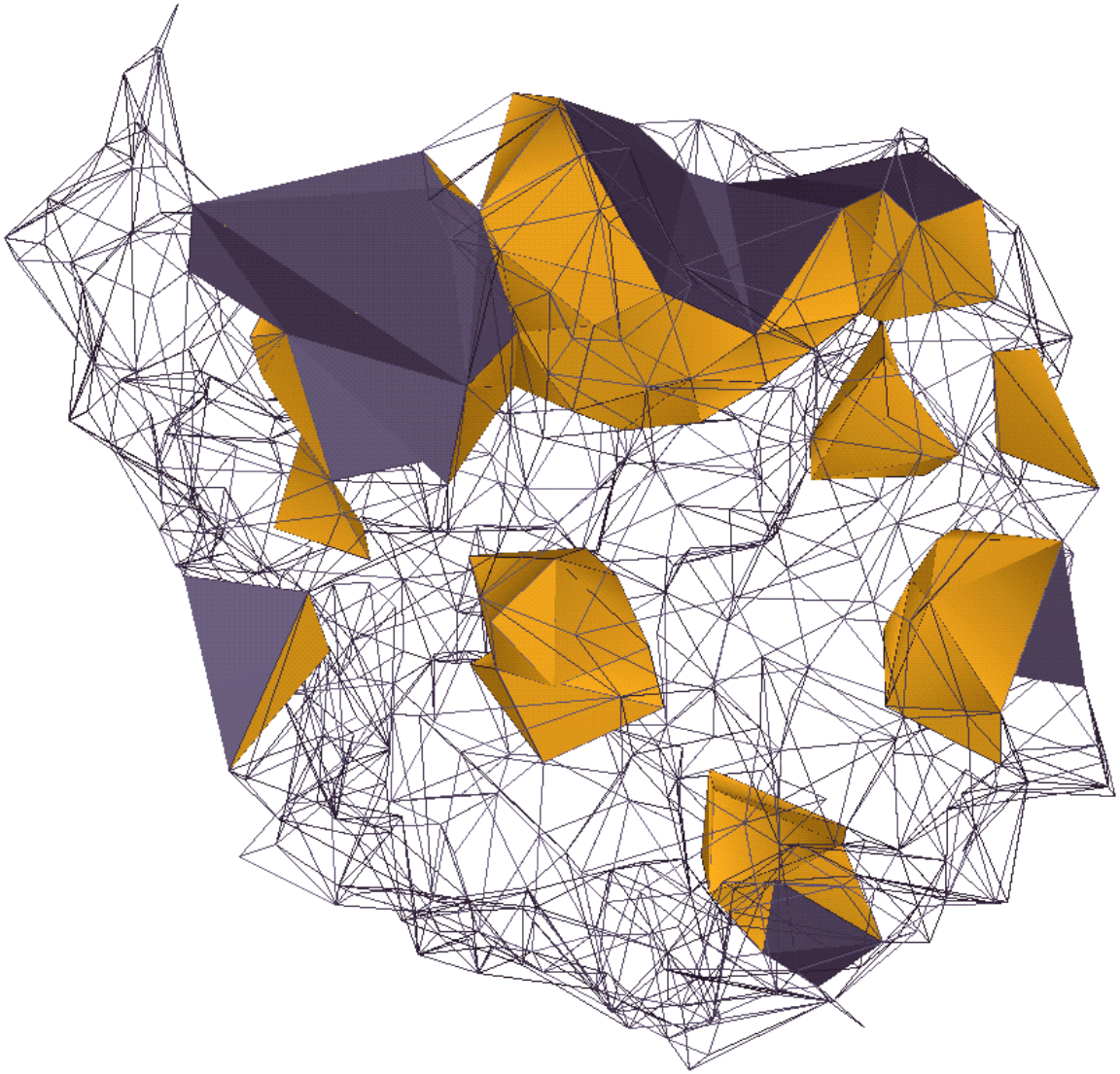


α-shapes: *success story and major tool in bioinformatics*

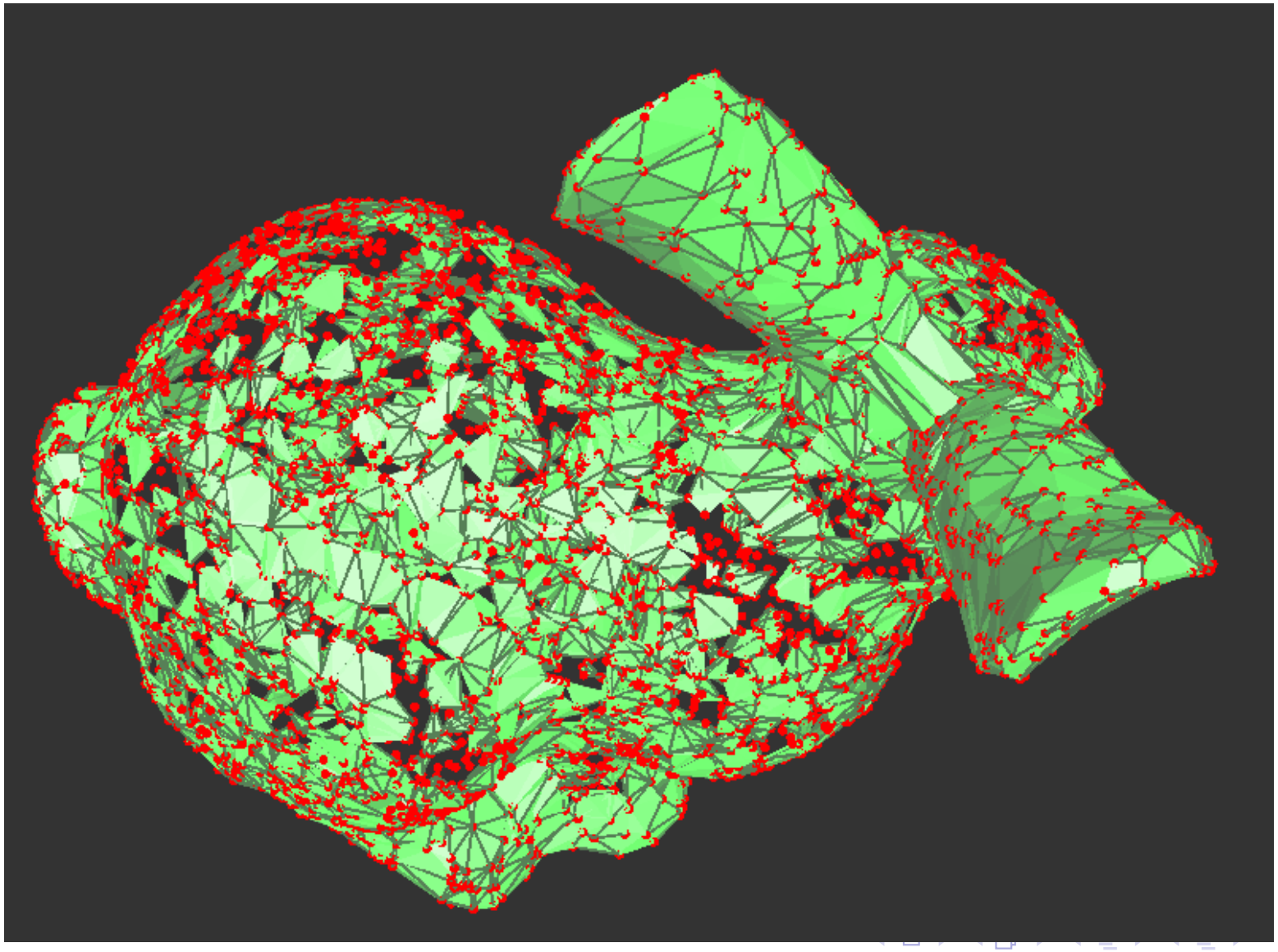


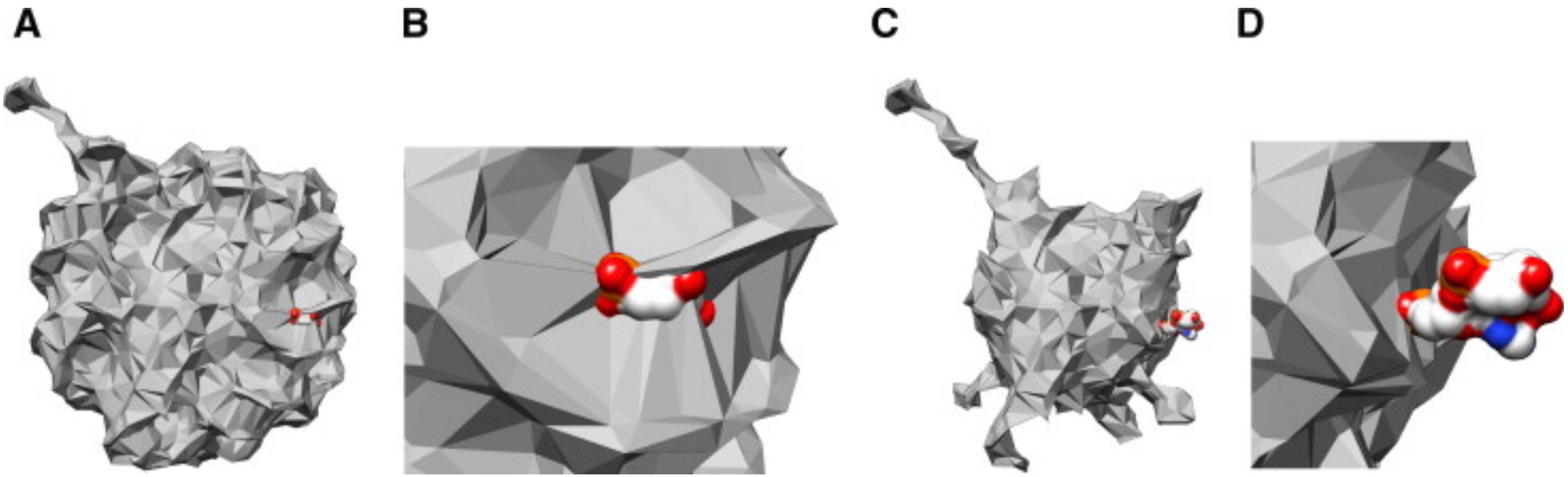

3D structure



CGAL



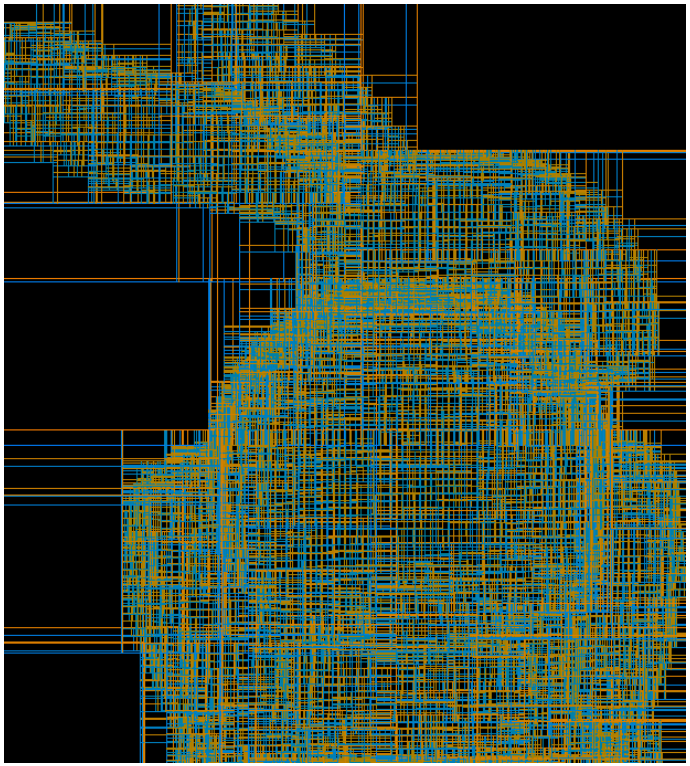
Different α

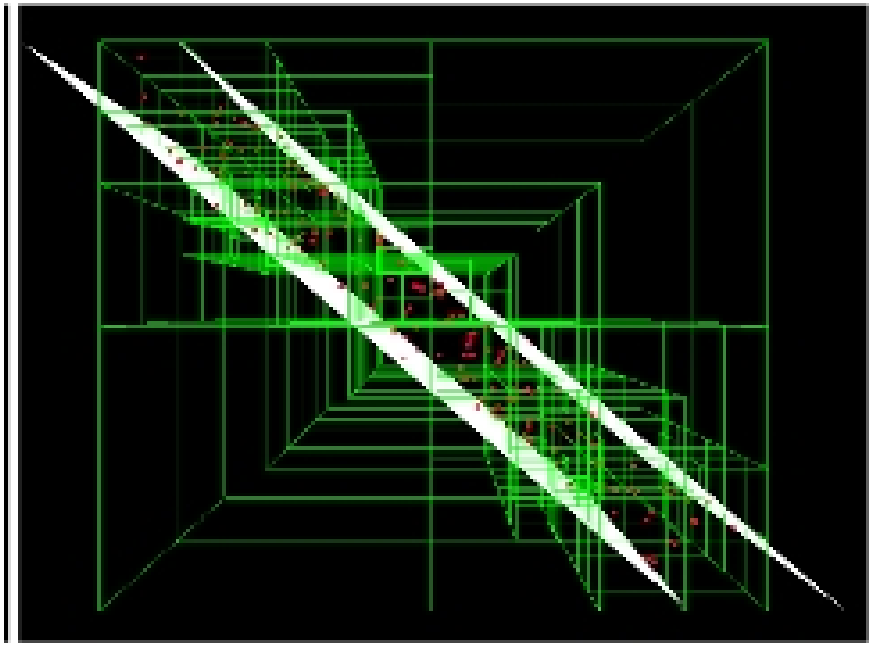


[Edelsbrunner]

Outline

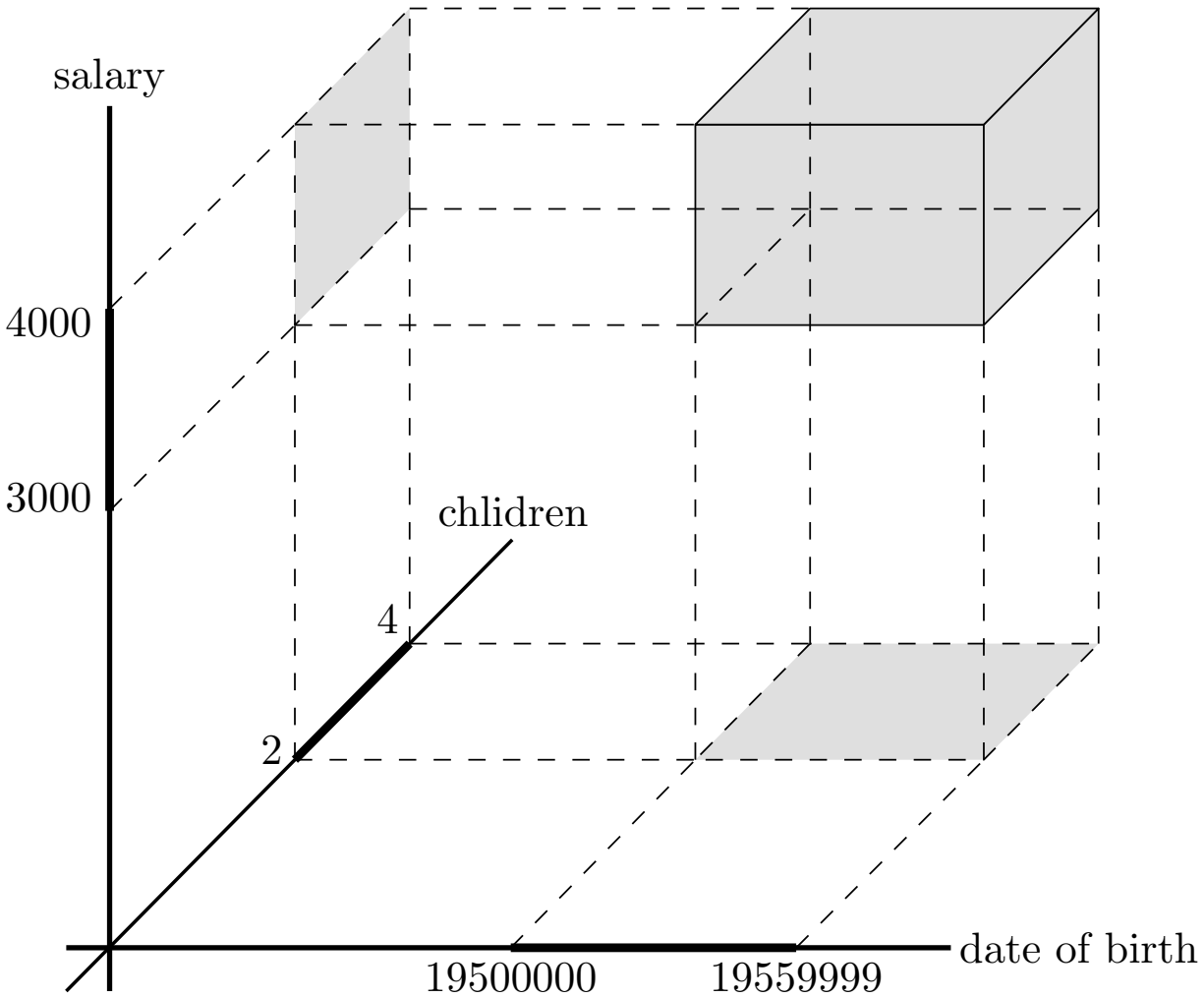
3D



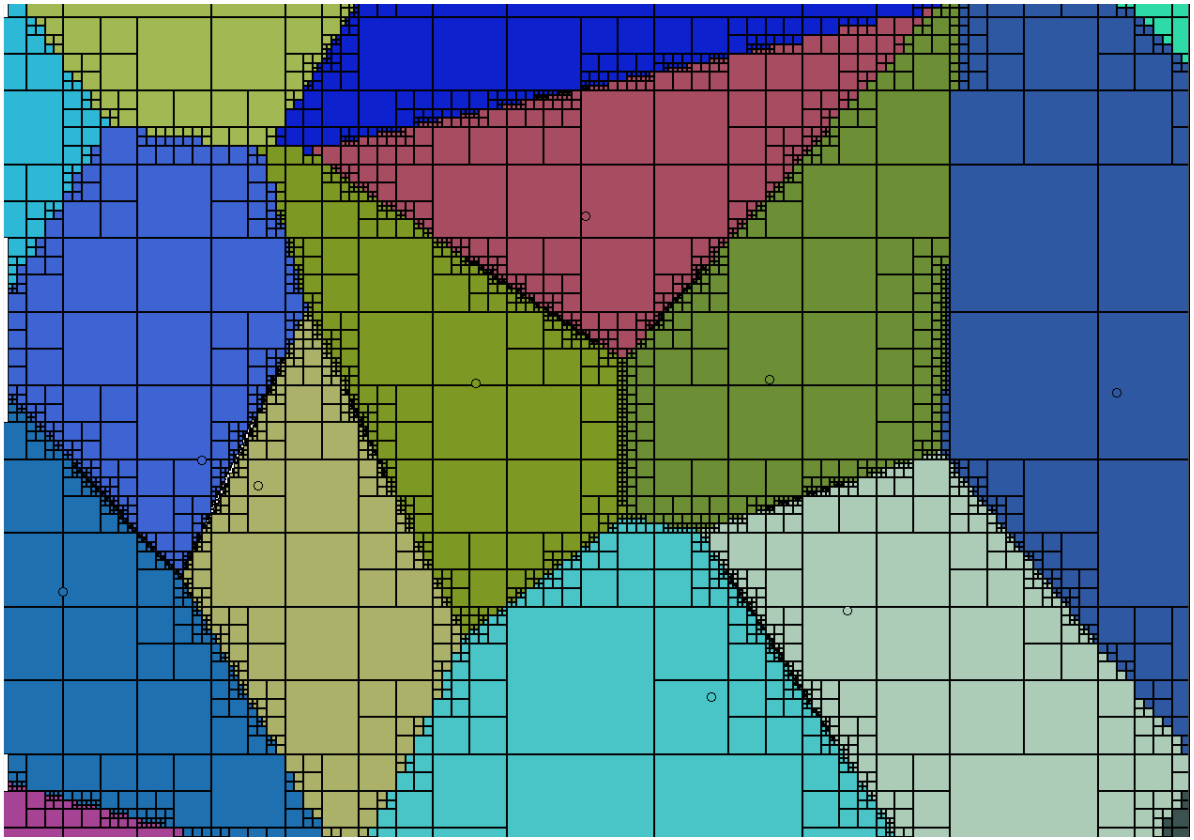


http://www.youtube.com/watch?v=OqVNQmX_J28

Databases

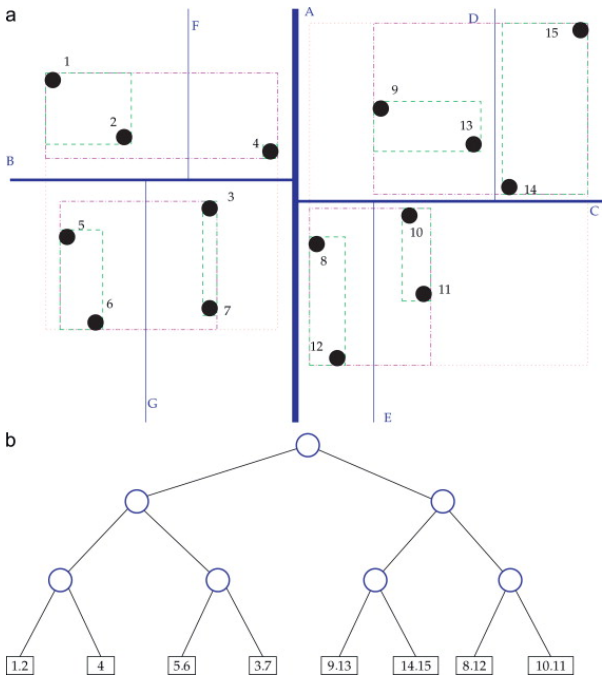

salary
4000
3000
chlidren
4
2
19500000
19559999
date of birth

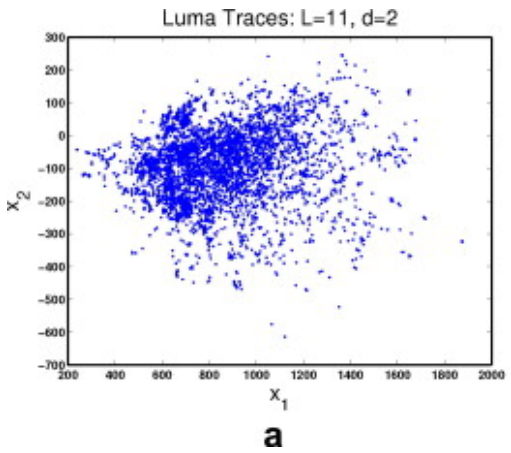
Approximate Voronoi diagram

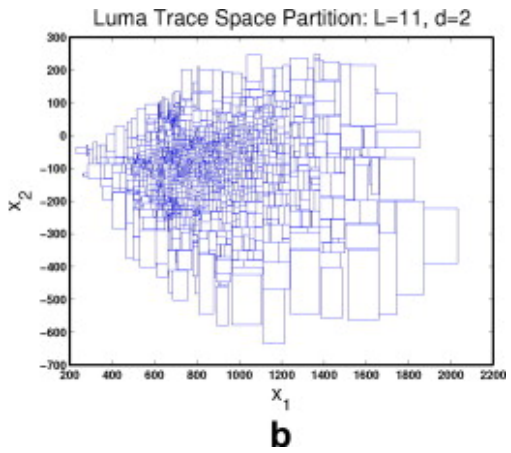


https://github.com/akonskarm/AVD

Data structures: *towards data mining*







ϵ-nearest neighbor in msec among 1M points in > 100 dimensions.

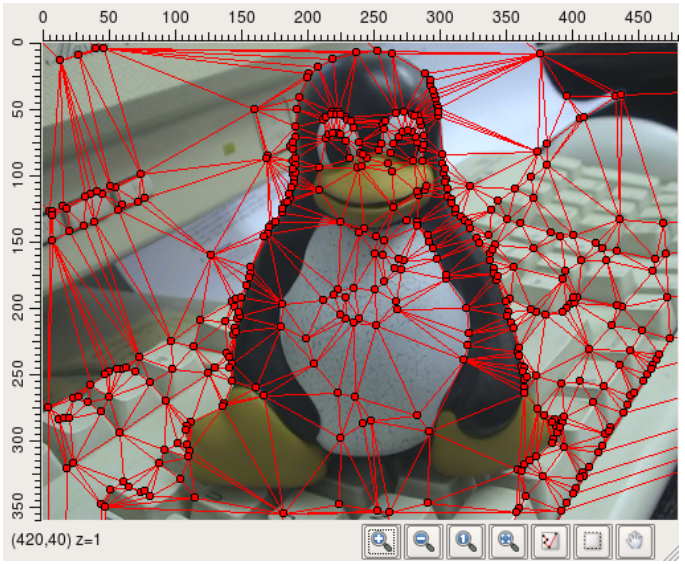
Curse of dimensionality

Tree based methods, e.g. AVD:
For fixed d, space $= O(n/\epsilon^d)$, ϵ-nearest neighbor in $O(\log(n/\epsilon))$.
[Mount, Har-Peled, et al.]

Locality sensitive hashing:
polynomial-time in d, exponential in ϵ.
[Indyk, Motwani, et al.]

High dimensions: *towards structure*

Bioinformatics, the Internet, GIS, Image processing:
they all easily provide data in e.g. 100, or 1000 dimensions.



Current goal: Exploit structure [Guibas]

And then?


Υπολογιστική
Γεωμετρία


ΕΚΔΟΣΕΙΣ ΚΛΕΙΔΑΡΙΘΜΟΣ
ΥΠΟΛΟΓΙΣΤΙΚΗ
ΓΕΩΜΕΤΡΙΑ
Μια σύγχρονη
αλγοριθμική προσέγγιση
Γιάννης Ζ. Εμίρης