MATH 529: Lecture 1 (01/13/2026)

Today: * syllabus, logistics
* two motivating applications

Call me Bala!

Introduction to **Computational** topology focus for this course

This course will be offered completely electronically:
- scribes will be posted as "course notes"; videos will also be posted.
- assignments to be turned in electronically
(you could submit scanned versions of handwritten assignments).
- web page has all the docs/info.

Topology "Topo" $\to$ place or space
"logos" $\to$ study or talk } in Greek

Topology talks about how space is connected.

topology
- point set topology (open/closed, connected, ...)
- algebraic topology (groups, addition, basis, ...)

We will concentrate on algebraic topology.

Computational topology

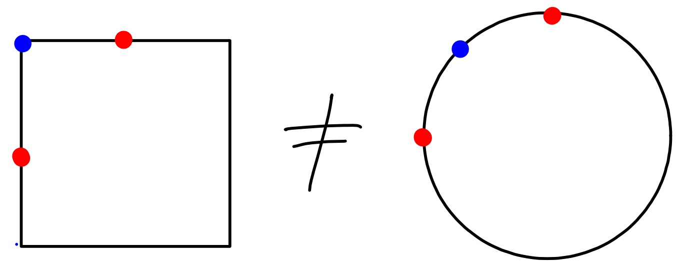
→ combine efficient algorithms and data structures with results from topology to analyze real-life data.

Let's start with an intuition for what we mean by connectivity of spaces.

An Example

According to geometry, the square and circle are not equal.

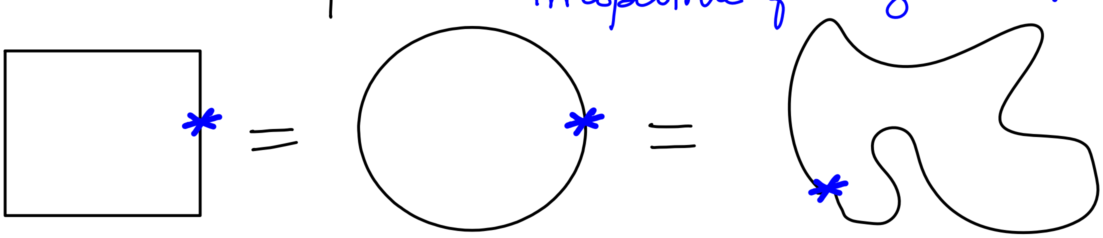
Size (length, here) is critical in geometry, but not so much in topology.



But topology says they are same as far as how they are connected!

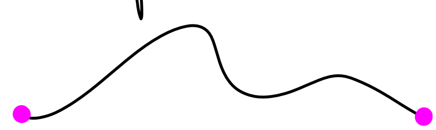
For instance, take a string, and tie a knot to make a loop.

We want to study connectivity irrespective of size here!



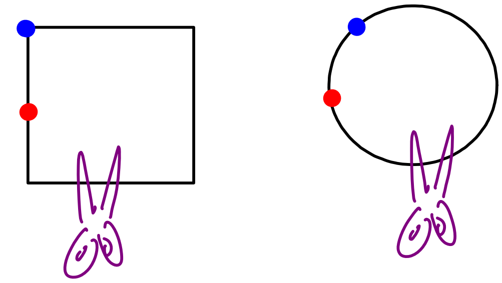
All these objects are same, i.e., they are connected the same way.

But, if you did not tie the knot, the loose string (open thread) differs from any of the above tied loops in connectivity.

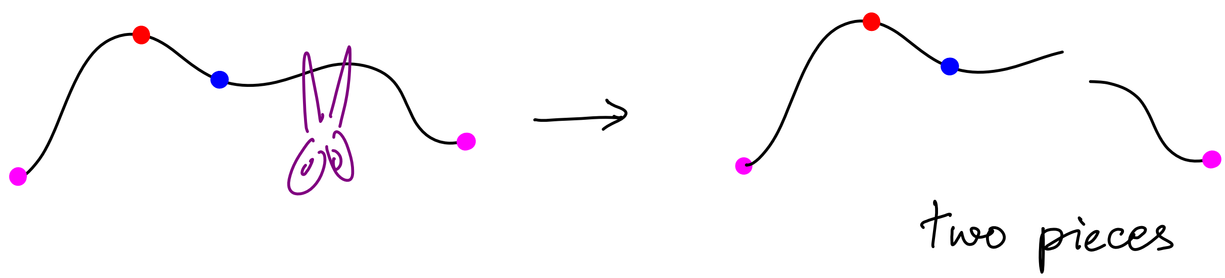


Note that the two end points have different "neighborhoods."

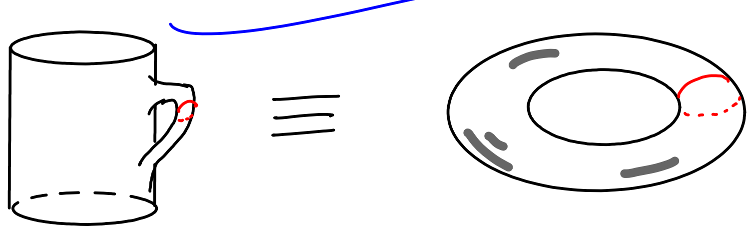
Here is another way to understand connectivity. Consider cutting the string (tied into a loop) once. Such a cut leaves the string in one piece, i.e., connected.



But cutting the open thread once leaves two pieces, i.e., it is disconnected.



A popular quote: "A topologist cannot distinguish the coffee cup from a donut!"



the coffee cup, if made of clay, could be deformed to look like the donut, without changing its connectivity.

A more practical example: how we are able to read (recognize) letters of the alphabet in different fonts.

A **A** $\mathcal{A}$ $\neq$ B **B** $\mathcal{B}$ $\neq$ C C **C**

Two Illustrations of Computational Topology

1. Patient antibiotic trajectories

https://doi.org/10.1145/3307339.3342143

? How do agents and doses affect length of stay?

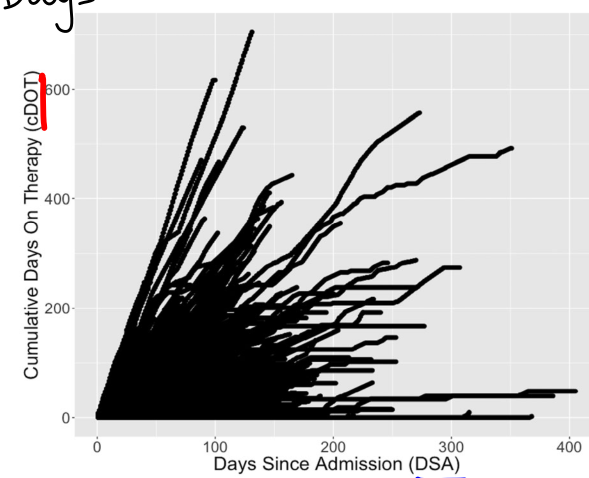
Number of hospitals	25
Number of hospital unit-categories	9
Number of distinct patient-admission records	349, 610
Number of adult patients	334, 207
Number of male patients	148, 540
Number of female patients	201, 052
Average LOS per admission	7 days
Longest LOS → length of stay	405 days
Number of antibacterials used	66
Most used antibacterial	Vancomycin
Average DOT per admission	6
Number of agent ranks	4
Most used agent rank	rank 3

DOT → Days On Therapy

If a patient gets one dose of an agent (antibiotic) that is counted as 1 Day On Therapy (DOT).

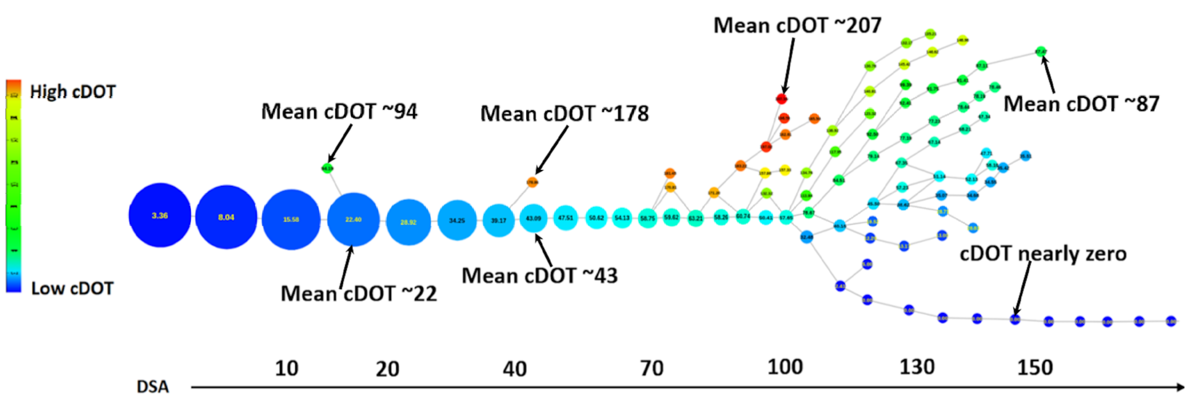
cDOT : cumulative Days On Therapy

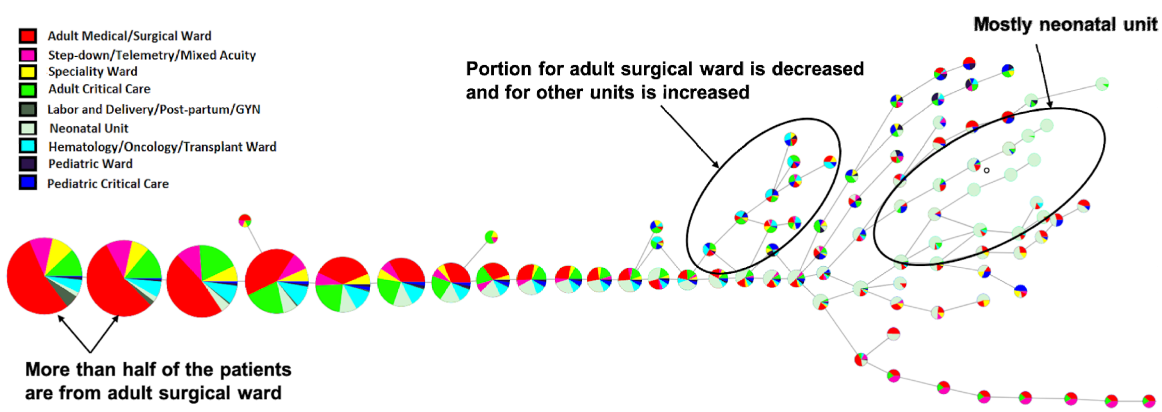
DSA : Days Since Admission.



This plot by itself is not very informative or insightful (even if we were to use color...)

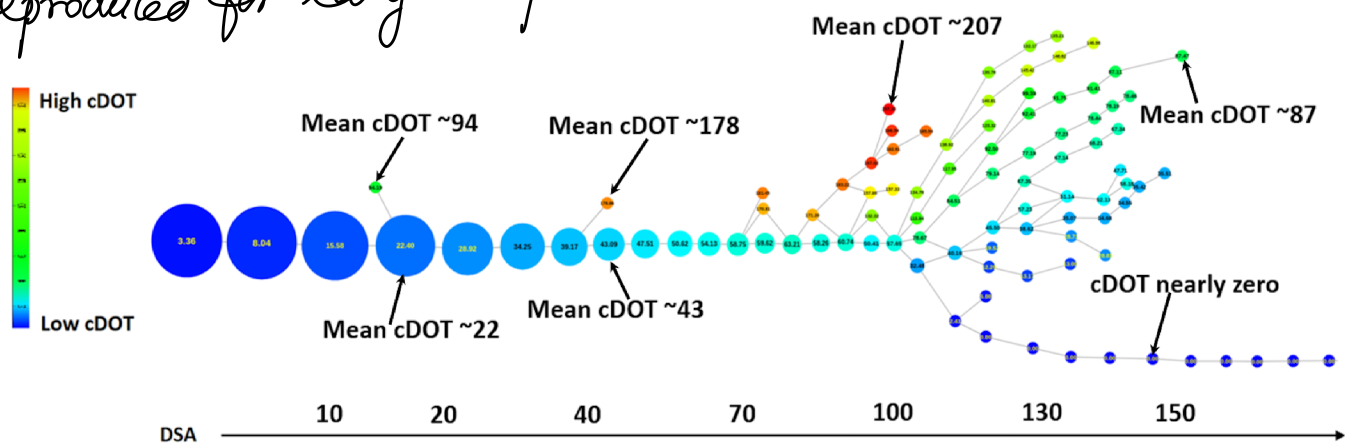
Here are two versions of a Mapper representation of the same data:

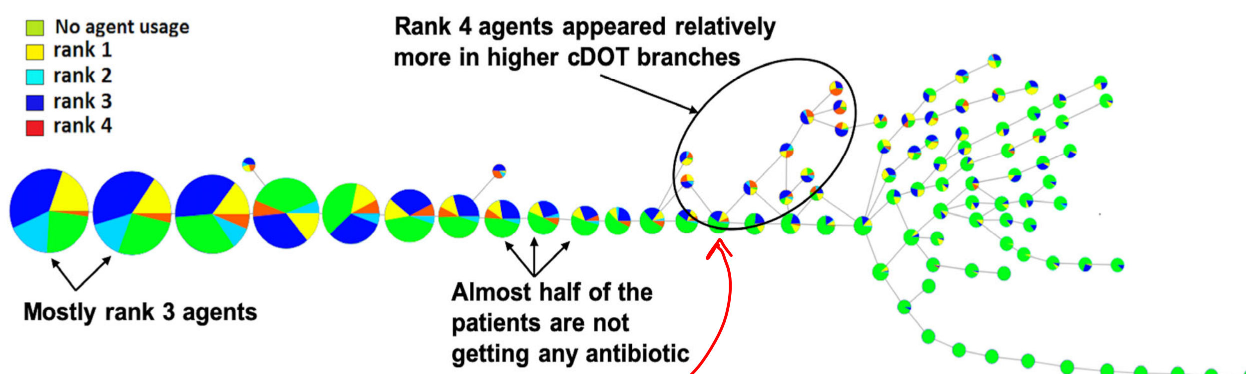




Each node represents a cluster of patient trajectories. A lot of the patients got low cDOTs and they had small(er) DSAs — as captured by the big clusters on the left. As the second Mapper shows, many of these patients were treated in the adult surgical ward — one of the most common types of admissions to hospitals.

Here is another version of the Mapper showing ranks (1–4, 4 is strongest) of the antibiotics. The first mapper (using cDOT) is reproduced for easy comparison.



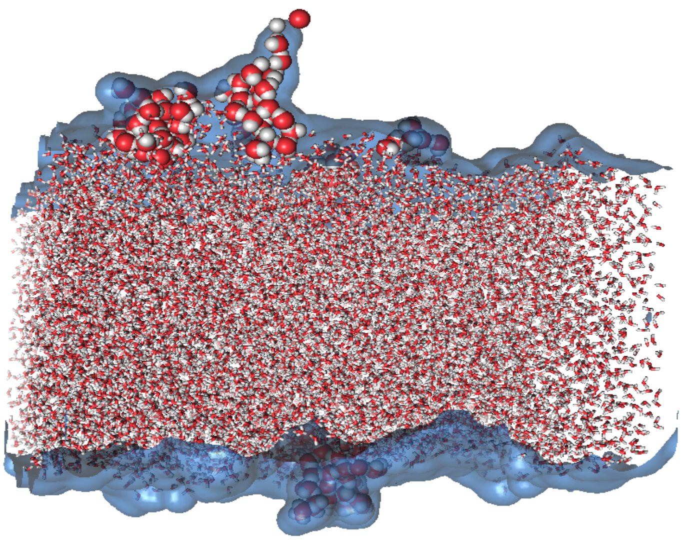


The high cDOT + high rank sub-branches had more patients in the other (higher risk) wards. Similarly, the much higher DSA group (120+) on the right end with relatively smaller cDOT values turned out to be patients in neonatal ICUs.

Note that these nontrivial subgroups are identified in an unsupervised manner — no learning is involved!

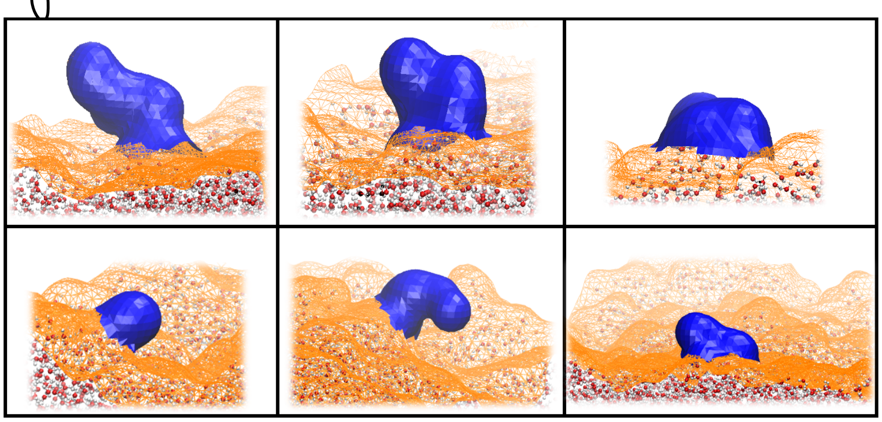
2. Interface features in Chemistry

https://doi.org/10.1021/acs.jctc.0c00260 (https://doi.org/10.26434/chemrxiv.11988048.v1) → preprint



An interface surface separates a water layer from a hexane (organic) layer. When a reagent is added, the reaction is initiated and the water molecules escape to the hexane layer through finger-like features in the interface called "protrusions". These features were identified manually (by observation!)

The goal was to identify and characterize protrusions using geometric measure theory and computational topology.



Which of these six features do you think are protrusions?

It is not easy to guess!

We use the notion of multiscale flat norm of surface T:

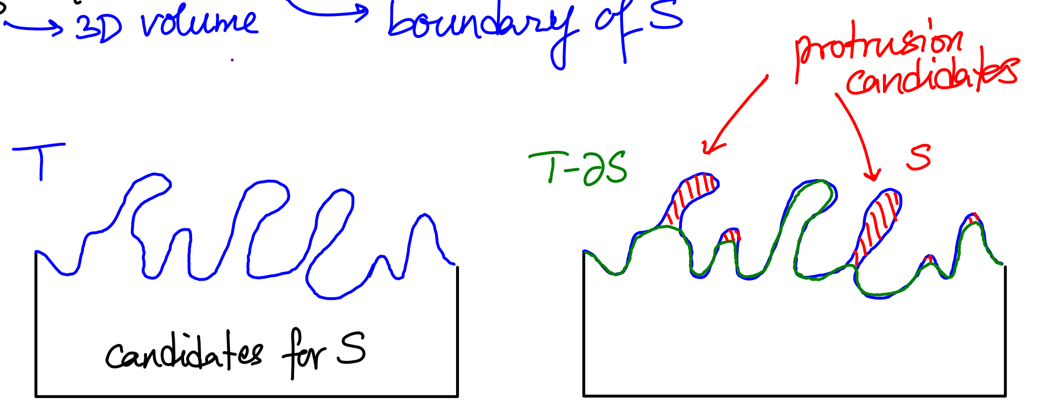
$$F_\lambda(T) = \min_{S} \{ \text{Area}(T - \partial S) + \lambda \, \text{Volume}(S) \}, \quad \lambda \geq 0$$

$\lambda \to$ scale parameter

$S \to$ 3D volume

$\partial S \to$ boundary of S

Illustration in 2D:

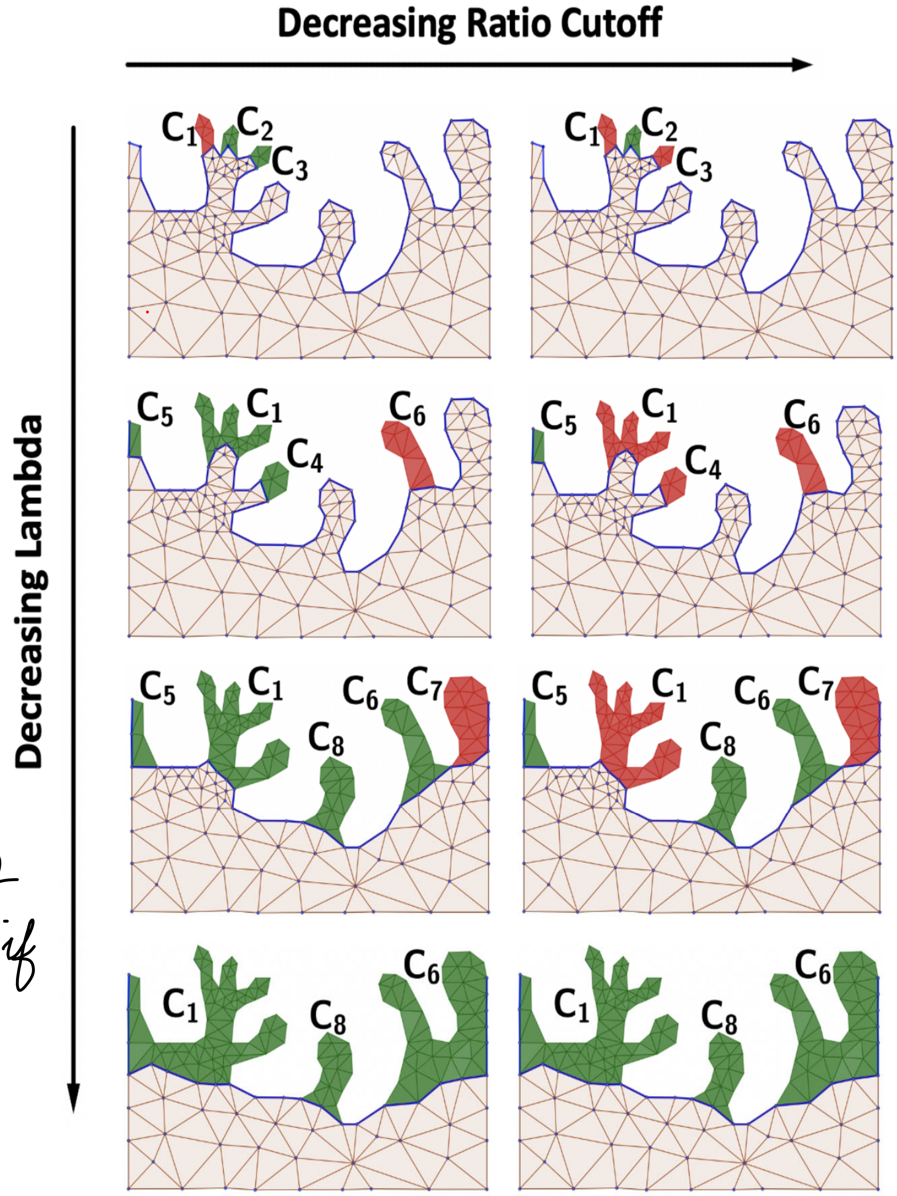


We keep track of connected components in S as $\lambda \downarrow$. We relabel them and also keep track of merging behavior.

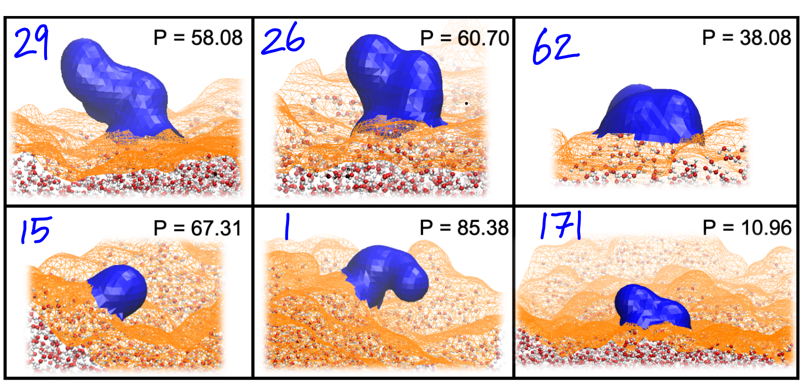
We also track the ratio $\frac{Vol(C)}{Vol(B(1/\lambda))}$ for each component C, where $B(1/\lambda)$ is the ball with radius $1/\lambda$ and $Vol(C)$ is the volume of component C.

A component C is "alive" at ratio cutoff r and scale λ if

$$\frac{Vol(C)}{Vol(B(1/\lambda))} > r.$$



The longer a component is alive, the more likely it is to be a protrusion.



It turned out that all six of these features were protrusions! The probabilities (as %'s) along with their ranks among 195 candidate features (lower rank $\Rightarrow$ more likely to be a protrusion) are shown here.

While this example is described in 3D, the underlying concepts are more general, and in fact generate certain key fundamental questions in geometric measure theory (GMT).

In fact, when we talk about applied algebraic topology the application could be to pure mathematics! We will talk about this aspect toward the end of the semester.

Note that we are showing a discrete version of the surface—in the form of a triangular mesh. Indeed, we need to discretize continuous spaces to perform computations!

Here is a notion of connectivity in the "discrete setting":

The neighborhood, i.e., the set of nearby points, of the two • points are different – they each have only one neighbor, while the • points all have two neighbors each.