COMPUTATIONAL TOPOLOGY
AN INTRODUCTION
Herbert Edelsbrunner and John Harer

Michal Kleinbort

Tel Aviv University, July 2013
Chapter I - Graphs

- **Graphs in Topology:**
  1-dimensional geometric objects where:
  - vertices = points
  - edges = curves connecting pairs of points

- **Graphs in Discrete-Math:**
  - vertices = abstract elements
  - edges = pairs of these abstract elements
Chapter I - Graphs

Topics:

- Connected Components
- Curves in the Plane
- Knots and Links
- Planar Graphs
Connected Components

Comparing the notion of connectedness in:
- Discrete graphs
- Continuous spaces
Simple Graphs

- An abstract graph is a pair \( G = (V, E) \).
- **Simple graph** - no two edges connect the same two vertices and no self loops.
- If \( G = (V, E) \) is a simple graph and \(|V| = n\) then \( m = |E| \leq \binom{n}{2} \).
- Let \( u = u_0, u_1, ..., u_k = v \) be a **path** between vertices \( u \) and \( v \), of length \( k \).
- A path can cross itself or backtrack.
- In a **simple path** the vertices are distinct.
Connectedness in Simple Graphs

- A simple graph is **connected** if there is a path between every pair of vertices.
- A **connected component** is a maximal subgraph that is connected.
- **Trees** are the smallest connected graphs (unique simple path between every pair of vertices).
- **Spanning tree** of $G = (V, E)$ is a tree $(V, T)$ where $T \subseteq E$ (requires a connected graph).
- A simple graph is connected iff it has a spanning tree.
- A **separation** is a non trivial partition of the vertices $V$ to two disjoint, non-empty sets $U, W$, such that no edge connects a vertex in $U$ with a vertex in $W$.
- A simple graph is connected if it has no separation.
Metric Spaces

What is a metric?

A metric \( d : X \times X \to \mathbb{R} \) on a set \( X \) is a function with the following properties:

- \( d(x, y) \geq 0 \) for all \( x, y \in X \)
- \( d(x, y) = 0 \) iff \( x = y \)
- \( d(x, y) = d(y, x) \) for all \( x, y \in X \)
- (Triangle inequality) \( d(x, y) + d(y, z) \geq d(x, z) \) for all \( x, y, z \in X \)

\( (X, d) \) is a metric space

Examples: the Euclidean metric, the discrete metric, the Hamming (counting) metric
Open Sets in Metric Spaces

- Let \((X, d)\) be a metric space
- Let \(B_d(x, \delta) = \{y | d(x, y) < \delta\}\)
- A subset \(U \subseteq X\) is **open** in the metric space \((X, d)\) iff for each \(y \in U\) there is a \(\delta > 0\) such that \(B_d(y, \delta) \subset U\)
Properties of Open Sets in Metric Spaces

If \((X, d)\) is a metric space then the following statements are true:

- \(\emptyset\) and \(X\) are open
- The union of any number (could be uncountable) of open subsets of \(X\) is open
- The intersection of a finite number of open subsets of \(X\) is open

Why is the intersection limited to a finite number of open sets?

Suppose we are working in \(\mathbb{R}^n\) with the Euclidean metric, then \(B(x, 1/j)\) is open, but \(\bigcap_{j=1}^{\infty} B(x, 1/j) = \{x\}\) is not!
Closed Sets in Metric Spaces

- A subset $U \subseteq X$ of a metric space $(X, d)$ is **closed** if the set $X - U$ is open.

- **Examples:**
  - The subset $[a, b]$ of $\mathbb{R}$ is closed because its complement $\mathbb{R} - [a, b] = (-\infty, a) \cup (b, +\infty)$ is open.
  - $\emptyset$ and $X$ are closed.
Properties of Closed Sets in Metric Spaces

(Deduced from the properties of open sets in metric spaces)

If \((X, d)\) is a metric space then the following statements are true:

- \(\emptyset\) and \(X\) are closed
- The intersection of any number (could be uncountable) of closed subsets of \(X\) is closed
- The union of a finite number of closed subsets of \(X\) is closed

What happens if we take a union of any number of closed sets?

We cannot say it is closed!

Suppose we are working in \(\mathbb{R}\) with the Euclidean metric then the interval \([1/i, 1]\) is closed, but \(\bigcup_{i=1}^{\infty} [1/i, 1] = (0, 1]\) is not!
Properties of Closed Sets in Metric Spaces

These properties of closed sets can be deduced from the properties of open sets in metric spaces.

For instance, let’s prove the last property, that is prove that: if $F_j$ is closed for all $1 \leq j \leq n$, then $\bigcup_{j=1}^{n} F_j$ is closed (the last bullet):

Since $F_j$ is closed then $X \setminus F_j$ is open for all $1 \leq j \leq n$. It follows that $X \setminus \bigcup_{j=1}^{n} F_j = \bigcap_{j=1}^{n} X \setminus F_j$ is open, and so $\bigcup_{j=1}^{n} F_j$ is closed.
Topological Spaces - The Continuous Model of Reality

Given a points set, the topology defines which points are near without specifying how near they are from each other.

Metric spaces are an example for topological space.
Topological Spaces - Formal Definition

A **topology** on a set $X$ is a collection $\mathcal{T}$ of subsets of $X$ having the following properties:

- $\emptyset$ and $X$ are in $\mathcal{T}$
- The union of the elements of any subcollection of $\mathcal{T}$ is in $\mathcal{T}$.
- The intersection of the elements of any finite subcollection of $\mathcal{T}$ is in $\mathcal{T}$.

A topological space is an ordered pair $(X, \mathcal{T})$ consisting of a set $X$ and a topology $\mathcal{T}$ on $X$. 
Open & Closed Sets in Topological Spaces

- If $(X, d)$ is a metric space we call the collection of open sets the topology \textit{induced} by the metric.
- If $(X, \mathcal{T})$ is a topological space we extend the notion of open set by calling the members of $\mathcal{T}$ open sets.
- In other words, if $(X, \mathcal{T})$ is a topological space a set $U \subseteq X$ is an \textbf{open set} if $U \in \mathcal{T}$.
- The definition for closed sets remains.
Example

Let $X = \{a, b, c\}$ be a three-element set. Possible topologies on $X$:

However, not any collection of subsets of $X$ is a topology on $X$:
Basis for a Topology

In the previous example we specified a topology by describing the entire collection $\mathcal{T}$ of open sets.

This is too difficult!

Instead, one can specify a smaller collection of subsets of $X$ and define the topology in terms of that.
Basis for a Topology - Definition

If \( X \) is a set, a **basis** for a topology on \( X \) is a collection \( \mathcal{B} \) of subsets of \( X \), called **basis elements**, such that:

- For each \( x \in X \), there is at least one basis element \( B \in \mathcal{B} \) containing \( x \)
- If \( x \in B_1 \cap B_2 \), then there is a basis element \( B_3 \) such that \( x \in B_3 \subset B_1 \cap B_2 \)

The topology \( \mathcal{T} \) generated by \( \mathcal{B} \) consists of all sets \( U \subseteq X \) for which \( x \in U \) implies there is a basis element \( B \in \mathcal{B} \) such that \( x \in B \subseteq U \) (each basis element is an element of \( \mathcal{T} \))

How to construct the topology explicitly? take all possible unions of all possible finite intersections of basis sets
Example

Consider the real line, $X = \mathbb{R}$, and let $\mathcal{B}$ be the collection of open intervals. This gives the usual topology of the real line.

Note that the intersection of the intervals $(-1/k, +1/k)$, for the infinitely many integers $k \geq 1$ is the point 0, which is not an open set.

Illustrating the restriction to finite intersections...
Options for Defining the Topology

So far, we mentioned three options for defining the topology of a set $X$:

- Specifying the entire collection of open sets
- Specifying the entire collection of closed sets
- Specifying the collection of basis elements

...and there are other options as well
Subspace Topology - Definition

If $\mathcal{T}$ is a topology of $X$ and $Y \subseteq X$, then the collection of sets $Y \cap U$, for $U \in \mathcal{T}$ is the **subspace topology** of $Y$.

Example:
Consider the closed interval $[0, 1] \subseteq \mathbb{R}$. In the previous example we saw that the open intervals form the basis for a topology of the real line. The intersections of open intervals with $[0, 1]$ form the basis of the subspace topology of the closed interval. Note that $(1/2, 1]$ is considered an open set in $[0, 1]$, but is not open in $\mathbb{R}$. 
Continuity - Definition

Let $X$ and $Y$ be topological spaces. A function $f : X \to Y$ is continuous if for each open subset $V \subseteq Y$, the set $f^{-1}(V)$ is an open subset of $X$.

Reminder: $f^{-1}(V)$ is the set of all points $x \in X$ for which $f(x) \in V$.

Example:
The function $f : \mathbb{R} \to \mathbb{R}$ that maps $(-\infty, 0]$ to 0 and $(0, \infty)$ to 1 is not continuous because for any $0 < \epsilon < 1$, the interval $(-\epsilon, \epsilon)$ is open, but $f^{-1}((-\epsilon, \epsilon))$ is not.
Path - Definition

- A **Path** is a continuous function $\gamma : [0, 1] \to X$.
- Connecting $\gamma(0)$ to $\gamma(1)$ in $X$.
- Self intersections are allowed (there can be $s, t \in [0, 1]$ such that $\gamma(s) = \gamma(t)$).
- No self intersections $\Rightarrow$ the function is **injective** and the path is **simple**.
Connectedness

- A topological space is **path-connected** if every pair of points is connected by a path.
- A **separation** of a topological space is a partition of $X$ into two non-empty, open, disjoint subsets (whose union is $X$).
- A topological space is **connected** if it has no separation.
- In other words: A topological space $X$ is **connected** iff the only subsets of $X$ that are both open and closed in $X$ are the empty set $\emptyset$ and $X$.
- Connectedness is weaker than path-connectedness.
Disjoint Set Systems

Given a graph, how can we decide whether it is connected or not?

We can use the disjoint-set data structure (AKA union-find), for instance.

The system:

- Stores each connected components of the graph as a subset of the vertices
- Edges are added one at a time
- The graph is connected iff in the end there is only one set left
Disjoint Set Systems - Operations

- **Find**(i) - return the name of the set that contains i
- **Union**(i, j) - if i, j belong to different sets, replace the two sets by their union.
Disjoint Set Systems - Data Structure

Data structure:
- Stores each set as a tree embedded in a linear array $arr$.
- Each tree node has a pointer to its parent, except for the root.
- $\text{Find}(i)$ - traverse the tree upward (starting at $arr[i]$) until we reach the root and return the root.
- $\text{Union}(i, j)$ - calls $\text{Find}(i)$ and $\text{Find}(j)$ and if the found roots are different assigns one root as the parent of the other.

Not very efficient, we may get very long paths!
Disjoint Set Systems - Data Structure

Improving the data structure:

- Always link the smaller to the larger tree
- A tree of $k$ nodes cannot have paths longer than log $k$ (the size of the subtree grows by at least a factor of 2 each time we join with a larger tree)
- Additional improvement: compress paths whenever we traverse them
  - In Find($i$): if $i$ is not the root, the function recursively finds the root of $i$ and makes the root the parent of $i$
Disjoint Set Systems - Analysis

- \( n \) - number of vertices

- The running time for executing a sequence of \( m \) union and find operations is \( O(m\alpha(n)) \), where \( \alpha(n) \) is the inverse of Ackermann function

- In practice - the algorithm takes constant average time per operation
Curves in the Plane

Capturing aspects of connectivity that go beyond components

- Closed curves
- Parity algorithm
- Polygon triangulation
- Winding number
Closed Curves

Closed curve - definitions:

- A path in which 0 and 1 map to the same point
- A map from the unit circle, $\gamma : S^1 \rightarrow X$, where $S^1 = \{x \in \mathbb{R}^2 | \|x\| = 1\}$
- Clearly, paths and closed curves capture different properties of topological spaces
Homeomorphism

Definition:
- Two topological spaces are **homeomorphic** (or topologically equivalent) if there exists a continuous bijection from one space to the other whose inverse is also continuous.
- Homeomorphism between two spaces gives a bijection between their open sets.

The unit interval and the unit circle are not homeomorphic.
Proof: removing the midpoint decomposes the interval into two components while removing its image leaves the circle connected - therefore, no bijection that is continuous in both directions exists.
Simple Closed Curves

A simple closed curve:
- Does not have self-intersections
- Decomposes the plane into two pieces (Inside / Outside)
Jordan Curve Theorem

Removing the image of a simple closed curve from $\mathbb{R}^2$ leaves two connected components:

- the inside (bounded)
- the outside (unbounded)

The inside together with the image of the curve is homeomorphic to a closed disk

The proof is not simple!

The theorem has generalizations to higher dimensions
Parity Algorithm

Let $C$ be a simple closed curve in the plane.

Input: a query point $x \in \mathbb{R}^2$

Output: $x$ lies inside, on, or outside the curve

- We assume a finite approximation of $C$, for instance:
  - specify $\gamma$ at a finite number of points and interpolate linearly between them
  - the result is a simple closed polygon
Parity Algorithm - Query

Query: a point \( x = (x_1, x_2) \)
Draw a half line emanating from \( x \) and count how many times it crosses the polygon
Assuming that \( x \) does not lie on the polygon
- Odd number of crossings \( \Rightarrow x \) lies inside
- Even number of crossings \( \Rightarrow x \) lies outside
We can assume that no 3 points are collinear and no two lie on a common vertical or horizontal line.

We can choose the horizontal half-line from $x$ to the right.

Now, we go over all the edges of the polygon performing the following steps:

▶ Suppose that $a = (a_1, a_2), b = (b_1, b_2)$ are the endpoints of an edge of the polygon, and $a_2 < b_2$

▶ Verify that the horizontal line crosses the edge (if $a_2 < x_2 < b_2$)

▶ If it does, check whether the crossing is to the left or to the right of $x$ - how?
Parity Algorithm - Query cont’d

We compute the determinant of the matrix:

\[ \Delta(x, a, b) = \begin{vmatrix} 1 & x_1 & x_2 \\ 1 & a_1 & a_2 \\ 1 & b_1 & b_2 \end{vmatrix} \]

It is positive iff \( x, a, b \) form a left-turn

We still need to handle the non-generic cases by slightly moving \( x \)
Polygon Triangulation

- A more structured representation of the inside of the polygon
- Decomposition into triangles
  - triangle vertices are a subset of the polygon vertices
  - triangle edges are a subset of the polygon edges and all possible diagonals
A Triangulation Always Exists

- Suppose that $n > 3$
- $b$ - the leftmost vertex of a polygon
  - We can either connect its two neighbors $a$ and $c$
  - Or connect $b$ to $d$, which is the leftmost vertex that lies inside $\triangle abc$

$n$-gon was decomposed into $n_1$-gon and $n_2$-gon
  - $n_1 + n_2 = n + 2$
  - $n_1, n_2 \geq 3 \Rightarrow n_1, n_2 < n$

- Use induction to complete the proof
Winding Number

- Let $\gamma : S^1 \to \mathbb{R}^2$ be a closed curve
- $x$ - a point not in the image of the curve
- $W(\gamma, x)$ (integer) = the total number of times that the curve travels counterclockwise around $x$
  - $W(\gamma, x) = (1 \times \text{number of counterclockwise turns}) + (-1 \times \text{number of clockwise turns})$
- If $\gamma$ is simple then the points inside the curve have the same winding number (either $+1$ or $-1$)
Winding Number of Non-simple Curves

- Suppose we move $x$ in the plane
  - $W(\gamma, x)$ does not change as long as $x$ does not cross the curve
  - If $x$ crosses the curve from left to right $W(\gamma, x)$ changes by $-1$
  - If $x$ crosses the curve from right to left $W(\gamma, x)$ changes by $+1$

- The region with locally maximum winding number lies to the left of all its boundary arcs
Bibliography

- Computational Topology, an Introduction by Edelsbrunner and Harer, Chapter I, Sections 1,2
- Topology by Munkres, Chapters 2,3
- Wikipedia
- Lecture notes for “Topology and Metric Spaces” course