Topological data analysis

Between topology and computer science

Neža Mramor Kosta

University of Ljubljana, Faculty of computer and information science
and
Institute for mathematics, physics and mechanics

Ljubljana, December 1, 2009

What is computational (algebraic) topology?

A field between topology and computer science dealing with

What is computational (algebraic) topology?

A field between topology and computer science dealing with

- applications of *topological concepts and models* in problems arising in *data analysis*, for example in reconstructions of curves, surfaces, terrains, objects, etc. from a sample of points

What is computational (algebraic) topology?

A field between topology and computer science dealing with

- applications of *topological concepts and models* in problems arising in *data analysis*, for example in reconstructions of curves, surfaces, terrains, objects, etc. from a sample of points
  describing complex models like configuration spaces, molecules,

What is computational (algebraic) topology?

A field between topology and computer science dealing with

- applications of *topological concepts and models* in problems arising in *data analysis*, for example in reconstructions of curves, surfaces, terrains, objects, etc. from a sample of points
  describing complex models like configuration spaces, molecules,
  in image and pattern alanysis

What is computational (algebraic) topology?

A field between topology and computer science dealing with

- applications of *topological concepts and models* in problems arising in *data analysis*, for example in reconstructions of curves, surfaces, terrains, objects, etc. from a sample of points
  describing complex models like configuration spaces, molecules,
  in image and pattern alanysis
- designing *algorithms* for *computing topological invariants* like homology and cohomology groups, cohomology rings, Steenrod operations, spectral sequences,

closely connected to the better established field of *computational geometry*.

Topological data analysis

- A topological model, for example a *simplicial* or more generally a *cellular complex* is associated to a “point cloud”

Topological data analysis

- A topological model, for example a *simplicial* or more generally a *cellular complex* is associated to a “point cloud”
- The idea goes far back: Dirichlet, Voronoi, Delaunay,

Topological data analysis

- A topological model, for example a *simplicial* or more generally a *cellular complex* is associated to a “point cloud”
- The idea goes far back: Dirichlet, Voronoi, Delaunay, Carlsson, Edelsbrunner . . .

Topological data analysis

- A topological model, for example a *simplicial* or more generally a *cellular complex* is associated to a “point cloud”
- The idea goes far back: Dirichlet, Voronoi, Delaunay, Carlsson, Edelsbrunner . . .
- A numbers of algorithms, implementations, libraries, exists.

Topological data analysis

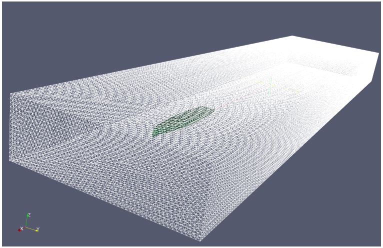
- A topological model, for example a *simplicial* or more generally a *cellular complex* is associated to a “point cloud”
- The idea goes far back: Dirichlet, Voronoi, Delaunay, Carlsson, Edelsbrunner . . .
- A numbers of algorithms, implementations, libraries, exists.
- The stress is on simplicial complexes (trianguations) or regular cell complexes

Topological data analysis

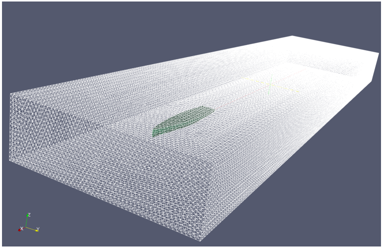
- A topological model, for example a *simplicial* or more generally a *cellular complex* is associated to a “point cloud”
- The idea goes far back: Dirichlet, Voronoi, Delaunay, Carlsson, Edelsbrunner . . .
- A numbers of algorithms, implementations, libraries, exists.
- The stress is on simplicial complexes (trianguations) or regular cell complexes
- This enables the use of a wide topological “arsenal”, for example homology, homotopy, topological simplifications.

- The topological models are often very big:

- The topological models are often very big:

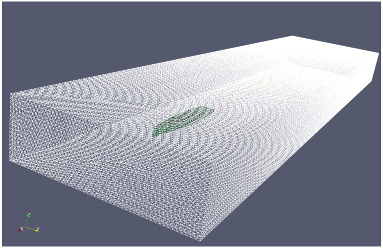


- The topological models are often very big:



- this calls for implementations of efficient (smaller) models, for example CW complexes instead of simplicial complexes with fewer cells

- The topological models are often very big:

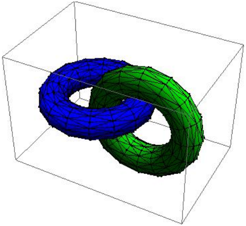


- this calls for implementations of efficient (smaller) models, for example CW complexes instead of simplicial complexes with fewer cells
- and efficient algorithms.

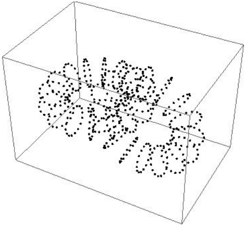
- The model can also be wrong.

- The model can also be wrong.

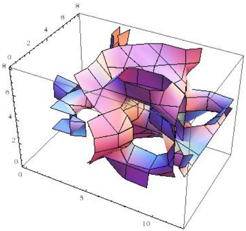
$\mathcal{M}$



sample

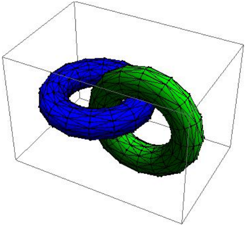


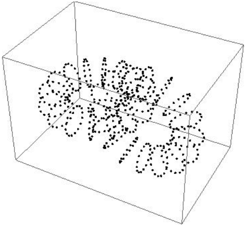
reconstruction

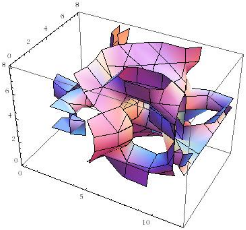


- The model can also be wrong.

$\mathcal{M}$ sample reconstruction







- The correctness can be checked by various topological invariants: number of connected components, dimension, number of “holes” (i.e. homology groups), etc.

Sampled functions

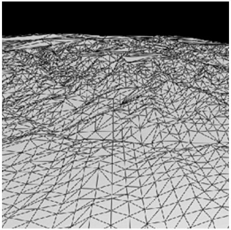
- $f : X \to \mathbb{R}$, given (measured) only in finitely many points $\{x_1, \ldots, x_n\} \subset X$

Sampled functions

- $f : X \to \mathbb{R}$, given (measured) only in finitely many points $\{x_1, \ldots, x_n\} \subset X$
- Example: reconstructing a terrain from a sample of measured altitudes

Sampled functions

- $f : X \to \mathbb{R}$, given (measured) only in finitely many points $\{x_1, \dots, x_n\} \subset X$
- Example: reconstructing a terrain from a sample of measured altitudes



Topological model: Morse theory

Marston Morse, 1925

Topological model: Morse theory

Marston Morse, 1925
Spectacular results and applications in topology and in dynamical systems (Smale, 1961).

Topological model: Morse theory

Marston Morse, 1925
Spectacular results and applications in topology and in dynamical systems (Smale, 1961).

- M a *smooth (compact) manifold*, that is, each point has a neighborhood which looks like the euclidean space, and the transition between two such neighbourhoods is a smooth function

Topological model: Morse theory

Marston Morse, 1925
Spectacular results and applications in topology and in dynamical systems (Smale, 1961).

- M a *smooth (compact) manifold*, that is, each point has a neighborhood which looks like the euclidean space, and the transition between two such neighbourhoods is a smooth function
- $f : M \to \mathbb{R}$ a *smooth function* on M, that is, in local coordinates it is differentiable

Topological model: Morse theory

Marston Morse, 1925
Spectacular results and applications in topology and in dynamical systems (Smale, 1961).

- M a *smooth (compact) manifold*, that is, each point has a neighborhood which looks like the euclidean space, and the transition between two such neighbourhoods is a smooth function
- $f : M \to \mathbb{R}$ a *smooth function* on M, that is, in local coordinates it is differentiable
- A *systematic analysis of the critical points* of f enables
  - a *topological reconstruction* of M as a CW complex with one cell of dimension i for each critical point of index i, where the index of a critical point is the number of independent directions in which the function values decrease, and

Topological model: Morse theory

Marston Morse, 1925
Spectacular results and applications in topology and in dynamical systems (Smale, 1961).

- M a *smooth (compact) manifold*, that is, each point has a neighborhood which looks like the euclidean space, and the transition between two such neighbourhoods is a smooth function
- $f : M \to \mathbb{R}$ a *smooth function* on M, that is, in local coordinates it is differentiable
- A *systematic analysis of the critical points* of f enables
  - a *topological reconstruction* of M as a CW complex with one cell of dimension i for each critical point of index i, where the index of a critical point is the number of independent directions in which the function values decrease, and
  - a decomposition of the domain M into the *Morse-Smale complex* with respect to the flow lines of the gradient vector field of f

Discrete Morse theory

Robin Forman, 1998

Discrete Morse theory

Robin Forman, 1998
One of the possible generalizations of Morse theory to non-smooth domains like simplicial complexes and general CW complexes

Discrete Morse theory

Robin Forman, 1998

One of the possible generalizations of Morse theory to non-smooth domains like simplicial complexes and general CW complexes

- M a (finite) simplicial or CW complex, that is, $M = \{c_i, i \in I\}$ is a finite set of cells, nicely glued together along the boundaries

Discrete Morse theory

Robin Forman, 1998

One of the possible generalizations of Morse theory to non-smooth domains like simplicial complexes and general CW complexes

- M a (finite) simplicial or CW complex, that is, $M = \{c_i, i \in I\}$ is a finite set of cells, nicely glued together along the boundaries
- $f: M \to \mathbb{R}$ (an indexing of the cells)

Discrete Morse theory

Robin Forman, 1998
One of the possible generalizations of Morse theory to non-smooth domains like simplicial complexes and general CW complexes

- M a (finite) simplicial or CW complex, that is, $M = \{c_i, i \in I\}$ is a finite set of cells, nicely glued together along the boundaries
- $f : M \to \mathbb{R}$ (an indexing of the cells)
- f almost restects dimension: for each cell $c \in M$, $f(a) < f(c)$ if $a \subset \partial c$ and $f(b) > f(c)$ if $c \subset \partial b$ with at most one exception.
- A critical cell of index i: a cell of dimension i, where there is no such exception.

Discrete Morse theory

Robin Forman, 1998

One of the possible generalizations of Morse theory to non-smooth domains like simplicial complexes and general CW complexes

- M a (finite) simplicial or CW complex, that is, $M = \{c_i, i \in I\}$ is a finite set of cells, nicely glued together along the boundaries
- $f : M \to \mathbb{R}$ (an indexing of the cells)
- f almost restects dimension: for each cell $c \in M$, $f(a) < f(c)$ if $a \subset \partial c$ and $f(b) > f(c)$ if $c \subset \partial b$ with at most one exception.
- A critical cell of index i: a cell of dimension i, where there is no such exception.

Discrete Morse theory

Robin Forman, 1998

One of the possible generalizations of Morse theory to non-smooth domains like simplicial complexes and general CW complexes

- M a (finite) simplicial or CW complex, that is, $M = \{c_i, i \in I\}$ is a finite set of cells, nicely glued together along the boundaries
- $f : M \to \mathbb{R}$ (an indexing of the cells)
- f almost restects dimension: for each cell $c \in M$, $f(a) < f(c)$ if $a \subset \partial c$ and $f(b) > f(c)$ if $c \subset \partial b$ with at most one exception.
- A critical cell of index i: a cell of dimension i, where there is no such exception.
- The discrete vector field on M induced by f: the set of all such exceptions, that is, the pairs (a, b) with $a \subset \partial b$ and $f(a) > f(b)$.

Advantages

- As in the smooth case we obtain a reconstruction of the domain M as a CW complex with one cell of dimension i for each critical cell of index i of f and a decomposition resembling the Morse-Smale decomposition.

Advantages

- As in the smooth case we obtain a reconstruction of the domain M as a CW complex with one cell of dimension i for each critical cell of index i of f and a decomposition resembling the Morse-Smale decomposition.
- A discrete Morse function is a very simple combinatorial object, well suited for implementations and algorithms.

Advantages

- As in the smooth case we obtain a reconstruction of the domain M as a CW complex with one cell of dimension i for each critical cell of index i of f and a decomposition resembling the Morse-Smale decomposition.
- A discrete Morse function is a very simple combinatorial object, well suited for implementations and algorithms.
- It incorporates a very simple mechanism for dealing with noise in the data.

Contributions

- King, Knudsen, M. (2005): An algoritm for extending function values measured in finitely many points of the domain *M* to a discrete Morse function on a triangulation (regular cell decomposition) of *M*

Contributions

- King, Knudsen, M. (2005): An algoritm for extending function values measured in finitely many points of the domain M to a discrete Morse function on a triangulation (regular cell decomposition) of M(more precisely, to a discrete vector field)

Contributions

- King, Knudsen, M. (2005): An algoritm for extending function values measured in finitely many points of the domain M to a discrete Morse function on a triangulation (regular cell decomposition) of M(more precisely, to a discrete vector field)

Contributions

- King, Knudsen, M. (2005): An algoritm for extending function values measured in finitely many points of the domain *M* to a discrete Morse function on a triangulation (regular cell decomposition) of *M*(more precisely, to a discrete vector field)(*M* je regularen celični kompleks)

Contributions

- King, Knudsen, M. (2005): An algoritm for extending function values measured in finitely many points of the domain M to a discrete Morse function on a triangulation (regular cell decomposition) of M(more precisely, to a discrete vector field)(M je regularen celični kompleks)
- Ayala, Jerše, M., Vilches (2009): a characterization of discrete vector fields on noncompact domains, which arise from a proper discrete Morse function

Contributions

- King, Knudsen, M. (2005): An algoritm for extending function values measured in finitely many points of the domain M to a discrete Morse function on a triangulation (regular cell decomposition) of M(more precisely, to a discrete vector field)(M je regularen celični kompleks)
- Ayala, Jerše, M., Vilches (2009): a characterization of discrete vector fields on noncompact domains, which arise from a proper discrete Morse function(on compact domains a characterization was given already by Forman)

Contributions

- King, Knudsen, M. (2005): An algoritm for extending function values measured in finitely many points of the domain *M* to a discrete Morse function on a triangulation (regular cell decomposition) of *M*(more precisely, to a discrete vector field)(*M* je regularen celični kompleks)
- Ayala, Jerše, M., Vilches (2009): a characterization of discrete vector fields on noncompact domains, which arise from a proper discrete Morse function(on compact domains a characterization was given already by Forman)
- Jerše, M. (2008): The formalization of the Morse-Smale complex in the context of discrete Morse functions, and an algorithm for its reconstruction

Contributions

- King, Knudsen, M. (2005): An algoritm for extending function values measured in finitely many points of the domain M to a discrete Morse function on a triangulation (regular cell decomposition) of M(more precisely, to a discrete vector field)(M je regularen celični kompleks)
- Ayala, Jerše, M., Vilches (2009): a characterization of discrete vector fields on noncompact domains, which arise from a proper discrete Morse function(on compact domains a characterization was given already by Forman)
- Jerše, M. (2008): The formalization of the Morse-Smale complex in the context of discrete Morse functions, and an algorithm for its reconstruction
- King, Knudsen, M. (2006): The definition of a parametric discrete Morse function and an algorithm for the bifurcation diagram of the critical cells

Applications in image analysis

- Žabkar, Jerše, Bratko, Planker, Schlemmer (2008): an algorithm for teaching a robot the concept of occlusion

Applications in image analysis

- Žabkar, Jerše, Bratko, Planker, Schlemmer (2008): an algorithm for teaching a robot the concept of occlusion
- Jerše, M. (2008): an example of tracing features in a sequnece of images

Applications in image analysis

- Žabkar, Jerše, Bratko, Planker, Schlemmer (2008): an algorithm for teaching a robot the concept of occlusion
- Jerše, M. (2008): an example of tracing features in a sequnece of images
- An obvious problem left to solve: simultaneous cancelling of pairs of critical cells.