

MR2572029 (2011e:00001) 00-02 05C10 52-02 55-02 57-02 65D18 68U05
Edelsbrunner, Herbert; Harer, John L.

★**Computational topology.**

An introduction.

American Mathematical Society, Providence, RI, 2010. xii+241 pp.
ISBN 978-0-8218-4925-5

This much needed book performs two basic tasks. Its primary aim is to form a basis for an introductory course in computational topology—a subject that is only about a decade old but which is already very rich in both content and applications. The second purpose is to provide a fairly complete account of topological persistence—a concept which has already found a multitude of applications in areas ranging from pure mathematics to biology, and to which the authors have made major foundational contributions. Much of what appears in the later chapters of the book is novel or the result of recent research. This means that the authors have set themselves the task of starting virtually from nothing and reaching the “state of the art” in a multidisciplinary subject in the space of just over 220 pages: a very tall order indeed, and one that can only be achieved by relying a great deal on the reader’s intuition as well as his willingness to devote a substantial amount of time to filling in missing details and looking up the literature referred to in the bibliographic notes. Fortunately the bibliographic notes are very extensive and informative. Each chapter ends with exercises for the reader.

“Computational topology”, in the sense that the authors use the term, is the study of algorithms that solve problems of a topological character, usually of the kind that have or are expected to have applications outside the area of topology. The “study of algorithms” is meant in the sense that has been established in computer science: our interest is not just in the existence of algorithms but also in their efficiency, and ultimately in creating concrete computer implementations that can be used to solve problems in science and engineering. Although the emphasis of such a study is on topology, the somewhat older subject of “computational geometry” often plays a key role in many of the algorithms, so the book also provides a substantial introduction into this field.

The authors state in the Preface that they assume no prior knowledge either of topology or of computer science. This is literally true in the former case, but not quite so in the latter. For example, many of the algorithms in the book are described in terms of pseudo-code which is never explicitly explained, though which would be clear enough to anyone with any kind of programming experience. Familiarity with concepts such as “data structures” and the basic principles of analysis of algorithms (time complexity) is also assumed.

On the other hand, the topology in the book really does begin with the definition of a topological space and proceeds on to the notions of simplicial complexes, manifolds (smooth and PL), homotopy types, homology and cohomology groups, Poincaré duality, Morse theory and many others. These topics, together with better known (at least to computer scientists) topics from graph theory and computational geometry (e.g. Voronoi diagrams, Delaunay triangulation, etc.), are covered in the first six chapters, which form the “introductory” part of the book.

Persistent homology is the key notion of the later sections of the book. The chapter on persistence follows a brief but quite extensive discussion of Morse functions and their

PL analogues. The basic example of persistence is provided by looking at the way the homology of the sublevel sets of a Morse function on a manifold changes when one passes through a critical point. At such points homology classes can either “be born” or “die”. “Persistence” attempts to measure the “duration” of a topological feature, such as a homology class, for example, using the difference between the values of the Morse function at the critical points where the class is born and where it dies. In applications the basic assumption is that more “persistent” topological features are important while the less persistent ones can be regarded as “noise”. The concept of persistence is readily extended to filtered simplicial complexes and simplicial homology. Information about persistent homology and persistent Betti numbers can be conveniently encoded in two-dimensional diagrams called persistence diagrams.

Besides having a convenient description in terms of diagrams, persistence can be computed efficiently. The book describes two algorithms that do so: one based on matrix reduction to the Smith Normal Form and one based on a sparse matrix representation.

The concept of extended persistence, which has important applications to fitting shapes to each other, is also considered in one of the more advanced sections of the book. It relies on a substantial amount of algebraic topology, including both Poincaré and Lefschetz duality.

An important mathematical property of persistence is stability under perturbations, in other words, topological objects which are in some sense close should have persistence diagrams that are also “close”. To formalize this one needs to define the distance between persistence diagrams. Two such notions, bottleneck distance and Wasserstein distance, are defined in chapter 8 and are used to prove stability theorems. These are then used to generalize classic results about curves. Algorithms for computing both kinds of distance are also described.

The last chapter (9th) is, at least for this reviewer, the most fascinating and the most difficult to review. It is concerned with applications of the tools developed in the earlier chapters to data analysis, and via that to numerous disciplines in science and engineering. Such data always contains some form of “noise”, which may reflect genuine properties of the measured phenomena but on a scale that lies outside our area of interest. The traditional approach involves some forms of “smoothing” the data, which results in their being changed—the new, “topological” approach aims at identifying and measuring the noise without changing the data.

Four applications are considered: gene expression, protein docking, image segmentation and the structure of plant roots. Not knowing anything about these fields, the reviewer cannot judge how significant these applications would appear to the experts, but at least they seem both serious and natural. Thus the book is another witness to the remarkable phenomenon that not so long ago seemed astonishing: more and more areas of mathematics, once considered “pure”, are finding serious applications in the “real world” thanks to the interaction between mathematics and computers.

Andrzej Kozłowski