J Stat Phys (2012) 149:1161–1163
DOI 10.1007/s10955-012-0655-y

Lorenzo Saitta, Attilio Giordana, Antoine Cornuéjols: Phase Transitions in Machine Learning

Cambridge University Press, 2011

Remi Monasson

Received: 15 November 2012 / Accepted: 22 November 2012 / Published online: 1 December 2012

Phase transitions, that is, abrupt changes in the macroscopic behavior of systems resulting from slight modifications of their control parameters, abound in Nature. Understanding what happens at or close to a critical point has therefore always been one of the key goals of statistical physics. In the past few decades it has become increasingly clear that the concepts and tools developed to tackle physical phase transitions, e.g., those taking place in condensed-matter systems, could also be relevant to problems at the boundaries of physics. One such example is the percolation transition, originally motivated by the motion of fluids through porous media, which gave birth over the last fifty years to a rich body of mathematical literature. Various aspects of percolation transitions, ranging from the appearance of a macroscopic connected components to the onset of more detailed structural properties, e.g., the presence of a large K-core, have been extensively studied in the context of random graph theory.

The experimental discovery of 'new' percolation transitions in random constraint satisfaction problems defined on random hypergraphs by computer scientists at the beginning of the 90's was one more illustration of the ubiquity of phase transitions, or zero-one laws in the language of probability theory, in random combinatorial structures. These constraint satisfaction problems also offered an ideal testing ground for the methods developed by statistical physicists working on mean-field disordered systems (spin-glasses) in the 80's. These approaches provided an understanding of the qualitative and quantitative features of phase transitions in constraint satisfaction problems.

That statistical physics could be useful to tackle mathematical or computer science-issued problems was not a novelty in itself at all. It is indeed appropriate to recall how the road to the applications of statistical physics ideas to out-of-physics problems was paved by two extremely influential works at the beginning of the 80's. On the numerical side, the simulated annealing algorithm, introduced by Kirkpatrick, Gelatt and Vecchi in 1983, first offered a versatile and powerful technique to deal with optimization problems based on

R. Monasson (✉)
Laboratoire de Physique Théorique de l'ENS, 24, rue Lhomond, 75231 Paris Cedex 05, France
e-mail: monasson@lpt.ens.fr

statistical physics ideas. A year earlier, in 1982, Hopfield proposed an Ising-like model for auto-associative memories, which had considerable conceptual impact in theoretical neuroscience and beyond. The introduction of the Hopfield model, and its statistical mechanics resolution by Amit, Gutfreund and Somplinsky two years later, unleashed a wealth of theoretical statistical physics studies about problems encountered in computer science, be they related to combinatorial optimization, information theory, machine learning, or statistical inference.

The book by Saitta, Giordana and Cornuéjols deals with one of those subjects, namely, machine learning. Remarkably, although all three authors are computer scientists (in Italy and France), they show a deep interest in the statistical physics approach to phase transitions in computer science. The core part of the book is devoted to the introduction of essential notions of machine learning and to the presentation of numerical investigations of phase transition-like phenomena in various applications, ranging from constraint satisfaction problems to the inference of regular grammar generating automata or to covering tests in propositional learning. Other chapters present basic notions of statistical physics and applications to what the authors call ‘complex’ systems, motivated by problems coming from social and biological sciences.

What is the right audience for this book? Generally speaking, I think this book will be useful to researchers interested in knowing more about the ongoing empirical developments about phase transitions in machine learning. Empirical means here that the strong point of the book is to report results from numerical investigations and insights about threshold-like phenomena, rather than the methods that can provide for their theoretical understanding. Statistical physicists who are not acquainted with machine learning will definitely extract valuable information from this book, especially from the chapters devoted to relational learning and grammatical inference. As for computer scientists wishing to deepen their knowledge of phase transition phenomena in their domain, they will find many entry points to the literature and a variety of examples that cover many aspects of machine learning and beyond, such as constraint satisfaction problems.

I am however more circumspect about the usefulness of the chapter that is devoted to statistical mechanics. Readers with a physics background will not learn much, and will likely be annoyed by frequent mistakes in the formulas (no, the entropy is not equal to minus the log of the number of micro-states) and by inaccurate statements (for instance, Parisi’s feat was not to propose replica symmetry, but to found replica symmetry-breaking theory). As for researchers working in the field of machine learning, I doubt that they will be able to understand disordered systems from a two-page presentation of the replica method, or that they will benefit much from a long (3-page) exposition of an approximate derivation of the value of the critical temperature of the 2D-Ising model with ferromagnetic nearest neighbor interactions. It is very hard to agree with the authors’ conclusion of the chapter, ‘From the practical point of view, for a computer scientist wanting to use the approaches mentioned in this chapter it is not necessary to become deeply acquainted with statistical physics; some basic notions, complemented by the mathematical tools sketched in this chapter, will be sufficient to apply these approaches fruitfully’. Along the same line, the chapter devoted to complex systems is likely to be too short and too superficial to be really useful to the reader.

This said, the core chapters are really interesting. I particularly enjoyed reading the chapter about grammatical inference. Regular grammars, in Chomsky’s classification, are languages that can be accepted (generated) by a finite automaton. Inferring this automaton from a set of strings is an interesting problem. The authors report evidence for the presence of sharp changes in the structure of the inferred automaton during the learning process. Starting from a prefix-tree automaton with many states, accepting the training set of strings and merging states together to accept more general languages, an abrupt transition from a poor generalization-large automaton regime to a good generalization-small automaton phase is observed. Understanding this behavior and quantifying it is, apparently, an open challenge for statistical physicists. Besides this specific example I cannot agree more with the authors's claim that investigating the properties of the learning process itself, and not only of its outcome, is very interesting and important. Learning defines a dynamical evolution, which is somewhat unusual from a physics point of view (absence of locality or non-Markovian features), and which deserves to be understood on its own. Such a statement is by no means new. Physicists made important contributions in the context of neural networks and on-line learning for instance about twenty years ago. However much more deserves to be done and the present book will provide the reader with challenging observations and should stimulate statistical physicists' curiosity and imagination.

In conclusion I recommend this book to researchers in statistical physics who are tempted to apply their skills and intuitions to machine learning problems. Though the book by itself cannot be considered as a self-contained introduction to machine learning (I would recommend for instance Data Mining: Practical Machine Learning Tools and Techniques by Witten, Frank and Hall for a more complete and non-theoretical introduction to the subject) it is definitely a source of stimulating information. Despite the weaknesses of the chapter(s) devoted to statistical mechanics I find it outstanding that computer scientists are fully aware that statistical mechanics can play a role in elucidating the non-trivial phenomena (be they dynamical or probabilistic) encountered in their field. From this point of view, the efforts of the authors will be fully rewarded if this book kindles new studies and interests among statistical physicists and enhances the interdisciplinary exchanges with the computer science community.