Preface

> ... like prisoners chained in a cave with their faces to the wall ...
>
> ... to them the truth would be literally nothing but the shadows on the cave wall ...
>
> Paraphrasing Plato's Parable of the Cave in *Republic*

The objective of this book is to supply examples—with rigorous mathematical proofs—of the following vague complexity law:

(i) discrete systems are either "simple" or they exhibit "advanced pseudorandomness" with or without constraints (even when there is no apparent independence);
(ii) and roughly speaking, *a priori* probabilities often exist, even when there is no intrinsic symmetry.

Part of the difficulty is how to clarify these vague statements. For example, here "advanced" (pseudorandomness) means roughly around the Central Limit Theorem, and the *a priori* probabilities are implicitly justified by laws of large numbers. An indirect evidence for the underlying "hidden randomness" is a mysterious phenomenon that I call *Threshold Clustering*—I give several illustrations.

This book grew, rather unexpectedly, out of a graph theory graduate course at Rutgers University in the spring of 2007. I am very grateful to my students for putting constant pressure on me to do something new in every class.

This is a spin-off of the much longer book *Combinatorial Games: Tic-Tac-Toe Theory* that was published by Cambridge University Press in 2008. Admittedly there are some overlaps, but this one has its own viewpoint, most of the results are new, and it is self-contained. This is an independent book with a much broader perspective (see Part A), but it can also be considered Volume 0 of the *Tic-Tac-Toe Theory* (despite the fact that this was written later). It is a good idea to read this book first and to move to *Tic-Tac-Toe Theory* after, if the reader is still interested.

A primary subject of this book is to understand randomness and complexity. The traditional approach to complexity—computational complexity theory—is to study very general complexity classes, such as P, NP, and PSPACE. What I do here is very different: I feel that studying interesting concrete systems, narrow subclasses, such as natural game classes, can give a new angle, new insights into the mystery of complexity.

In the last part of the book I focus on Games and Graphs, in particular on the new game-theoretic concept of Surplus. In most sports, winning means outscoring the opponent; this motivates the definition of Surplus. The Surplus can be defined for every hypergraph, but here, for the sake of simplicity, I focus just on graphs. The Surplus is a graph parameter like the chromatic number or the independence number, but the Surplus is even more difficult to determine. We don't know the exact value of the Surplus even for the simplest dense graphs, including the complete graph.

Games have the following natural classification: (1) Games of Chance, (2) Games of Incomplete Information (like Poker and the "coin-hiding" game), and (3) Games of Complete Information (like Chess and Tic-Tac-Toe).

A study of (1) led to the developments of classical probability theory (Pascal, Fermat, de Moivre, Laplace—to name only a few).

A study of (2) led to the developments of traditional game theory (von Neumann, Nash, etc.).

By von Neumann's minimax theorem, an optimal strategy in a Game of Incomplete Information (usually) demands Random Play (to compensate for the lack of information). For example, in the "coin-hiding" game the optimal strategy is to hide the coin in the left or right fist randomly.

From studying (3) comes the big surprise. For Games of Complete Information (with no chance moves) Random Play seems to be a useless concept (since we don't have to compensate for the lack of information); nevertheless, for large infinite classes of games (e.g., grown-up versions of Tic-Tac-Toe) we can achieve an optimal strategy by a derandomization of the Random Play; in fact, Random Play with or without Constraints (Endgame Policy).

I'd better admit up-front that the theory of computational complexity is simply beyond the scope of this book. I will hardly discuss, or even mention, results about complexity classes. Besides my ignorance of the subject, I have the following reasons/excuses: (1) I wanted the book to be as short as possible; (2) the precise definitions of the complexity classes, and the intuition behind the definitions, is rather space-consuming; (3) I don't want to go into the delicate issue of interpretation, about the "meaning" of the conditional results; and (4) I have nothing new to say about the traditional complexity classes.

Also, I have to point out that, even if "chaos" and "randomness" are almost synonyms, this book focuses entirely on discrete mathematics and has nothing to do with chaos theory (i.e., sensitive dependence on initial conditions in continuous mathematics).

Part A of the book is mostly an essay; Part B is partly new results with proofs and partly a survey (including a summary of the main results from my previous book *Tic-Tac-Toe Theory*). Part C is mostly new results with proofs. Part A is an easy read; Part B is harder (because I expect the reader to fully understand the proofs); and Part C is much harder (because of the difficult proofs in Chapters 17–21).

To make it available to a wider audience, the book is more or less self-contained.

Next I say a few words about the notation, which is rather standard. The sets of integers, rational numbers, and real numbers are denoted, respectively, by $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$. If S is a finite set, then $|S|$ denotes the number of elements of S. I use $\|x\|$ to denote the distance of a real number x from the nearest integer (so $0 \le \|x\| \le 1/2$), and I also use the even more standard notation $\{x\}$, $\lfloor x \rfloor$, and $\lceil x \rceil$, which mean, in this order, the fractional part of x and the lower and upper integral parts of x. The natural (base e) logarithm is denoted by $\log x$, so $\int \frac{1}{x}\,dx = \log x$; in this book I don't use $\ln x$ at all. Also, $\log_2$ and $\log_{10}$ stand, respectively, for the base 2 and 10 logarithms.

As usual, c, c_0, c_1, c_2, $\ldots$ or const denote absolute constants (usually positive) that I could, but do not care to, determine. Also, I adopt the standard notation involving O and o: for functions f and nonnegative functions g, $f = O(g)$ means $|f| \leq c \cdot g$ and $f = o(g)$ means $f/g \to 0$ in the limit.

I am sure there are many (hopefully minor) errors in the book. I welcome any corrections, suggestions, and comments.

Last but not least, I would like to thank the National Science Foundation and the Harold H. Martin Chair at Rutgers University for the research grants supporting me during this work. I am especially grateful for the continuous generosity of Mr. Harold H. Martin over the last fifteen years.

József Beck (jbeck@math.rutgers.edu)
New Brunswick, New Jersey
March 2009