Preface

This monograph is the outcome of our work on probabilistic combinatorial optimization since 1994. The first time we heard about it, it seemed to us to be a quite strange scientific area, mainly because randomness in graphs is traditionally expressed by considering probabilities on the edges rather than on the vertices. This strangeness was our first motivation to deal with probabilistic combinatorial optimization. As our study progressed, we have discovered nice mathematical problems, connections with other domains of combinatorial optimization and of theoretical computer science, as well as powerful ways to model real-world situations in terms of graphs, by representing reality much more faithfully than if we do not use probabilities on the basic data describing them, i.e., the vertices.

What is probabilistic combinatorial optimization? Basically, it is a way to deal with aspects of robustness in combinatorial optimization. The basic problematic is the following. We are given a graph (let us denote it by $G(V, E)$, where V is the set of its points, called vertices, and E is a set of straight lines, called edges, linking some pairs of vertices in V), on which we have to solve some optimization problem Π. But, for some reasons depending on the reality modelled by G, Π is only going to be solved for some subgraph G' of G (determined by the vertices that will finally be present) rather than for the whole of G. The measure of how likely it is that a vertex $v_i \in V$ will belong to G' (i.e., will be present for the final optimization) is expressed by a probability p_i associated with v_i. How we can proceed in order to solve Π under this kind of uncertainty?

A first very natural idea that comes to mind is that one waits until G' is specified (i.e., it is present and ready for optimization) and, at this time, one solves Π in G'. This is what is called *re-optimization*. But what if there remains very little time for such a computation? We arrive here at the basic problematic of the book. If there is no time for re-optimization, another way to proceed is the following. One solves Π in the whole of G in order to get a feasible solution (denoted by S), called *a priori solution*, which will serve her/him as a kind of benchmark for the solution on the effectively present subgraph G'. One has also to be provided with an algorithm that modifies S in order to fit G'. This algorithm is called *modification strategy* (let us denote it by `M`). The objective now becomes to compute an *a priori* solution that, when modified by `M`, remains "good" for any subgraph of G (if this subgraph is the one where Π will be finally solved). This amounts to computing a solution that optimizes a kind of expectation of the size of the modification of S over all the possible subgraphs of G, i.e., the sum of the products of the probability that G' is the finally present graph multiplied by the value of the modification of S in order to fit G' over any subgraph G' of G. This expectation, depending on both the instance of the deterministic problem Π, the vertex-probabilities, and the modification strategy adopted, will be called the *functional*. Obviously, the presence-probability of G' is the probability that all of its vertices are present.

Seen in this way, the probabilistic version PΠ of a (deterministic) combinatorial optimization problem Π becomes another equally deterministic problem Π', the solutions of which have the same feasibility constraints as those of Π but with a different objective function where vertex-probabilities intervene. In this sense, probabilistic combinatorial optimization is very close to what in the last couple of years has been called "one stage optimisation under independent decision models", an area very popular in the stochastic optimization community.

What are the main mathematical problems dealing with probabilistic consideration of a problem Π in the sense discussed above? We can identify at least five interesting mathematical and computational problems dealing with probabilistic combinatorial optimization:

1) write the functional down in an analytical closed form;

2) if such an expression of the functional is possible, prove that its value is polynomially computable (this amounts to proving that the modified problem Π' belongs to **NP**);

3) determine the complexity of the computation of the optimal *a priori* solution, i.e., of the solution optimizing the functional (in other words, determine the computational complexity of Π');

4) if Π' is **NP**-hard, study polynomial approximation issues;

5) always, under the hypothesis of the **NP**-hardness of Π', determine its complexity in the special cases where Π is polynomial, and in the case of **NP**-hardness, study approximation issues.

Let us note that, although curious, point 2 in the above list in neither trivial nor senseless. Simply consider that the summation for the functional includes, in a graph of order n, 2^n terms (one for each subgraph of G). So, polynomiality of the computation of the functional is, in general, not immediate.

Dealing with the contents of the book, in Chapter 1 probabilistic combinatorial optimization is formally introduced and some old relative results are quickly presented.

The rest of the book is subdivided into two parts. The first one (Part I) is more computational, while the second (Part II) is rather "structural". In Part I, after formally introducing probabilistic combinatorial optimization and presenting some older results (Chapter 1), we deal with probabilistic versions of four paradigmatic combinatorial problems, namely, PROBABILISTIC MAX INDEPENDENT SET, PROBABILISTIC MIN VERTEX COVER, PROBABILISTIC LONGEST PATH and PROBABILISTIC MIN COLORING (Chapters 2, 3, 4 and 5, respectively). For any of them, we try, more or less, to solve the five types of problems just mentioned.

As the reader will see in what follows, even if, mainly in Chapters 2 and 3, several modification strategies are used and analyzed, the strategy that comes back for all the problems covered is the one consisting of moving absent vertices out of the *a priori* solution (it is denoted by `MS` for the rest of the book). Such a strategy is very quick, simple and intuitive but it does not always produce feasible solutions for any of the possible subgraphs (i.e., it is not always feasible). For instance, if it is feasible for PROBABILISTIC MAX INDEPENDENT SET, PROBABILISTIC MIN VERTEX COVER and PROBABILISTIC MIN COLORING, this is not the case for PROBABILISTIC LONGEST PATH, unless particular structure is assumed for the input graph. So, in Part II, we restrict ourselves to this particular strategy and assume that either `MS` is feasible, or, in case of unfeasibility, very little additional work is required in order to achieve feasible solutions. Then, for large classes of problems (e.g., problems where feasible solutions are subsets of the initial vertex-set or edge-set satisfying particular properties, such as stability, etc.), we investigate relations between these problems and their probabilistic counterparts (under `MS`). Such relations very frequently derive answers to the above mentioned five types of problems. Chapter 7 goes along the same lines as Chapter 6. We present a small compendium of probabilistic graph-problems (under `MS`). More precisely we revisit the most well-known and well-studied graph-problems and we investigate if strategy `MS` is feasible for any of them. For the problems for which this statement holds, we express the functional associated with it and, when possible, we characterize the optimal *a priori* solution and the complexity of its computation.

The book should be considered to be a monograph as in general it presents the work of its authors on probabilistic combinatorial optimization graph-problems. Nevertheless, we think that when the interested readers finish reading, they will be perfectly aware of the principles and the main issues of the whole subject area. Moreover, the book aims at being a self-contained work, requiring only some mathematical maturity and some knowledge about complexity and approximation theoretic notions. For help, some appendices have been added, dealing, on the one hand, with some mathematical preliminaries: on sets, relations and functions, on basic concepts from graph-theory and on some elements from discrete probabilities and, on the other hand, with elements of the complexity and the polynomial approximation theory: notorious complexity classes, reductions and **NP**-completeness and basics about the polynomial approximation of **NP**-hard problems. We hope that with all that, the reader will be able to read the book without much preliminary effort. Let us finally note that, for simplifying reading of the book, technical proofs are placed at the end of each chapter.

As we have mentioned in the beginning of this preface, we have worked in this domain since 1994. During all these years many colleagues have read, commented, improved and contributed to the topics of the book. In particular, we wish to thank Bruno Escoffier, Federico Della Croce and Christophe Picouleau for having working with, and encouraged us to write this book. The second author warmly thanks Elias Koutsoupias and Vassilis Zissimopoulos for frequent invitations to the University of Athens, allowing full-time work on the book, and for very fruitful discussions. Many thanks to Stratos Paschos for valuable help on LaTeX.

Tender and grateful thanks to our families for generous and plentiful support and encouragement during the task.

Finally, it is always a pleasure to work with Chantal and Sami Menasce, Jon Lloyd and their colleagues at ISTE.

Cécile Murat and Vangelis Th. Paschos
Athens and Paris, October 2005