of mathematics in a very skillful and accessible form. It will certainly be useful for a wide range of specialists.

Mathematical Institute
of the Czechoslovak Academy of Sciences,
Prague

BOHDAN MASLOWSKI

References

1. Arkin, V. I. and Saksonov, M. T.: Necessary optimality conditions for stochastic differential equations, *Soviet Math. Dokl.* **20** (1979), 1–5.
2. Bensoussan, A.: *Lectures on stochastic control*, Lecture Notes in Mathematics **972** (1982), 1–62.
3. Haussmann, U. G.: General necessary conditions for optimal control of stochastic systems, *Math. Programming Studies* **6** (1976), 30–48.
4. Kushner, H. J.: Necessary conditions for continuous parameter stochastic optimization problems, *SIAM J. Control* **10** (1972), 550–565.

P. J. M. van Laarhoven and E. H. L. Aarts: *Simulated Annealing: Theory and Applications*, D. Reidel, Dordrecht, 1987, 198 pp., ISBN 90-277-2513-6, Dfl. 120.

The book deals in depth with an optimization algorithm called 'simulated annealing'. Applications to various fields are discussed.

The review is written from a mathematician's point of view. To begin with, I would like to elaborate a little bit about the term 'simulated annealing'.

In various disciplines, one often encounters the problem of finding global minima of a function $U(x)$ defined on a state space (configuration space) S. For discrete variables, when the cardinality of S is large, it is not feasible to search through the whole space for the minima. The Travelling Salesman Problem is such an example. As for the continuous case, one usually finds the local minima instead.

Kirkpatrick *et al.* [7] and Cerny [1] introduced into this minimization problem the idea of 'annealing' from statistical physics. They considered U as the energy function of a physical system with state space S. For each temperature T, the thermal equilibrium of the system is described by the Gibbs distribution π_T with energy U and temperature T. One may simulate the time evolution of the system to the equilibrium π_T by generating a sequence of states X_t which form a homogeneous (ergodic) Markov chain. This is the famous Metropolis method. Note that π_T settles down to a probability distribution π_0 concentrating on the global minima of U as the temperature cools down to zero.

The idea of simulated annealing is to mimic the Metropolis method but, in the meantime, decrease the temperature slowly. Randomness induced by the thermal perturbations enables the process to escape from being trapped in the local

minima. Time is used to exchange for space and the annealing procedure will naturally lead to global minima.

The authors mainly discussed the discrete case, except Section 8.2. Different types of combinatorial optimization algorithms are discussed in the introduction. Simulated annealing can be viewed as a general approximation algorithm.

In Chapter 2, the mathematical model of simulated annealing is described. Given a neighborhood system, simulated annealing is a Markov chain that transforms the current configuration into one of its neighbors. The corresponding transition probability can be defined as

$$P_{ij}(T) = \begin{cases} G_{ij}(T)A_{ij}(T), & \text{for } j \neq i, \\ 1 - \sum_{l \neq i} P_{il}(T), & \text{for } j = i. \end{cases} \tag{1}$$

The temperature T might depend on time. Note that the transition probability is a product of two conditional probabilities. $G_{ij}(T)$ is the probability of generating j from i, usually $G_{ij}(T) = 0$, if j is not in the neighborhood of i. The number of neighbors for each configuration is far less than that of the configuration space. $A_{ij}(T)$ is the probability of accepting j, once it has been generated from i. Usually, $A_{ij}(T)$ is defined by

$$A_{ij}(T) = \exp - \frac{(U(j) - U(i))^+}{T}, \tag{2}$$

i.e. if $U(j) \leq U(i)$, then we accept j with probability 1; when $U(j) > U(i)$, there is still a positive probability of accepting j, but this probability decreases as T cools down.

The authors considered two algorithms. The homogeneous algorithm is described by a sequence of homogeneous Markov chains. Each Markov chain is generated by (1) with a fixed T and T is decreased in between subsequent Markov chains. The inhomogeneous algorithm is described by one Markov chain with transition probability (1), with $T \equiv T(k)$ depending on time k.

Let M denote the set of all globally minimal configurations. Under suitable conditions, the simulated annealing algorithm converges to M. Chapter 3 is devoted to the study of the asymptotic convergences.

The homogeneous case is quite simple. Basically if one has the existence of the stationary distribution for each T, then the asymptotic property is just the convergence of these stationary distributions.

As for the inhomogeneous case, the necessary and sufficient condition for X_t to converge in probability to M is described in Theorem 6. If $T(k)$ is of the form $c/\log k$, then there exists a constant c_0 such that X_t converges to M in probability if and only if $c \geq c_0$. The authors do not clearly explain the 'convergence' here. Actually, there are two crucial constants involved. I will come back to this point later.

The relation with statistical physics is discussed in Chapter 4. The algorithm is analyzed in Chapter 5 (implementation), and Chapter 6 (performance). These chapters are very well written. People who are not familiar with these topics should be able to understand these chapters without much difficulty.

I like Chapter 7. The authors did a good job by describing in detail many examples, especially those from computer-aided circuit design. Note that for each application one has to define the configuration space, choose a proper cost function and create a neighborhood system (a generating mechanism for transitions). Usually, this is not easy.

Parallel implementations and continuous optimization are discussed briefly in Chapter 8.

I quite agree with the authors' concluding remark

> As a final conclusion we state that the simulated annealing algorithm is a generally applicable, flexible, robust and easy-to-implement approximation algorithm, that is able to obtain near-optimal solutions for a wide range of optimization problems. However, computation time can be long and in a number of cases valuable tailored algorithms are available that can be executed in far less computation time. For those problem areas where no tailored algorithms are available, we consider simulated annealing to be a powerful optimization tool.

I would like to comment a little bit on the mathematical aspect. The convergence of the algorithm should be in probability and could not be with probability 1, as mentioned by the authors.

In the initial formulation, i.e. $A_{ij}(T)$ is defined by (2), there are two crucial constants c_0 and c_1 with $c_0 \leq c_1$. c_0 defined by Hajek is the one for convergence in probability. The constant c_1 appeared in Freidlin and Wentzell's work [4], is connected with the weak convergence of X_t to π_0. Hence, "Solutions, obtained by simulated annealing, do not depend on the initial configuration . . . " stated in Chapter 1, page 5, is ambiguous.

One may use random perturbations of dynamical systems to model the simulated annealing for the continuous configuration space. This gives the connection to Friedlin–Wentzell's work. The authors did not discuss this approach in the book.

This is the first book on this subject. The authors did an excellent job in putting together the theoretical results, working examples and an extensive bibliography. People from different disciplines including computer scientists, mathematicians, physicists, electric engineers will definitely find the book worth reading.

Academia Sinica CHII-RUEY HWANG

References

1. Cerny, V.: 'A Thermodynamical Approach to the Travelling Salesman Problem: an Efficient Simulation Algorithm', Inst. Phys. & Biophys., Comenius Univ., Bratislava, 1982.

2. Chiang, T.-S. and Chow, Y.: 'A Limit Theorem for a Class of Inhomogeneous Markov Processes', Tech. Report, Inst. of Math., Academia Sinica, Taipei, 1987.
3. Chiang, T.-S., Hwang, C.-R., and Sheu, S.-J.: 'Diffusion for Global Optimization in $\mathbf{R}^n$', *SIAM J. Control Optim.* **25** (1987), 737–753.
4. Freidlin, M.I. and Wentzell, A.D.: *Random Perturbations of Dynamical Systems*, Springer-Verlag, Berlin, New York, 1984.
5. Geman, S. and Hwang, C.-R.: 'Diffusions for Global Optimization', *SIAM J. Control Optim.* **24** (1986), 1031–1043.
6. Hwang, C.-R. and Sheu, S.-J.: 'Large Time Behaviors of Perturbed Diffusion Markov Processes with Applications III: Simulated Annealing', Tech. Report, Inst. of Math. Academia Sinica, Taipei, 1986.
7. Kirkpatrick, S., Gelatt, C. D. Jr., and Vecchi, M. P.: 'Optimization by Simulated Annealing', *Science* **220** (1983), 671–680.