Global Optimization by Simulated Tempering *

Matthew Shore

Abstract

A study of *Simulated Tempering* (ST), a Monte Carlo optimization procedure, is presented. A non–trivial, yet analytically solvable model, the two–dimensional Ising Model of Ferromagnetism (IMF), is used as a test case. A comparison of the Internal Energies obtained numerically and analytically suggests that ST is, indeed, a viable scheme. In addition, the ability of ST to tunnel, that is, to freely sample ordered states ($T < T_c$) and disordered states ($T > T_c$), is illustrated.

1 Introduction

Global optimization is relevant to many areas of science and has evolved, with the advent of computers, into a prominent area of research [5, 7, 10]. This discipline is concerned with systems (functions) that have sofar not been optimized analytically. It includes those problems which, although in theory are solvable directly, impose unrealistic computational demands. Two examples of hard optimization problems are the Traveling Salesman Problem and the two–dimensional Ising Spin Glass Model [1].

ST is a relatively new addition to the family of numerical methods that aim to approximate solutions to difficult optimization problems. ST has been applied to two such problems: the Random Field Ising Model [8] and the two–dimensional Ising Spin Glass Model [6, 11, 12]. As with any numerical

*Submitted in partial fulfillment of the undergraduate honors degree program in Computer Science.

method, one desires to be reasonably confident with the results. While many possible avenues exist to warrant the correctness of numerical simulations, the present work examines one.

The validity of ST is demonstrated by applying it to an analytically–solvable problem, the IMF. The IMF has several attributes that make it an interesting test case. First, it is easily adapted for computer simulation. A lattice of binary spins can be represented by an array of integers where each element corresponds to a spin and is assigned a $+1$ for an up spin ($\uparrow$) or a -1 for a down spin ($\downarrow$). Second, the IMF displays an interesting phenomenon present in *real* systems - an order–disorder transition, also known as a phase transition. Finally, and perhaps most importantly, this model has been solved exactly [3, 9] (For a discussion of the solution, see [2]). A point worth noting is that the IMF is easily modified into the Ising Spin Glass Model for which exact solutions are not known.

2 Theory and Formulation

ST is a Monte Carlo procedure for the simulation of (finite) systems with many coexisting states. (The IMF at $T < T_c$, for instance, exhibits spin alignment where the majority of the spins can be up *or* down.) One of its salient features is the ability to sample systems ranging from highly disordered to highly ordered (i.e., systems at various temperatures) while not disturbing equilibrium. This has significant advantages over conventional Simulated Annealing (SA). Although in theory it was shown that SA can always reach an optimal solution [4], an implementation, as prescribed by [4], will prove unrealistic. Such a solution would require establishing an annealing schedule in which β's [1] are changed by infinitesimally small amounts. For a simulation, the above requirement implies an infinite investment of computer time and is absurd. It is interesting to note that this β parameter, which is pivotal in the SA scheduling dilemma, also becomes a crucial factor for the successful ST method. The remainder of this section is devoted to the theoretical formulation of ST.

For a finite system, the configuration space, X, is defined by the assignment of independent variables, ς_i, to each lattcie site, $1, \ldots, N$. The variables ς_i take on only two values, $\varsigma_i = \pm 1$, and represnet two possible states of the

[1] $\beta = \frac{1}{Tk}$

lattice site. Next, it is necessary to expand the configuration space by introducing a new parameter, m. (For simplicity, m will be defined over a specified range, $m = 1, \ldots, M$.) With each value of m there is an associated β (inverse temperature). This is a critical point, as it suggests that temperature is a second stochastic parameter. An immediate ramification of the reformulated state space is that a new method for describing system dynamics is required. To be precise, the Hamiltonian of the Ising Model [2],

$$H(X) = -J \sum_{<i,j>} s_i s_j, \tag{1}$$

where J refers to the coupling strength between adjacent sites, becomes

$$H(X,m) = -\beta_m \sum_{<i,j>} s_i s_j - g_m = \beta_m H(X) - g_m. \tag{2}$$

Here g_m is

$$g_m = H(X, \bar{m}), \tag{3}$$

where $\bar{m}$ corresponds to some β between the current β, β_m, and the projected β, say β_{m+1} or β_{m-1}. That is, g_m is an intermediate energy value. Now since β is incorporated into the the calculation of the energy function, H acquires dimensionality, and, for clarity will henceforth be denoted as $\tilde{H}$.

The probability distribution is now described by the relation

$$P(X,m) \propto e^{-H(X,m)}. \tag{4}$$

Indeed, for specific m, $\tilde{H} \equiv H(X,m) = H(X)$, the Hamiltonian of the Ising Model at some temperature, $\beta = \beta_m$. Thus, at any m, a state of equilibrium is preserved.

The partition function for the new system is given by

$$\begin{aligned} \tilde{Z} &= \sum_{X,m} e^{-\tilde{H}(X,m)} \\ &= \sum_{X,m} e^{-(\beta_m H(X) + g_m)} \end{aligned} \tag{5}$$

[2] The following characterization of the Hamiltonian assumes a zero magnetic field, as this is the case for which analytic solutions are known.

$$= \sum_m Z(\beta_m) e^{-g_m},$$

where $\tilde{Z}$ corresponds to the partition function associated with $\tilde{H}$. Equation (5) is, in fact, a restatement of the assertion that ST does not disturb the equilibrium of the system, as for a specific m, (5) reduces to

$$Z(\beta) = \sum_{all \quad configs.} e^{-\beta H(X)}, \tag{6}$$

the usual sum over all configurations of the corresponding exponentiated Hamiltonians.

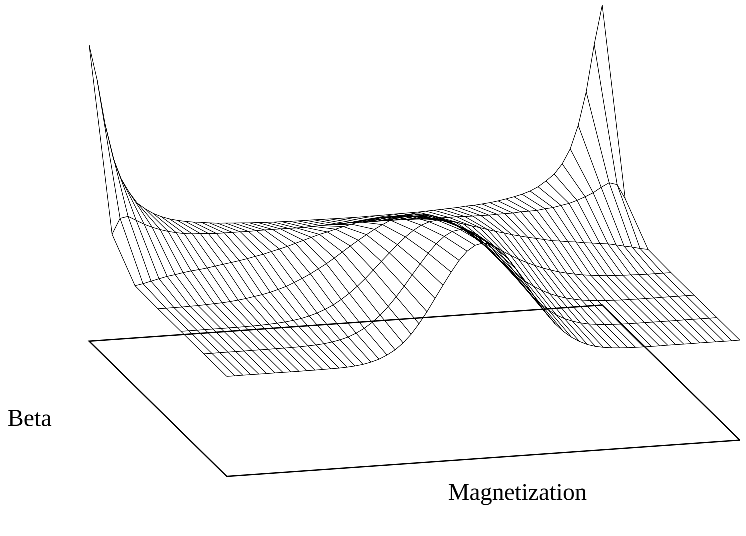

Figure 1: Histogram of magnetization versus β. The two peaks correspond to ground states where most of the spins are up and down, respectively. The region where the peaks first begin to form occurs at $T = T_c$. This is a simulation of an 8×8 lattice.

There is an important subtlety in this formulation. Namely, it is suggested that ST simulates some number (M) of separate models distinguishable by the prescribed β's. It now becomes clear that the parameters m, each with an associated β, and g_m are directly responsible for the ability of the algorithm to freely sample systems at different temperatures. In fact, it is desirable to tune the algorithm so that the probability of visiting systems at all m's approaches a constant value [6, 8].

β	Numerical	Analytic	% Difference
0.03000000	-0.057685185	-0.060090102	1.041
0.03330000	-0.065740741	-0.066723264	1.014
⋮	⋮	⋮	⋮
0.10000000	-0.206013147	-0.203377394	0.337
0.11111112	-0.227636826	-0.226869892	0.337
0.12345680	-0.254595080	-0.253314402	0.503
0.13717422	-0.283791820	-0.283172671	0.220
0.15241581	-0.313938831	-0.317012655	0.979
0.16935090	-0.358689979	-0.355544983	0.877
0.18816768	-0.394593904	-0.399676441	1.029
0.20907520	-0.447417316	-0.450591039	1.007
0.23230578	-0.517002989	-0.509878180	1.379
0.25811754	-0.575974295	-0.579747069	1.007
0.28679728	-0.665661549	-0.663414250	0.337
0.31866365	-0.765798611	-0.765887361	1.000
0.35407073	-0.878739316	-0.895893374	1.020
0.39341194	-1.052083333	-1.074612157	1.021
⋮	⋮	⋮	⋮
0.85000000	-1.990972222	-1.990097027	0.999
0.89999998	-1.992569444	-1.993493517	1.000

Figure 2: A comparison of the internal energies per spin on a 24×24 lattice obtained by simulation and analysis.

β	Number of *sweeps*		Number of *visits*	
	g_m refined	g_m not refined	g_m refined	g_m not refined
0.05882	412400	915600	4124	9156
0.06920	417800	819000	4178	8190
0.08141	431400	709800	4314	7098
0.09578	435500	597500	4355	5975
0.11268	458556	511344	4326	4824
0.13257	454322	416337	4246	3891
0.15596	439884	301968	4073	2796
0.18349	443085	187698	4065	1722
0.21587	461340	93830	4194	853
0.25396	469840	39200	4195	350
0.29878	466440	14260	4056	124

Figure 3: The effect of iterative refinement of the g_m parameter on algorithm performance. This was a simulation of a 24×24 lattice. A total of 50000 sweeps were performed.

3 Conclusions and Discussion

ST [3] is able to reproduce analytic results (internal energies) obtained by the procedure described by Ferdinand and Fisher [3] with impressive accuracy (see Figure 2). At this point, an intuitive discussion may be useful. As Equation (1) suggests, when $J > 0$ ($J = 1$ for these experiments) it is energetically favorable for neighboring spins to be aligned. Consider, for instance, the state space at $T = 0$. Since only nearest neighbors contribute to the Hamiltonian, one would expect a local solution of 2 or -2, that is, both spins up ($\uparrow\uparrow$) or both spins down ($\downarrow\downarrow$). Generally, this sort of cooperative phenomenon can be anticipated at sufficiently low temperatures. In other words, through nearest–neighbor interactions, a magnetic moment can influence the alignment of spins that are separated by a macroscopic distance. This long–ranged order, known as *spontaneous magnetization*, induces a magnetization on the lattice even in the absence of a magnetic field (see Figure 1). Parenthetically, it should be noted that these simulation results arose from separate measurements of H, since, as the reader may recall, the for-

[3] For a detailed description of the algorithm used in this work see [11].

mulation of $\tilde{H}$, the energy function for ST, absorbs the inverse temperature (β).

Though the estimation of parameters β and g_m was not the focus of this work, their importance cannot be understated [6, 8, 11, 12]. Due to the simplicity of the IMF, the rather primitive parameter determination described in [11] proved to be sufficient. Nevertheless, an iterative refinement procedure for the g_m's improved the performance of the algorithm. Here, in subsequent visits to β_m, the newly computed g_m is averaged in with the previously determined value. The effect of this approach is illustrated in Figure 3.

Acknowledgements

Thanks goes to Dr. Bernd Berg for suggesting the premise for this research; to Dr. R.C. Lacher for many helpful comments regarding the technical writing aspects of this paper; to Brian Gorman and Claus Vohwinkel for numerous intriguing discussions; to Dr. Saul Youssef for more encouragement and support than would be customary to list here. All of the simulations were performed on IBM RS6000 workstations at the Supercomputer Computations Research Institute.

References

[1] Bachas, C. P. Spin–glass model of crystal surfaces. *Phys. Rev. Lett.*, 54(1):53–55, 1985.

[2] Chandler, D. *Introduction to Modern Statistical Mechanics.* Oxford University Press, New York, 1987.

[3] Ferdinand, A. E. and Fisher, M. E. Bounded and inhomogeneous ising models. i. specific–heat anomaly of a finite lattice. *Phys. Rev.*, 185(2):832–846, 1969.

[4] Geman, S. and Geman, D. Stochastic relaxation, gibbs distributions and the bayesian restoration in images. *IEEE Trans. Patt. Anal. Mac. Int.*, 6:721–741, 1984.

[5] Ingber, L. Simulated annealing: Practice versus theory. *Statistics and Computing*, 1993.

[6] Kerler, W. and Rehberg, P. Simulated–tempering approach to spin–glass simulations. *(preprint)*, 1993.

[7] Kirkpatrick, S., Gellat, C. D., Jr., and Vecchi, M. P. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[8] Marinari, E. and Parisi, G. Simulated tempering: A new monte carlo scheme. *Europhys. Lett.*, 19:451, 1992.

[9] Onsager, L. Crystal statistics i. a two–dimensional model with an order–disorder transition. *Phys. Rev.*, 65:117–149, 1944.

[10] Otten, R. H. J. M. and van Ginneken, L. P. P. P. *The Annealing Algorithm.* Kluwer Academic Publishers, Dordrecht, 1989.

[11] Shore, M. Coddington, P. D., Fox G. C., Marinari, E. and Wu, C. H. A new automatic simulated annealing–type global optimization scheme. *Journal of Undergraduate Research in High Performance Computing*, 3, 1993.

[12] Wu, C. H. On a study of simulated tempering as an optimization method. *Technical Report, NPAC, Syracuse University*, 1993.