Appl. Math. and Comp. Sci., 1994, vol.4, No.3, 423–431

STATISTICAL PHYSICS APPROACH TO OPTIMIZATION PROBLEMS

LESZEK CIEPLIŃSKI*, CZESŁAW JEDRZEJEK*

Methods of statistical physics are applied to the Travelling Salesman Problem. The starting point is the Hopfield-type Hamiltonian and the most representative benchmark is the 318-city TSP. We find that the recent Ising model neural network implementation by Mehta and Fulop (1993) can be made fully equivalent to the Potts representation proposed by Peterson and Södeberg (1989). Our calculations using the mean-field method for the Potts representation are more effective (average cost 58000, the best 55000) than the Ising model neural network implementation by Mehta and Fulop (the cost from 64552 to 61337). In terms of the tour cost a genetic-type algorithm always gives better results than the Hopfield approach. We relate distribution of energy in the population during the evolution to the quality of genetic algorithm.

1. Introduction

The purpose of this work is the application of statistical physics methods to optimization problems and operation research, in particular to the Travelling Salesman Problem (TSP) (Lawler *et al.*, 1985).

The most important problems in this field belong to the class of NP-complete ones, i.e. time of their solution grows exponentially with the number of elements. The TSP can be stated as follows: given distribution of distances between N cities find the shortest tour that visits each city exactly once. Although the exact solutions are known for certain class of problems up to 2392 cities (Padberg and Rinaldi, 1991), for large N the cost becomes prohibitive. In need of acceptable although not exact solutions statistical physics methods play the prominent role. Various methods are in use including the neural nets of the Hopfield (Hopfield and Tank, 1985) or Kohonen (e.g. Angeniol *et al.*, 1988) type, elastic net (Durbin and Willshaw, 1987), simulated annealing (Kirkpatrick *et al.*, 1983; Kirkpatrick, 1984) or the mean-field. In this work we concentrate on the comparison of two different approaches to the Hopfield-type Hamiltonian that claimed relative success in optimizing the TSP. We mostly benchmark the methods with the use of the popular 318-city TSP, for which the best known cost is 41345 (Lin and Kerninghan, 1973; Crowder and Padberg, 1980). We also study genetic algorithms in context of the relation between distribution of energy in the population during the evolution and the quality of algorithm. The direction seems to be promising in view of the existing relationship between the stochastic dynamics and the statistical mechanics of disordered systems.

* Institute of Physics, Jagellonian University, ul. Reymonta 4, 30–059 Kraków, Poland

2. Analysis of the Hopfield and Tank Type Hamiltonians

Application of neural networks to optimization problems started with Hopfield and Tank (1985) who proposed the following energy function

$$E_{HT} = \frac{A}{2}\sum_{x,i\neq j} n_{xi}n_{xj} + \frac{B}{2}\sum_{i,x\neq y} n_{xi}n_{yi} + \frac{C}{2}\left(\sum_{x,i} n_{xi} - N\right)^2 + \frac{D}{2}\sum_{x,y,i} d_{xy}n_{xi}(n_{y,i+1} + n_{y,i-1}) \quad (1)$$

where n_{xi} is a pseudospin (occupation) variable for visiting city x in the n-th step, d_{xy} is the distance between city x and city y, in the last term the i variable is modulo N. The first (second) term is zero if there is no more than one 1 in each row (column). The third term is zero if there are exactly N 1's, and the last term is equal to the cost of the route. The coefficients A, B, C, D are positive, but apart from this not precisely defined. Hopfield and Tank used $A = B = 500$, $C = 200$ and $D = 500$ in their work. In spite of initial claims the method generally failed for $N > 10$. Various remedies were proposed. Attempts to improve the Hopfield-Tank model and remove the illegal configurations by artificially large weights d_{xx} (Wilson and Pawley, 1988) that made it possible to eliminate xx sequences or by adding extra terms to the Hamiltonian (Xu and Tsai, 1991), eliminating the xyx sequences were not fully satisfactory.

Recently, Mehta and Fulop (1993) proposed the following Lyapunov function that removed some problems inherent in the original Hamiltonian of Hopfield and Tank (1985)

$$E_{MF} = \frac{H}{2}\sum_{x,i\neq j} n_{xi}n_{xj} + \frac{V}{2}\sum_{i,x\neq y} n_{xi}n_{yi} - G\sum_{x,i} n_{xi} + \frac{D}{2}\sum_{x,y,i} d_{xy}n_{xi}(n_{y,i+1} + n_{y,i-1}) + \frac{S}{2}\sum_{x,i} n_{xi}n_{xi} \quad (2)$$

More precisely, this energy was used to the Hamiltonian cycle problem which is a special case for TSP. It can be transformed into TSP by completing the connection graph to a fully connected graph and setting weights 0 to the original edges and weights 1 to new ones. The problem is defined by specifying the cost matrix d_{xy} and average connectivity between cities c.

In contrast to the Hopfield and Tank parametrization, Mehta and Fulop also gave a relation for coefficients that produce stable solutions. For $G = 1.0$, $V = H = 0.7$, $S = 0.6$ and $D = 0.05 - 0.5$ they achieved 100% of legal solutions.

Among the Hopfield-type Hamiltonians the Potts representation proposed by Peterson seems to be the most attractive (Peterson and Södeberg, 1989; Peterson, 1990)

$$E_{PS} = -\frac{\beta}{2}\sum_x\sum_i n_{xi}^2 + \frac{\alpha}{2}\sum_i\Big(\sum_x n_{xi}\Big)^2 + \sum_{x,y}\frac{d_{xy}}{2}\sum_i n_{xi}(n_{y,i+1} + n_{y,i-1}) \tag{3}$$

The advantage of this approach is that the Potts variable exactly fulfils the legality condition for the i-th variable

$$\sum_i n_{xi} = 1 \tag{4}$$

The corresponding mean-field equation is as follows:

$$\bar{n}_l = \left(1 + \exp\Big(-\frac{1}{k_B T}\frac{\partial E[\bar{n}]}{\partial \bar{n}_l}\Big)\right)^{-1} \tag{5}$$

where l denotes two indices x, i, k_B is the Boltzmann constant, T temperature, and $\bar{n}_l$ is average population at site l.

For d_{xy} normalized to a unit square and parameter values $\alpha = 1.0$ and $\beta = 0.5$ one rarely obtains 100% of legal solutions (for large problem size) but these solutions are not too far from optimal so a hybrid approach e.g. with the use of *greedy heuristics* or *2-opt* does not make the cost function considerably worse.

Below we will compare the Mehta – Fulop (1993) and Peterson – Södeberg (1989) Hamiltonians. Before using the Potts constraint $\sum_i n_{xi} = 1$ the original Peterson and Södeberg Hamiltonian $E_{PS}^{(0)}$ is of the form

$$E_{PS}^{(0)} = \frac{\beta}{2}\sum_x\sum_{i\neq j} n_{xi}n_{xj} + \frac{\alpha}{2}\sum_i\Big(\sum_x n_{xi} - 1\Big)^2 + \sum_{x,y}\frac{d_{xy}}{2}\sum_i n_{xi}(n_{y,i+1} + n_{y,i-1}) \tag{6}$$

Writing the second term as (and dropping 1 as the constant)

$$\begin{aligned}\Big(\sum_x n_{xi} - 1\Big)^2 &= \sum_{xy} n_{xi}n_{yi} - 2\sum_x n_{xi} + 1 \\ &\approx \sum_x {n_{xi}}^2 + \sum_{x\neq y} n_{xi}n_{yi} - 2\sum_x n_{xi}\end{aligned} \tag{7}$$

we obtain

$$\begin{aligned}E_{PS} &= \frac{\beta}{2}\sum_x\sum_{i\neq j} n_{xi}n_{xj}\sum_{x,y}\frac{d_{xy}}{2}\sum_i n_{xi}(n_{y,i+1} + n_{y,i-1}) \\ &+ \frac{\alpha}{2}\sum_x\sum_i {n_{xi}}^2 + \frac{\alpha}{2}\sum_{x\neq y}\sum_i n_{xi}n_{yi} - \alpha\sum_x\sum_i n_{xi}\end{aligned} \tag{8}$$

This has the same Ising spin form as the Mehta and Fulop Hamiltonian if the following equivalences are used

$$D \longrightarrow 1.0, \quad H \longrightarrow \beta, \quad V \longrightarrow \alpha, \quad G \longrightarrow \alpha, \quad S \longrightarrow \alpha.$$

As a consequence one can compare the methods of solutions of the same effective Hamiltonian.

For values of coefficients scaling the Mehta and Fulop Hamiltonian, by taking $D = 1.0$ and multiplying the rest of coefficients by $\frac{10}{7}$, we obtain

$$H = 1.0 = 2\beta, \quad V = 1.0 = \alpha, \quad G = 1.3 \approx \alpha, \quad S = 0.86 \approx \alpha$$

The largest difference (by a factor of 2) appears in the first term. The difference in the third term is less important because it is dropped from the Potts Hamiltonian, anyway.

We verify that using $\beta = 1$ for the Potts Hamiltonian makes the results worse. We first compare our results for the mean field Potts model with the Södeberg and Peterson results for 50 and 100 cities with $\alpha = 1, \beta = 0.5$ (see Fig. 1). They are approximately of the same quality (strictly, they are better than in (Södeberg and Peterson, 1989) because of breaking the symmetry of the direction of a tour and taking $n_{x1} = 1$ for one selected city x, but a little worse than in (Peterson, 1990), where only one distribution of cities was used. Our calculations using the mean-field method for the Potts representation for the 318 city problem are more effective (average cost 58000, the best 55000) than the Ising model neural network implementation by Mehta and Fulop (1993) (the cost from 64552 to 61337). We did not recalculate the Ising scheme here. It would be interesting to investigate why the Potts scheme still ends with illegal matrices using the mean-field approach while Mehta and Fulop reported 100% of legal matrices.

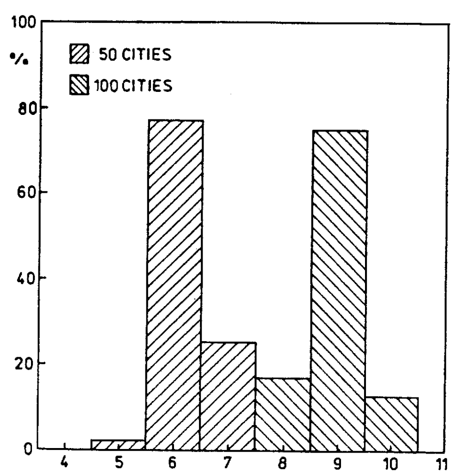

Fig. 1. Histograms based on 50 experiments for 50 and 100 cities TSP.

3. Genetic algorithm

Genetic algorithms are increasingly popular in solving classical optimization problems (Grefenstette *et al.*, 1985; Muhlenbein *et al.*, 1991; Boseniuk *et al.*, 1987; Bac and Perov, 1993; Pal, 1993; Arabas, 1993) and different solutions have been proposed concerning the representation of each individual, the size of the population, the crossover and mutation operators, and the initialization strategy. The most complete investigation in this respect to the TSP was performed by Prinetto *et al.* (1993). Hybrid approaches are also used with the goal of including heuristic techniques into the pure genetic algorithm schemes. In the coding strategy we use two representations: (1) path representation for initialization (Grefenstette *et al.*, 1985) defined as follows: the tour is described by a vector of N integers whose i-th element holds the value j if the city j is reached at the i-th step; (2) adjacency representation after initialization (Grefenstette *et al.*, 1985) defined as: the tour is described by a vector of N integers whose i-th element holds the value j if i precedes j in the tour.

We use heuristic crossover (Grefenstette *et al.*, 1985). The starting city for the tour is randomly chosen; then the next visited city is the nearest one along one of the parent tours that has not been visited in yet the offspring tour. In most simulations we use a modified version of the above, in which the choice between the two possibilities is made at random with weights inversely proportional to the corresponding distances (Pal, 1993).

We perform some experiments on different forms of mutation. Similarly as in (Pal, 1993), we use (1) 2-opt mutation or (2) series of 2-opt mutations until no such mutation can improve the tour. Using the genetic algorithm with the first mutation scheme it is relatively easy to bring the result to 43000 for the 318-city problem. The second mutation scheme, as already established by Pal (1993), eventually brings the result to the exact length of 41345 (not 41269 as given by Mehta and Fulop, 1993), see Fig. 2. Here we concentrate on energy density distributions (which in the TSP case are equivalent to the tour length distributions) during the population evolution as functions of type of an algorithm. It was shown recently by Prügel-Bennett and Shapiro (1994) that the energy density distribution has the dominant effect on the speed of convergence. We study the same effect for type (1) mutation scheme obtained using 10000 individuals. In Figs. 3 and 4 we show the energy distribution during the evolution for probability of mutation equal to 0.1 and 0.9, respectively. It is seen that energy distribution looses symmetry rapidly for both cases. The reason for this is the type of the genetic algorithm we used. Unlike Prügel-Bennett and Shapiro, we use steady-state replacement, i.e. we introduce a new individual to the population immediately after its creation rather than build a new population that replaces the old one (Beasley *et al.*, 1993). The second important fact to notice is that in both instances the energy distribution is very similar during the beginning stages of evolution. Then the crossover is much more effective than mutation. The 2-opt mutations are efficient only for tours not far from the optimum.

4. Conclusions

In this work we concentrate on the comparison of two different approaches to the Hopfield-type Hamiltonian that claimed relative success in optimizing the TSP. We

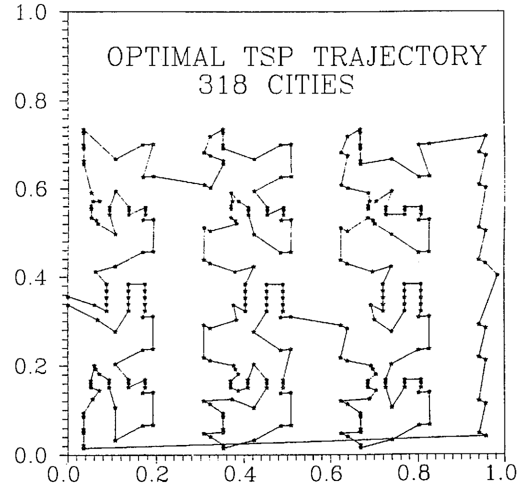

Fig. 2. Optimal trajectory for 318-cities TSP problem of Lin and Kerninghan (1973).

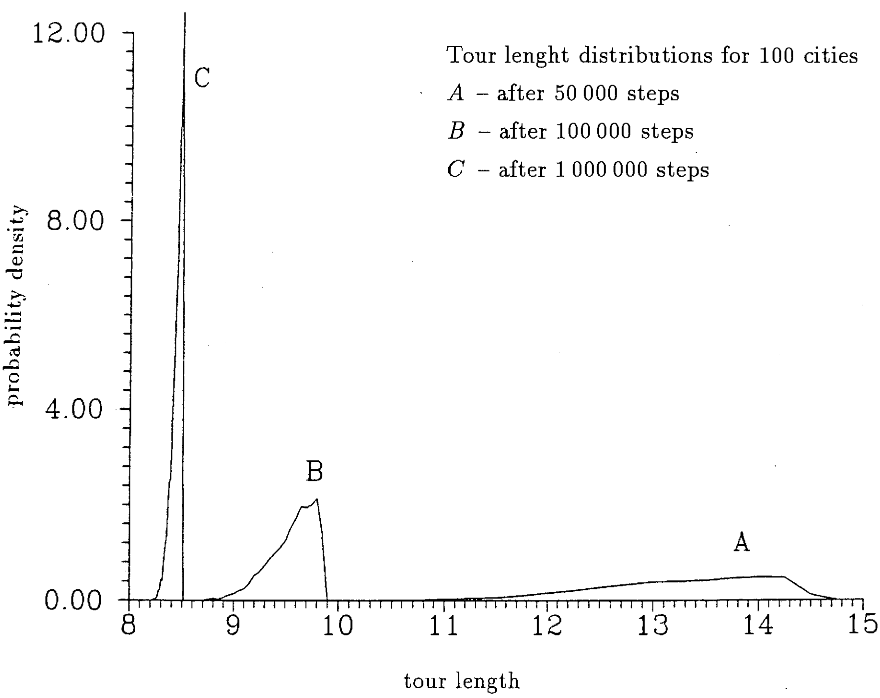

Fig. 3. Distribution of energy (tour lengths) for 10 000 individuals and probability of mutation 0.1.

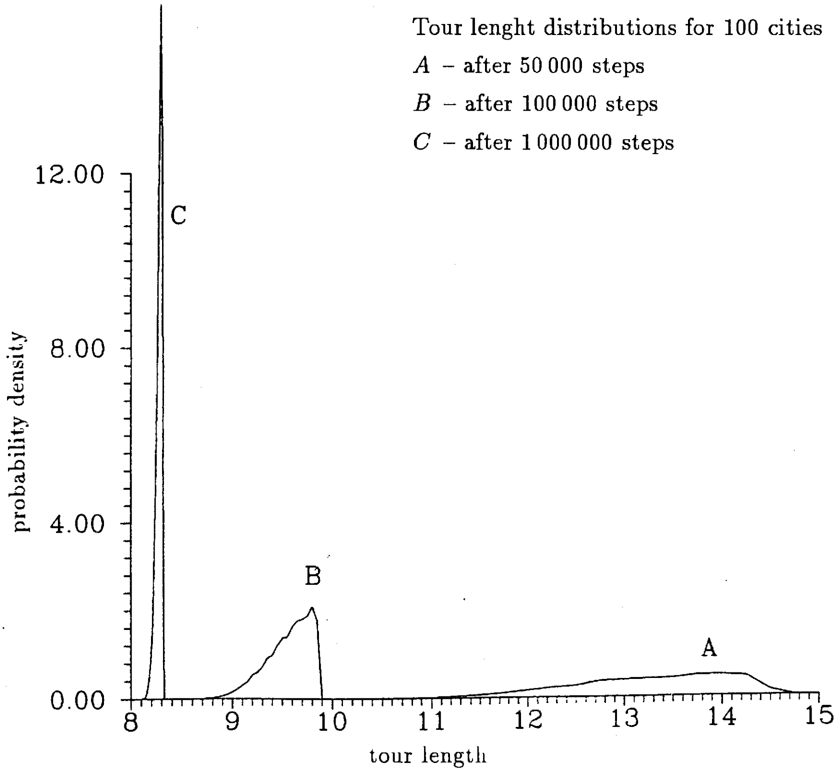

Fig. 4. Distribution of energy (tour lengths) for 10 000 individuals and probability of mutation 0.9.

mostly benchmark the methods with the use of the popular 318-city TSP for which the best known cost is 41345. For the Hopfield type Hamiltonian we find that the mean field approach for the Potts implementation (however with illegal matrices that are next corrected by the 2-opt) gives better results than the neural network Ising type implementation. We also study genetic algorithms in context of the relation between distribution of energy in the population during the evolution and the quality of algorithm. We show that the 2-opt mutations cannot improve starting tours consisting of random sequences of cities and they are efficient only for tours not far from the optimum. We plan a systematic study of effect of various crossover and mutation schemes on distribution of energy in the population during the evolution.

One way to improve any scheme for the Hopfield type Hamiltonian is to include the long range correlations between the n_{xy} variables e.g. by the use the local density method (Jedrzejek and Ciepliński, 1994). The same correlations will be useful for genetic algorithms that presently give better results than those coming from the neural network approach (Bac and Perov, 1993; Pal, 1993). It has recently been shown that such correlations seem to be much more important for optimization problems, such as office assignment problem compared to the TSP (Altschuler *et al.*, 1994).

References

Altschuler E.L, Williams T.J., Ratner E.R., Dowla F. and Wooten F. (1994): *Method of constrained global optimization.* — Phys. Rev. Lett., v.72, pp.2671–2674.

Angeniol B., Vaubois G. and Texier J. (1988): *Self-organizing feature maps and the Traveling Salesman Problem.* — Neural Networks, v.1, pp.289–293.

Arabas J. (1993): *A genetic approach to the Hopfield neural network in the optimization problems.* — Proc. 1st Nat. Conf. *Circuit Theory and Electronic Circuits*, Kołobrzeg, (Poland), pp.594–599.

Bac F.Q. and Perov V.L. (1993): *New evolutionary genetic algorithms for NP-complete combinatorial problems.* — Biol. Cybern., v.69, pp.229–234.

Beasley D., Bull D.R. and Martin R.R. (1993): *An overview of genetic algorithms: Part 1, Fundamentals.* — University Computing, v.15, pp.58–69.

Boseniuk T., Ebeling W. and Engel A. (1987): *Boltzmann and Darwin strategies in complex optimization.* — Physics Letters A, v.125, pp.307–310.

Crowder H. and Padberg M.W. (1980): *Solving large scale symmetric travelling salesman problems to optimality.* — Management Sci., v.26, pp.496–509.

Durbin R. and Willshaw D. (1987): *An analog approach to the travelling salesman problem using an elastic net method.* — Nature, v.326, pp.681–691.

Grefenstette J., Gopal R., Rosmaita B. and Van Gucht D. (1985): *Genetic Algorithms for the Traveling Salesman Problem.* — Proc. 1st Int. Conf. *Genetic Algorithms*, Pittsburgh, PA (USA), pp.160–168.

Hopfield J.J. and Tank D.W. (1985): *Neural computation of decisions in optimization problems.* — Biol. Cybern., v.52, pp.141–152.

Jedrzejek C. and Ciepliński L. (1994): *Application of statistical physics methods to optimization problems.* — Proc. 1st Nat. Conf. *Neural Networks*, Częstochowa, (Poland), pp.287–292.

Kirkpatrick S. (1984): *Optimization by simulated annealing: quantitative studies.* — J. Statistical Physics, v.34, pp.975–986.

Kirkpatrick S., Gelatt C.D. and Jr. Vecchi M.P. (1983): *Optimization by simulated annealing.* — Science, v.220, pp.671–680.

Lawler E.L., Lenstra J.K., Rinnoy A.G.H. and Shmoys D.B. (1985): *The Travelling Salesman Problem: A Guided Tour of Combinatorial Optimization.* — New York: J. Wiley.

Lin S. and Kerninghan B.W. (1973): *An effective heuristic algorithm for the travelling salesman problem.* — Operations Research, v.21, pp.498–516.

Mehta S. and Fulop L. (1993): *An analog neural network to solve the hamiltonian cycle problem.* — Neural Networks, v.6, pp.869–881.

Muhlenbein H., Schomisch M. and Born J. (1991): *The parallel genetic algorithm as function optimizer.* — Parallel Computing, v.17, pp.619–632.

Padberg M.W. and Rinaldi G. (1991): *A branch-and-cut algorithm for the resolution of large-scale symmetric travelling salesman problem.* — SIAM Review, v.33, pp.1–100.

Pal K.F. (1993): *Genetic algorithms for the travelling salesman problem based on a heuristic crossover operation.* — Biol. Cybern., v.69, pp.539–546.

Peterson C. (1990): *Parallel distributed approach to combinatorial optimization: benchmark studies on traveling salesman problem.* — Neural Comp., v.2, pp.261–269.

Peterson C. and Södeberg B. (1989): *A new method for mapping optimization problems onto neural networks.* — Int. J. Neural Systems, v.1, pp.3–22.

Prinetto P., Rebaudengo M. and Reorda M.S. (1993): *Hybrid genetic algorithms for the traveling salesman problem.* — Proc. Int. Conf. *Artificial Neural Nets and Genetic Algorithms*, Insbruck, (Austria), pp.559–566, Berlin: Springer-Verlag.

Prügel-Bennett D. and Shapiro J.L. (1994): *Analysis of genetic algorithms using statistical mechanics.* — Phys. Rev. Lett., v.72, pp.1305–1309.

Wilson G.V. and Pawley G.S. (1988): *On the stability of the travelling salesman problem algorithm of Hopfield and Tank.* — Biol. Cybern., v.58, pp.63–70.

Xu X. and Tsai W.T. (1991): *Effective neural algorithm for the travelling salesman problem.* — Neural Networks, v.4, pp.193–205.