Scientific Research

Information in the Traveling Salesman Problem

Gilad Barach[1], Hugo Fort[2], Yoni Mehlman[1], Fredy Zypman[1]

[1]Department of Physics, Yeshiva University, New York, USA
[2]Departamento de Física, Facultad de Ciencias, Montevideo, Uruguay
Email: zypman@yu.edu

ABSTRACT

In the Simulated Annealing algorithm applied to the Traveling Salesman Problem, the total tour length decreases with temperature. Empirical observation shows that the tours become more structured as the temperature decreases. We quantify this fact by proposing the use of the Shannon information content of the probability distribution function of inter-city step lengths. We find that information increases as the Simulated Annealing temperature decreases. We also propose a practical use of this insight to improve the standard algorithm by switching, at the end of the algorithm, the cost function from the total length to information content. In this way, the final tour should not only be shorter, but also smoother.

Keywords: Traveling Salesman Problem; Shannon Entropy

1. Introduction

The Traveling Salesman Problem (TSP) seeks to find the tour order in which to visit all the cities of a given set so that the total length traveled is the smallest possible. The problem was posed centuries ago and has enjoyed constant attention and consistent progress en route for better solutions [1,2]. Generally speaking, the optimal solution can be found for sets containing a few dozens cities. However, for sets larger than that, and due the non-polynomial growth of tours permutations with set size, solutions to the TSP are approximate: in this case one is satisfied with finding a good tour instead of the best tour. Although perhaps not fully theoretically appealing, this practical approach to the TSP is the one taken by applied mathematicians when providing tours to delivery trucks that need to visit a prescribed set of addresses, or microscope probes that need to reach specific sample locations. Practical solutions of these types reduce waste and are thus of great value to the commercial and/or scientific enterprise.

Here we consider the randomly generated TSP instances problem with Euclidean distance [3], for which the optimal tour length grows as $0.712\sqrt{N}$ with the number of cities N [4,5].

The last ten years have seen a renewed surge in the interest in the TSP in its original form and variations, and the algorithms developed are so efficient that one doubts anything new can be said on the topic. However, there is an unexploited perspective that can give new insights into the problem and suggest improved algorithms. This perspective is proposed here and it is based on the information content in a given tour. If one picks a tour by a random procedure, first most likely that tour will be long because long tours are common while the short ones are rare, and second the longer tours by its own nature of passing through more between city points are less structured, even less visually pleasing due to presence of a large number of sharp turns, than shorter tours. This qualitative description can be made quantitative by appealing to the Negentropy, or information content [6]. We suggest here the use of the Negentropy of a tour as a measure of its structural order. The paper is organized as follows. In Section 2, we study the probability density function (PDF) of step lengths in a tour. These PDFs are then studied as a function of time, that is the time step of a tour optimization algorithm. We identify an analytical expression for these PDFs at all times. In Section 3, we propose to quantify the information content by the use of Negentropy [7], which is defined through Shannon's information [8]. We find an analytical expression of the information content. We also monitor the evolution of both total tour length and Negentropy during optimization. In Section 4, we suggest practical applications where by monitoring the tour information content can provide smoother tours. Section 5 presents the conclusions.

2. Simulated Annealing and PDF of Steps

The TSP has been optimized by genetic algorithms, direct steepest descendant, insect swarm algorithms, etc.

Here we use a program written by us based on the extensively tested Simulated Annealing (SA) adapted to the TSP [9]. In a generic minimization by SA [10], a cost depending on many parameters is to be minimized. First, random values of the parameters are chosen and the corresponding cost is evaluated. Subsequently elemental changes of the parameters are performed. If the effect of the change is to reduce the cost, then the change is accepted. If, on the other hand, the effect of the change is to increase the cost, then the change is accepted with the Boltzmann's probability $\exp(-\Delta C/T)$, where T is a parameter representing an absolute temperature and Boltzmann's constant has been chosen as unity in that scale. At the beginning of the algorithm, the temperature is kept sufficiently high so that the system explores a large region of parameter space without getting trapped in a local minimum. Eventually, after a preset number of steps for example a thousand, the temperature is reduced to a fraction of its previous value, typically 90%. Then the process is repeated, and as the temperature decreases, the system gets confined in subregions of parameter space. The program stops when the cost value stagnates. In the particular implementation of SA to TSP, we use parameter changes that revert the direction of travel between two randomly chosen nodes, or that two randomly chosen consecutive nodes are visited at a different part of the path, also randomly chosen [11,12].

Thus the SA-TSP algorithm generates a tour at each time step. We analyze each such tour and produce histogram of steps—that is, a histogram of all the consecutive-city distances that the traveler takes in the tour. **Figure 1** shows such a histogram, in this case for the initial randomly chosen tour, before the SA-TSP program begins.

The mean length is known to be

$\frac{1}{3}\ln(1+\sqrt{2})+\frac{1}{15}(2+\sqrt{2})\approx 0.52$ [13], for distances between random points on the unit square, quite in agreement with our numerical results as seen in **Figure 1**. We would like to underscore the large width of the PDF, relative to its mean. As the SA-TSP algorithm progress, the tours tend to contain steps more narrowly distributed. **Figure 2** shows the PDF for random initial conditions for 2048 cities, after the program was run during 86 temperature decrements.

The increased narrowness of the PDFs for shorter tours is a consequence of the smoother nature of better tours. This can be seen in **Figure 3**. There we show four snapshot of SA-TSP with 64 cities: at the beginning of the program, at two intermediate times, and at the end. We see that as the tour becomes shorter, the step size distribution becomes less spread. In addition, the figure gives further insight; there is a suggestion that the better tours convey a message while the longest tours barely
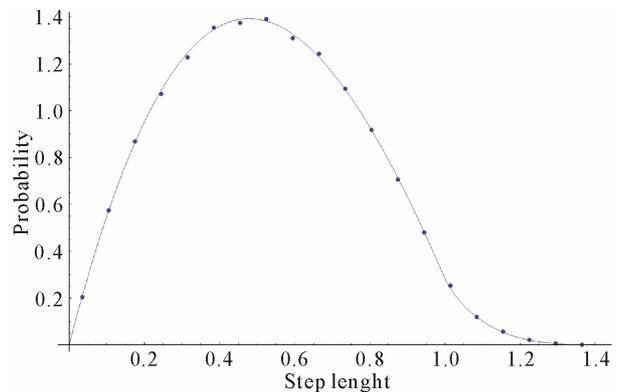
Figure 1. Histogram (dots) of step lengths for 10^5 nodes randomly distributed in the unit square. No optimization has yet been made. The continuous line is from the analytical expression for the PDF of step lengths for tightly packed points on a square [13].
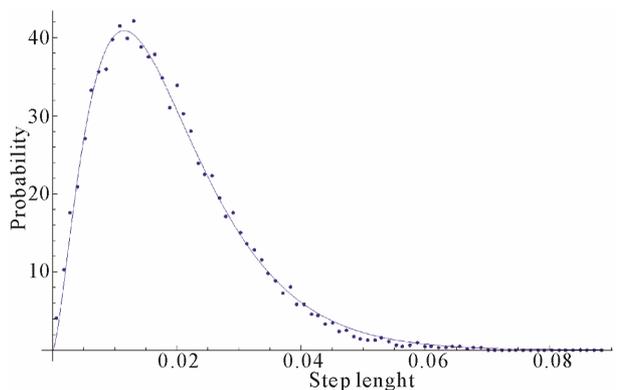
Figure 2. Histogram of steps for optimum SA for a 2048 cities tour TSP problem. The continuous probability distribution function is explained in Section 3.

communicate anything. It seems as if the shortest tours carried more information. Although this is a vague statement, we will make it quantitative in the next section.

3. Negentorpy and Physical Insight

In this section we follow Wheeler's inspiration that "all things physical are information-theoretical in origin" [14]. To quantify the visual appearance of the tours, we introduce the tour information measure, or Negentropy [8],

$$I = \sum_j p_j \ln(p_j) \qquad (1)$$

where p_j is the probability of having a step length ℓ_j for a given tour of total length

$$L = \sum_j \ell_j \qquad (2)$$

With this definition, we could use the histograms for the tour step lengths introduced in the previous section

64 Cities, Step:9, Length: 22.06

64 Cities, Step:18, Length: 11.3741

64 Cities, Step:25, Length: 8.52285

64 Cities, Step:32, Length: 7.11776
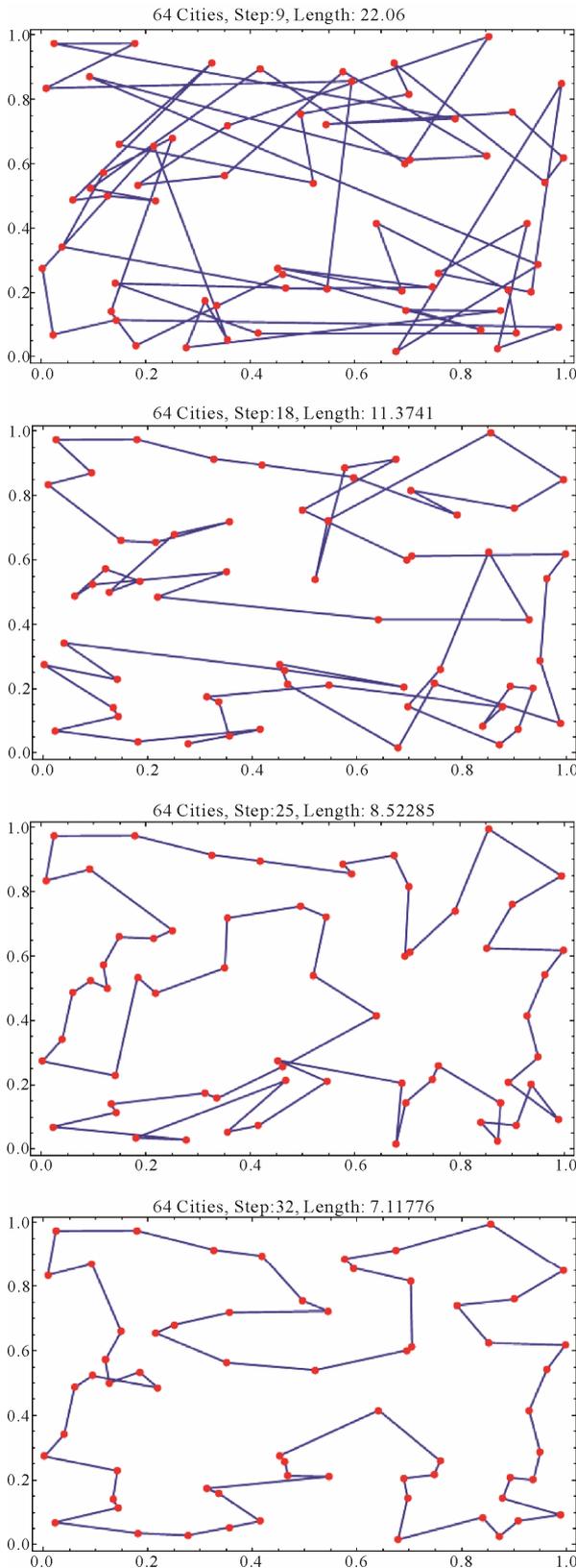
Figure 3. Evolution of TSP tours for four SA time steps, from initially completely random on the top left, to optimum on the bottom right.

however, to gain more insight we shall introduce an analytical description of the problem.

In **Figure 2** we show the histogram of step lengths for the optimal tour (dots), and we introduce a continuous PDF curve of the form

$$P(\ell) = A\ell^{\beta}e^{-\gamma\ell} \tag{3}$$

where A is a normalization constant, $\beta = 1.6$, and $\gamma = 4.3$.

We then studied the SA TSP from 64 to 2048 cities and found that Equation (3) consistently represents the PDF of tour steps at any stage of the SA annealing algorithm. While the form of Equation (3) remains valid throughout the evolution of the SA, the values of β and γ change as functions of the SA time stage.

Therefore, we take Equation (3) as the PDF of TSP tour steps at any stage of the SA. In addition, the Negentropy of a given PDF cannot depend either on its area under the curve (total tour length) or on its mean step size. The Negentropy, on the other hand, should be dependent on the variance and higher moments since they carry (dis)order information. Thus we consider the area and mean normalized version of Equation (3)

$$P(\ell) = \frac{(\beta+1)^{\beta+1}}{\Gamma(\beta+1)}\ell^{\beta}e^{-(\beta+1)\ell} \tag{4}$$

where the parameter γ, that carries information of the mean, disappeared.

This PDF has already been recognized as that of the near neighbors for a one-dimensional gas with logarithmic pair potential. In addition, it is a particular case (rank 2) of a daisy model, in which the PDF is derived by the Poisson spectrum and then removing every r^{th} intermediate levels, where r defines the rank of the derived distribution [15].

Using Equation (1) for the PDF in (4) one obtains,

$$I = 1 + \int_{0}^{\infty} P(\ell)\text{Log}(P(\ell))\text{d}\ell \tag{5}$$

where we have added the arbitrary constant 1 to the definition of Negentropy for later convenience. Then the Negentropy becomes explicitly,

$$I = \beta\left[\Psi(\beta+1)-1\right] + \log\left[\frac{\beta+1}{\Gamma(\beta+1)}\right] \tag{6}$$

where Ψ is the logarithmic derivative function [16].

Figure 4 shows the Negentropy, I for β in the range $0 < \beta < 2$ of interest for the TSP. In the SA program β increases as a function of time and, according to the figure so does I.

4. Practical Considerations

In addition to being able to quantify the previously vague

concept of tour appearance, can the ideas of Section 3 be of any practical use? We argue here that indeed, we can refine the optimum solution obtained by standard SA TSP.

Figure 5 shows both total tour length and negative Negentropy (Entropy) as functions of the SA temperature. As expected, both quantities decrease as the SA temperature decreases: the tour length becomes smaller, and the information increases. Also both quantities tend asymptotically to constant values at small temperatures.

However, we also see that their convergence rates are different: the total length saturates faster than the negentropy. Therefore, one could run the SA program monitoring the total length until it does not change substantially. At that point one would switch to monitor the negentropy instead. However, while the total length is easy to compute, $L = \sum_j \ell_j$, the negentropy would require to build a histogram and then use Equation (1). However, this would be too time consuming for the program. Instead, we use the analytical results of Section 3, to expedite the calculations. The variance of the PDF in Equation (4) is

$$\sigma^2 = \frac{1}{1+\beta} \qquad (7)$$
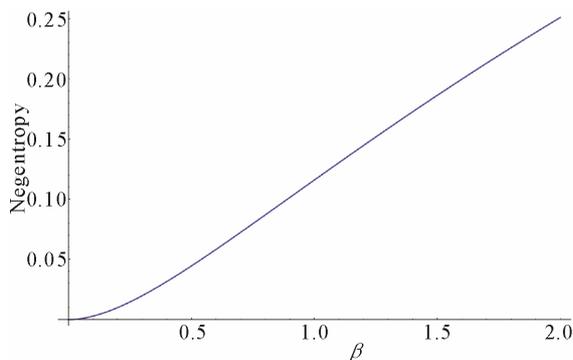
Figure 4. Negentropy as a function of *β*.
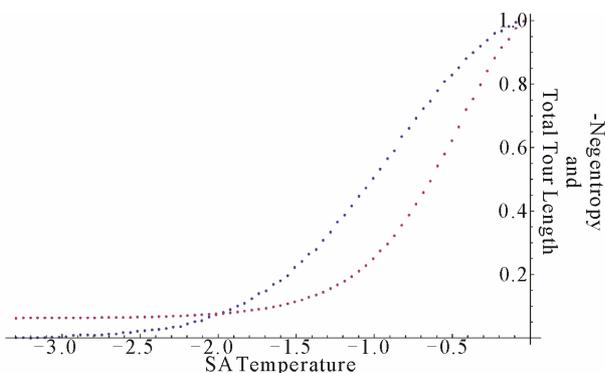
Figure 5. Total tour length (red dots) and entropy (blue dots) as functions of the SA temperature. Both graphs are normalized so they have the same value 1 at high temperatures, before the algorithm begins.

So given the set of step lengths in the program, one computes $\sigma^2 = \left\langle \ell_j^2 \right\rangle - \left\langle \ell_j \right\rangle^2$ and then substitutes in

$$I = \left[\frac{1}{\sigma^2} - 1 \right] \left[\Psi\left(\frac{1}{\sigma^2}\right) - 1 \right] - \log\left[\sigma^2 \Gamma\left(\frac{1}{\sigma^2}\right) \right] \qquad (8)$$

Thus we have a fast way to evaluate information in terms of the firs two moments of the step-length.

The proposed improvement may provide a marginal gain in tour length. But while indeed the total tour length will stay practically unchanged the algorithm will search, out of those many good solutions, for the ones with maximum information. In the end, the traveler will not only travel less, but will also travel more comfortably.

5. Conclusion

We have introduced a new perspective from which to study the Traveling Salesman Problem. Besides using the total tour length as the cost function, we propose to also use the tour information content. The Shannon entropy has been used here to quantify the intuitive fact that generally, shorter tours contain more information. The analytical expression for the information content provides insight into the TSP structure. We have also proposed a practical improvement to the SA TSP where, when an acceptable short tour has been found by the standard algorithm, the program remains running to search among short tours for those with more information. This provides the additional benefit that the tour selected will not only be short but also smoother. This could be relevant, for example, when the TSP is used to design solutions that require mechanical manipulation of an arm to visit points on a surface. Such could be the case when visiting preset locations on a sample with Atomic Force Microscopy [17]. The additional requirement of maximizing information of the tour makes the trajectory less taxing to the mechanical components of the microscope. We end by speculating on a possible fruitful use of the viewpoint proposed here. The use of Negentropy as a measure of the goodness of a tour may find connections with neuroscience since it has been empirically established that animals are able to find good tours instinctively: for example pigeons have been found to obtain acceptable tours [18,19]. Although it has not been established that there is a cognitive map that allows the pigeons to act for future planning, it is clear that whatever their means, animals do need to minimize foraging tours to avoid being prayed.

REFERENCES

[1] W. J. Cook, "In Pursuit of the Traveling Salesman: Mathematics at the Limits of Computation," Princeton University Press, Princeton, 2011.

[2]   H. Fort, M. Kornbluth and F. Zypman, "Traveling Sales-man Problem for Finite-Size Cities," *Mathematical Structures in Computer Science*, Vol. 20, No. 6, 2010, pp. 1067-1078. doi:10.1017/S096012951000037X

[3]   M. N. Ghosh, "Expected Travel among Random Points," *Calcutta Statistical Association*, Vol. 2, 1949, pp. 83-87.

[4]   J. Beardwood, J. H. Halton and J. M. Hammersley, "The Shortest Path through Many Points," *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 55, No. 4, 1959, pp. 299-327. doi:10.1017/S0305004100034095

[5]   A. G. Percus and O. C. Martin, "Finite Size and Dimensional Dependence of the Euclidean Traveling Salesman Problem," *Physical Review Letters*, Vol. 76, 1996, pp. 1188-1191. doi:10.1103/PhysRevLett.76.1188

[6]   L. Brillouin, "The Negentropy Principle of Information," *Journal of Applied Physics*, Vol. 24, No. 9, 1953, p. 1152. doi:10.1063/1.1721463

[7]   P. Fleurquin, H. Fort, M. Kornbluth, R. Sandler, M. Segall and F. Zypman, "Negentropy Generation and Fractality in the Dry Friction of Polished Surfaces," *Entropy*, Vol. 12, No. 3, 2010, pp. 480-489. doi:10.3390/e12030480

[8]   C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, Vol. 27, 1948, pp. 379-423.

[9]   W. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, "Numerical Recipes in C," 2nd Edition, Cambridge University Press, Melboune, New York, 1992.

[10] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, "Optimization by Simulated Annealing," *Science*, New Series, Vol. 220, 1983, pp. 671-680.

[11] M. Held and R. M. Karp, "A Dynamic Programming Approach to Sequencing Problems," *Journal of the Society for Industrial and Applied Mathematics*, Vol. 10, No. 1, 1962, pp. 196-210. doi:10.1137/0110015

[12] S. Lin, "Computer Solutions of the Traveling Salesman Problem," *Bell Systems Technical Journal*, Vol. 44, 1965, pp. 2245-2269.

[13] J. Philip, "The Probability Distribution of the Distance between Two Random Points in a Box," Technical Report Number TRITA MAT 07 MA 10, 1991. http://www.math.kth.se/~johanph/habc.pdf

[14] J. A. Wheeler, "Information, Physics, Quantum: The Search for Links," *Proceedings of 3rd International Symposium on the Foundations of Quantum Mechanics*, Tokyo, 1989, pp. 354-368.

[15] H. Hernández-Saldaña, J. Flores and T. H. Seligman, "Semi-Poisson Statistics and Beyond," *Physical Review E*, Vol. 60, 1999, pp. 449-452. doi:10.1103/PhysRevE.60.449

[16] E. Jahnke and F. Emde, "Table of Functions with Formulae and Curves," Dover Publications, New York, 1943, page 18.

[17] F. Zypman, "Fast Atomic Force Microscopy," *Encyclopedia of Nanoscience and Nanotechnology*, Vol. 3, 2004, p. 307.

[18] B. M. Gibson, E. A. Wasserman and A. C. Kamil, "Pigeons and People Select Efficient Routes when Solving a One-Way Traveling Salesperson Task," *Journal of Experimental Psychology: Animal Behavior Processes*, Vol. 33, No. 3, 2007, pp. 244-261. doi:10.1037/0097-7403.33.3.244

[19] H. Miyata and K. Fujita, "Route Selection by Pigeons (Columba livia) in 'Traveling Salesperson' Navigation Tasks Presented on an LCD Screen," *Journal of Comparative Psychology*, Vol. 124, No. 4, 2010, pp. 433-446. doi:10.1037/a0019931