On the Spacefilling Curve Heuristic for the Euclidean Traveling Salesman Problem

Dimitris Bertsimas*
Sloan School of Management, Rm. E53-359
Massachusetts Institute of Technology
Cambridge, MA 02139

Michelangelo Grigni†
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA 02139

18 December 1988

Abstract

Bartholdi and Platzman [3] proposed the spacefilling curve heuristic for the Euclidean Traveling Salesman Problem and proved that their heuristic returns a tour within an $O(\lg n)$ factor of optimal length. They conjectured that the worst-case ratio is in fact $O(1)$. In this note we exhibit a counterexample showing the $O(\lg n)$ upper bound is tight.

*This research is partially supported by the National Science Foundation under grant ECS-8717970.

†This material is based upon work supported under a National Science Foundation Graduate Fellowship.

1 Introduction

Bartholdi and Platzman [3] proposed a heuristic for the Euclidean Traveling Salesman Problem (ETSP) based on a spacefilling curve. Their curve ϕ is a uniformly continuous map from the unit interval to the unit square. For our purposes we define ϕ as a total linear ordering $<_\phi$ on the points of the unit square $[0,1]^2$. This ordering may be defined by a recursive procedure; for any two distinct points (x,y) and (x',y') the procedure will eventually decide which point comes first in the order.

In the unit square, if $x - y \le 0 < x' - y'$, then $(x,y) <_\phi (x',y')$. In figure 1(a), this means all points from the lower right triangle precede points from the upper left triangle. Otherwise, both points fall in the same triangle, and we apply a recursive ordering on points in such a triangle. First rotate and enlarge the triangle into standard position with vertices at $(0,0)$, $(1,0)$, and $(1,1)$ (for this transformation to be uniquely defined, we have to orient the triangles; in our figures this orientation is denoted by an arrow). In the following we refer to this triangle as the 'unit triangle'. If $x + y \le 1 < x' + y'$ in the unit triangle, then $(x,y) <_\phi (x',y')$. In figure 1(b) this means all points from the lower left subtriangle precede points from the upper right subtriangle. Otherwise, both points fall in the same subtriangle. Now this (oriented) subtriangle is similar to the original unit triangle, so we may recurse. After t iterations of taking subtriangles, the subtriangle we are considering has a hypotenuse of length $2^{(1-t)/2}$ (in the scale of the original unit square), hence the process is guaranteed to halt after $O(\log(1/d))$ iterations, where d is the Euclidean distance between the two input points (x,y) and (x',y').

Now we describe the heuristic of [3]: given a set S of n points in the unit square, visit the points in the order ϕ defined above, and finish the tour by returning to the first point. As remarked in [4], this heuristic is very fast. Given point (x,y) input as a pair of k-bit fractions, a $2k$-bit sorting key t may be computed in $O(k)$ bit operations (t is really an inverse $\phi^{-1}(x,y)$ where ϕ is defined as a continuous map from the unit interval onto the unit square). Given n such points, all the keys may be computed and then sorted (by radix-sort [1]) in $O(kn)$ bit operations, i.e. in time linear in the input size.

Let $L^\phi(S)$ be the length of the tour produced by this heuristic, and let $L^*(S)$ be the length of the optimal tour. In [5] they proved that $L^\phi(S)/L^*(S) = O(\lg n)$. They further conjecture that the worst case ratio is in fact $O(1)$. We refute this conjecture by exhibiting a simple set S_n of n points with $L^\phi(S_n)/L^*(S_n) = \Theta(\lg n)$.

2 A $\Theta(\lg n)$ Example

Consider n points uniformly spaced along the line from $(1/3, 1/3)$ to $(1, 1/3)$. Precisely, define the set of points

$$S_n = \left\{ (x_i, y_i) = \left(\frac{1}{3} + \frac{2}{3} \cdot \frac{2i-1}{2n},\ \frac{1}{3} \right) \text{ for } 1 \le i \le n \right\}. \quad (1)$$

This definition of S_n is convenient for our proof, but in fact any reasonably uniform distribution on the line $y = 1/3$ would suffice. The optimal tour on S_n visits the points in the order of increasing i, and has length $L^*(S_n) < 4/3$. It now suffices to show:

Theorem 2.1 *For $n = 2^k$, traversing S_n in the order of curve ϕ produces a tour of length $L^\phi(S_n) > 2k/9$.*

Proof: The basic idea is to recursively decompose the order that the spacefilling curve visits S_n, and observe that at each level there are significantly long 'jumps'.

Let a_k be the length of the path produced by traversing S_n in the order of curve ϕ (note $a_k < L^\phi(S_n)$ since $L^\phi(S_n)$ counts the extra edge used to close the path into a cycle). Let b_k be the length of the path traversed by the heuristic on the following similar set of $n = 2^k$ points:

$$S'_n = \left\{ (x_i, y_i) = \left(\frac{2}{3} + \frac{1}{3} \cdot \frac{2i-1}{2n},\ \frac{2}{3} \right) \text{ for } 1 \le i \le n \right\}. \quad (2)$$

By the recursive definition of ϕ, the points of S'_n are visited in the same order as those in S_n; hence $b_k = a_k/2$.

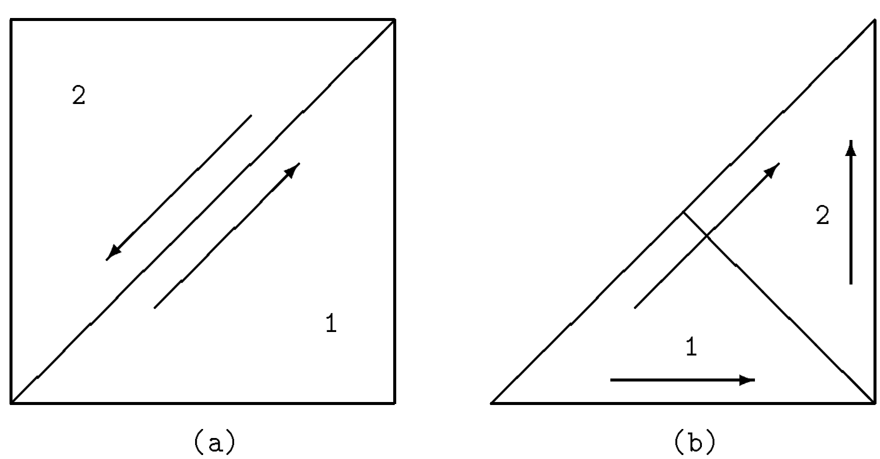

Figure 1: In both the square (a) and the triangle (b), points of subtriangle 1 precede points of subtriangle 2.

We now derive a recursion for a_k. In figure 2, let T denote the entire unit triangle, and let T_1, T_2, T_3, T_4 denote the subtriangles in the order visited by ϕ. The points of S_n lie in subtriangles T_1, T_2, and T_3. Hence a_k is the sum of the path lengths within each subtriangle, together with the lengths of the jumps between consecutive subtriangles.

The $n/4$ points of $S_n \bigcap T_1$ form a half-scale image of $S'_{n/4}$, contributing $b_{k-2}/2$ to a_k. Similarly the $n/4$ points of $S_n \bigcap T_2$ form a *reversed* half-scale image of $S'_{n/4}$. Since ϕ is reversible (i.e. 'reversing the arrow' of a triangle exactly reverses the order of points within that triangle), the length of their path is the same as the path in T_1, contributing another term of $b_{k-2}/2$ to a_k. Finally the $n/2$ points of $S_n \bigcap T_3$ form a half-scale reversed image of $S_{n/2}$, contributing $a_{k-1}/2$ to a_k. Hence

$$a_k = b_{k-2} + a_{k-1}/2 + j_{12} + j_{23} = (a_{k-2} + a_{k-1})/2 + j_{12} + j_{23}$$

where j_{12} is the length of the jump from the last point in T_1 to the first point in T_2, and j_{23} is the length of the jump from the last point in T_2 to the first point in T_3. To estimate these jump lengths we need to know the first and last points of S_n visited in each subtriangle.

Lemma 2.2 *For $n = 2^k$, $k \geq 1$, the first point in S_n (and by similarity S'_n) under order ϕ is (x_1, y_1), and the last point is (x_p, y_p) where $p = (n/2) + 1$.*

Proof: Let the points of S_n be indexed as in the definition (1). By inspection for $k \leq 2$. For $k \geq 3$, we again use figure 2. The first point of S_n in T is the first point

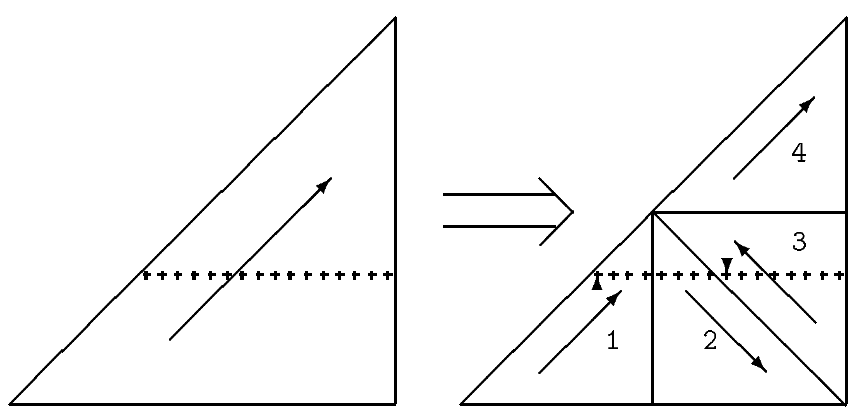

Figure 2: The set S_{16} in the unit triangle T decomposes into a copy of S'_4 in T_1, a reversed copy of S'_4 in T_2, and a reversed copy of S_8 in T_3. Curve ϕ visits the leftmost point first and the ninth point last (marked by arrowheads).

in T_1. The points in $S_n \bigcap T_1$ are $\{(x_i, y_i),\ 1 \le i \le n/4\}$, with identical indices. They form a half-scale image of $S'_{n/4}$, and inductively the first point among them is (x_1, y_1).

The last point of S_n in T is the last point in T_3. The points in $S_n \bigcap T_3$ are a half-scale reversed image of $S_{n/2}$, label them $\left\{(x'_i, y'_i) = (x_{(n/2)+i}, y_{(n/2)+i}),\ 1 \le i \le n/2\right\}$. Since the order is reversed, the last point visited in T_3 corresponds to the *first* point in $S_{n/2}$. Inductively the first point visited in $S_{n/2}$ would be (x'_1, y'_1); this is point $(x_{(n/2)+1}, y_{(n/2)+1})$ of S_n. □

Now we find the jumps distances j_{12} and j_{23}. For $k \ge 3$, we may apply the lemma to the find the first and last points of S_n in T_1, T_2, and T_3. The last point in T_1 has index $i = n/8 + 1$, and the first point in T_2 (since it is reversed) has index $3n/8$. Similarly the last point in T_2 has index $i = n/2$, and the first point in T_3 has index $3n/4 + 1$. Hence $j_{12} = (n/4 - 1)(2/3n) = 1/6 - 2/3n$, $j_{23} = (n/4 + 1)(2/3n) = 1/6 + 2/3n$, so for $k \ge 3$ we have the simple recurrence:

$$a_k = \frac{1}{2}(a_{k-1} + a_{k-2}) + \frac{1}{3}.$$

From the base cases $a_1 = 1/3$, $a_2 = 2/3$ follows that $a_k = 2k/9 + (4/27)(-1/2)^k + 5/27$. Finally we have $L^\phi(S_n) = a_k + 1/3 > 2k/9$ as claimed. □

3 Application to the PTSP

Bertsimas [2] applies the spacefilling curve heuristic to the Euclidean Probabalistic Traveling Salesman Problem which is defined as follows: we are given a set S of n points and a probability p_i for each point $(x_i, y_i) \in S$. A random instance $X \subset S$ is then generated by including each point (x_i, y_i) in X independently with probability p_i. For a tour τ on S, let $L^\tau(X)$ be the length of the tour generated by visiting the points of X in the order followed in τ. Suppose tour σ minimizes $E[L^\sigma(X)]$; the goal is to choose tour τ to get $E[L^\tau(X)]$ close to $E[L^\sigma(X)]$. Let $L^\phi(X)$ be the length when τ was chosen by the spacefilling heuristic. Then by a similar analysis as in section 2 we may show that in the worst case, $E[L^\phi(X)]/E[L^\sigma(X)] = \Theta(\lg n)$.

4 Concluding Remarks

Similar examples hold for other curves in the unit square, in particular the Hilbert curve (figure 3(a)) and 'zig-zag' curve (figure 3(b)) mentioned in [4]. For the Hilbert curve, take n points along the diagonal line $x + y = 2/3$. For the zig-zag curve, take points uniformly along the horizontal line $y = 1/2$. Numerical experiments with the curve ϕ strongly suggest that n points uniformly spaced on a 'random' line across the unit square have expected tour length $E[L^{\phi}] = \Theta(\lg n)$.

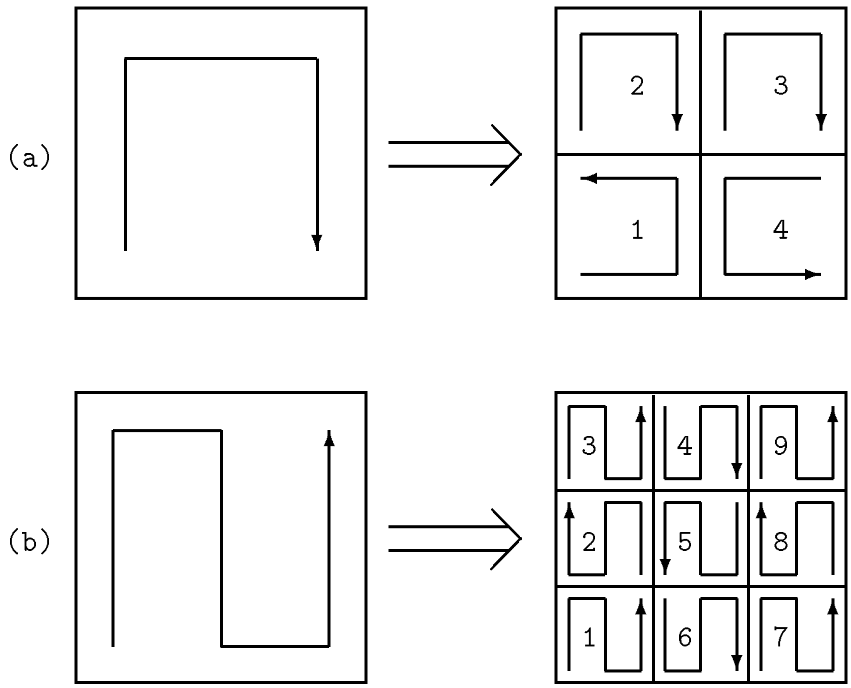

Figure 3: Recursively define the Hilbert curve (a) and the zig-zag curve (b) as several copies of themself. The numbers indicate the relative order of the subcurves.

We conjecture that this is true for 'all' spacefilling curves. More concretely, we conjecture that for any ordering of the vertices of the $n \times n$ mesh, there is some subset of vertices such that the length of their induced tour is an $\Omega(\lg n)$ factor off from optimal.

References

[1] A. Aho, J. Hopcroft, J. Ullman. *The Design and Analysis of Computer Algorithms.* Addison-Wesley, 1974.

[2] D. Bertsimas. Probabilistic Combinatorial Optimization Problems. Ph.D thesis, Technical report 194 (1988), Operations Research Center, Massachusetts Institute of Technology, Cambridge, Mass.

[3] J. J. Bartholdi, L. K. Platzman. An $O(N \log N)$ Planar Traveling Salesman Heuristic Based on Spacefilling Curves. Operations Research Letters, **1** (1982), 121–125.

[4] J. J. Bartholdi, L. K. Platzman. Heuristics Based on Spacefilling Curves for Combinatorial Problems in Euclidean Space. Management Science, **34**, no. 3, (1988), 291–305.

[5] L. K. Platzman, J. J. Bartholdi. Spacefilling Curves and Planar Traveling Salesman Problem. PDRC technical report 83-02, Georgia Institute of Technology, 1983 (to appear in JACM).