MATHEMATICS OF OPERATIONS RESEARCH
Vol. 6, No. 3, August 1981
Printed in U.S.A.

COMPLETE CONVERGENCE OF SHORT PATHS AND KARP'S ALGORITHM FOR THE TSP*†

J. MICHAEL STEELE

Stanford University

Let X_i, $1 \leqslant i < \infty$, be uniformly distributed in $[0, 1]^2$ and let T_n be the length of the shortest closed path connecting $\{X_1, X_2, \ldots, X_n\}$. It is proved that there is a constant $0 < \beta < \infty$ such that for all $\epsilon > 0$

$$\sum_{n=1}^{\infty} P\left(|T_n/\sqrt{n} - \beta| > \epsilon\right) < \infty.$$

This result is essential in justifying Karp's algorithm for the traveling salesman problem under the independent model, and it settles a question posed by B. W. Weide.

1. Introduction. The main objective of the present note is to solve a problem proposed by Weide (1978) concerning the complete convergence of certain random variables associated with Karp's probabilistic analysis of the traveling salesman problem (Karp (1976), (1977)).

To set the problem precisely let X_i, $1 \leqslant i < \infty$, be independent random variables uniformly distributed on the unit square $[0, 1]^2$, and let T_n denote the length of the shortest closed path (in the usual Euclidean distance) which connects each element of $\{X_1, X_2, \ldots, X_n\}$.

It was proved by Beardwood, Halton, and Hammersley (1959) that

$$\lim_{n\to\infty} T_n/\sqrt{n} = \beta$$

with probability one for a finite constant β. This fact was central to the motivation behind Karp's algorithm, but as Weide (1978) points out the Karp algorithm actually calls for the following stronger result to be proved here:

THEOREM 1. *There is a constant β such that for all $\epsilon > 0$, one has*

$$\sum_{n=1}^{\infty} P\left(|T_n/\sqrt{n} - \beta| > \epsilon\right) < \infty.$$

This type of convergence is usually called complete convergence, and Theorem 1 stands in a similar relation to the Beardwood-Halton-Hammersley Theorem as the Hsu-Robbins Theorem stands in relation to the strong law of large numbers (Lukacs (1968), Hsu and Robbins (1947)). The "easy-half" of the Borel-Cantelli lemma shows that Theorem 1 implies the Beardwood-Halton-Hammersley Theorem and the "hard-half" of the Borel-Cantelli lemma shows how Theorem 1 is necessary in modeling contexts where problems of increased size are generated independently of previous problems. (For a full discussion of *independent* versus *incrementing* models for random problems one should consult Weide (1978).)

By an elementary lemma due to Few (1955) the variables $T_n/\sqrt{n}$ are bounded, so the almost sure convergence in the Beardwood-Halton-Hammersley Theorem neces-

* Received August 21, 1979; revised May 15, 1980.
AMS 1980 subject classification. Primary 60D05, Secondary 90C42.
OR/MS Index 1978 subject classification. Primary: 491 Networks/graphs traveling salesman.
Key words. Traveling salesman problem, complete convergence, subadditive processes, subadditive Euclidean functionals, jackknife, Efron-Stein inequality.
† Research supported in part by Office of Naval Research Grant N00014-76-C-0475.

0364-765X/81/0603/0374$01.25

sarily entails $ET_n \sim \beta\sqrt{n}$ as well. This fact, Markov's inequality, and a standard 2ϵ argument show that Theorem 1 is an easy consequence of the following result.

THEOREM 2. *For all $k \geqslant 0$ there is a constant c_k such that*

$$E(T_n - ET_n)^k \leqslant c_k \tag{1.1}$$

for all $n \geqslant 1$.

Actually for Theorem 1 it would suffice to show (1.1) for any $k > 2$, but the present method extends without effort to give the full result. The proof of Theorem 2 is based upon a systematic recursive application of a recent jackknife inequality due to Efron and Stein (1978). This inequality is described in the next section. The third section gives the proof of Theorem 2, and the last section touches briefly on some extensions of these results.

2. Two inequalities. If $S(x_1, x_2, \ldots, x_{n-1})$ is any symmetric function of $n-1$ vectors x_i, and X_i are independent identically distributed random vectors we define new random variables by

$$S_i = S(X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$$

and

$$S_. = \frac{1}{n}\sum_{i-1}^{n} s_i.$$

Efron and Stein (1978) proved the following inequality

$$\operatorname{Var} S(X_1, X_2, \ldots, X_{n-1}) \leqslant E\sum_{i=1}^{n}(S_i - S.)^2. \tag{2.1}$$

This remarkable inequality will no doubt find many applications in the asymptotic analysis of nonlinear processes. By its application in the present case it was possible to both simplify and sharpen an earlier proof of Theorem 1 due to the author. The earlier proof was based on the theory of independent subadditive processes.

The second inequality concerns the variable $d_n = \min\{|X_i - X_1| : 1 < i \leqslant n\}$ where X_i are independent and uniformly distributed in $[0,1]^2$. It is that for all $k \geqslant 0$ there is a constant A_k such that

$$Ed_n^k \leqslant A_k n^{-k/2} \quad \text{for all } n. \tag{2.2}$$

The proof of (2.2) is easy. One just notes by geometry that $P(d_n > x) \leqslant (1 - \alpha x^2)^{n-1}$, even with $\alpha = \pi/4$. Then by generous bounds,

$$Ed_n^k \leqslant \int_0^\infty x^{k-1} e^{-\alpha x^2(n-1)}\,dx \leqslant A_k n^{-k/2}.$$

3. Proof of Theorem 3. The proof of Theorem 3 can now be given. By the Efron-Stein inequality applied to $S = \phi(T_{n-1})$ where ϕ is a smooth function one obtains

$$\operatorname{Var}\phi(T_{n-1}) \leqslant E\sum_{i=1}^{n}(\phi_i - \phi.)^2 \tag{3.1}$$

where $\phi_i = \phi(T(X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$ and $\phi. = (1/n)\sum_{i=1}^{n}\phi_i$. Replacing $\phi.$ by any other function of $X_1, X_2, \ldots, X_n$ only increases the right side of (3.1), so we have

$$\operatorname{Var}\phi(T_{n-1}) \leqslant E\sum_{i=1}^{n}(\phi_i - \phi(T_{n-1}))^2. \tag{3.2}$$

By the mean value theorem it is possible to choose a value $T_{n-1}(i)$ between $T(X_1,X_2,\ldots,X_{i-1},X_{i+1},\ldots,X_n)$ and $T(X_1,X_2,\ldots,X_{n-1})$ such that

$$\phi_i - \phi(T_{n-1}) = \phi'(T_{n-1}(i))\Delta_n(i) \tag{3.3}$$

where $\Delta_n(i) = T(X_1,X_2,\ldots,X_{i-1},X_{i+1},\ldots,X_n) - T_{n-1}$. Applying the representation (3.3) to (3.2) and using the identical distribution of the X_i's will yield

$$\operatorname{Var}\phi(T_{n-1}) \leqslant nE(\phi'(T_{n-1}(1))\Delta_n(1))^2. \tag{3.4}$$

We will now make some special choices of ϕ and proceed inductively. First consider simply $\phi(x) = x$, then (3.4) becomes

$$\operatorname{Var} T_{n-1} \leqslant nE\big(\Delta_n^2(1)\big) \leqslant 2A_2 \tag{3.5}$$

where A_2 is the constant provided by inequality (2.2) since $|\Delta_n| \leqslant d_{n-1}$. Next, choose $\phi(x) = (x - ET_{n-1})^2$ and apply Schwarz' inequality to (3.4) to obtain

$$\begin{aligned}\operatorname{Var}\big((T_{n-1} - ET_{n-1})^2\big) &\leqslant nE\big((2(T_{n-1}(1) - ET_{n-1})\Delta_n(1))^2\big)\\ &\leqslant 2^2 n\big(E(T_{n-1}(1) - ET_{n-1})^4\big)^{1/2}\big(E\Delta_n(1)^4\big)^{1/2}.\end{aligned} \tag{3.6}$$

Now since $(a + b)^4 \leqslant 16(a^4 + b^4)$ we have

$$\begin{aligned}E\big((T_{n-1}(1) - ET_{n-1})^4\big) &\leqslant 16E(T_{n-1}(1) - T_{n-1})^4 + 16E\big((T_{n-1} - ET_{n-1})^4\big)\\ &\leqslant 16E\Delta_n^4(1) + 16E(T_{n-1} - ET_{n-1})^4.\end{aligned}$$

Applying $|\Delta_n(1)| \leqslant d_{n-1}$ and (2.2) the last inequality reads

$$E\big((T_{n-1}(1) - ET_{n-1})^4\big) \leqslant 16E(T_{n-1} - ET_{n-1})^4 + 16A_4(n-1)^{-2}.$$

But since

$$\begin{aligned}\operatorname{Var}\big((T_{n-1} - ET_{n-1})^2\big) &= E(T_{n-1} - ET_{n-1})^4 - \big(E(T_{n-1} - ET_{n-1})^2\big)^2\\ &= E(T_{n-1} - ET_{n-1})^4 - (\operatorname{Var} T_{n-1})^2\end{aligned}$$

equations (3.5) and (3.6) imply

$$\begin{aligned}E(T_{n-1} - ET_{n-1})^4 \leqslant 4A_2^2 + 2^2 n\big(16E(T_{n-1} - ET_{n-1})^4\\ + 16A_4(n-1)^{-2}\big)^{1/2} A_4^{1/2}(n-1)^{-1}.\end{aligned} \tag{3.7}$$

Dividing both sides of (3.7) by $(E(T_{n-1} - ET_{n-1})^4)^{1/2}$ is good enough to show that $E(T_{n-1} - ET_{n-1})^4$ is bounded by a constant which is independent of $n \geqslant 2$. This proves the theorem for $k = 4$, and this special case has been done in such detail because the case $k = 4$ is sufficient to prove Theorem 1.

The proof of the general case will be by induction. We begin by supposing the existence of c_k and proceed to show the existence of c_{2k}. By Jensen's inequality and the induction principle, this will complete the proof. In order to avoid concerning ourselves with irrelevant constants the Vinogradov symbol $a_n \ll b_n$ will be used to indicate that there is a constant c not depending on n such that $a_n \leqslant cb_n$ for all n sufficiently large.

This time we take $\phi(x) = (x - ET_{n-1})^k$ and apply (3.4) to obtain

$$\begin{aligned}\operatorname{Var}\phi(T_{n-1}) &\ll nE\big((T_{n-1}(1) - ET_{n-1})^{k-1}\Delta_n(1)\big)^2\\ &\ll nE\big((T_{n-1} - ET_{n-1})^{k-1}\Delta_n(1)\big)^2.\end{aligned} \tag{3.8}$$

Applying Hölders inequality with $p = 2k/2(k-1)$ and $q = k$ one obtains

$$\operatorname{Var}\phi(T_{n-1}) \ll n\Big(E\big((T_{n-1} - ET_{n-1})^{2k}\big)\Big)^{2(k-1)/2k}\big(E\Delta_n(1)^{2k}\big)^{1/k}.$$

By inequality (2.2) we can check that $E\Delta_n^{2k}(1) \ll n^{-k}$ so

$$\operatorname{Var}\phi(T_{n-1}) \ll \Big(E\big((T_{n-1} - ET_{n-1})^{2k}\big)\Big)^{2(k-1)/2k}. \tag{3.9}$$

Since $\operatorname{Var}\phi(T_{n-1}) = E((T_{n-1} - ET_{n-1})^{2k}) - (E(T_{n-1} - ET_{n-1})^k)^2$ the induction assumption and (3.9) imply

$$E\big((T_{n-1} - ET_{n-1})^{2k}\big) \ll c_k^2 + \big(E(T_{n-1} - ET_{n-1})^{2k}\big)^{2(k-1)/2k}. \tag{3.10}$$

Now just as in (3.7) the presence of the small power on the right side leads to a uniform bound on $E(T_{n-1} - ET_{n-1})^{2k}$ and thus proves the existence of c_{2k}. As mentioned at the beginning, this suffices to complete the proof.

4. Further results. In the previous section the aim was to give the most direct proof possible of the result conjectured by Weide (1978). There are many natural extensions of that result, and some of those will be stated here. The proofs of these results can be given along exactly the same lines used in §3. The most obvious extension is the following.

THEOREM 3. *Suppose $\{X_i\}$ are independent and uniformly distributed in $[0,1]^d$ and $T_n = T(X_1, X_2, \ldots, X_n)$ is the length of the shortest closed path through $\{X_1, X_2, \ldots, X_n\}$, then*

$$E(T_n - ET_n)^k \leq c_k n^{(k/2)(1-d/2)} \text{ for some } c_k \text{ and all } n \geq 1 \tag{4.1}$$

and

$$\sum_{n=1}^{\infty} P\big(|T_n/n^{(d-1)/d} - \beta| \geq \epsilon\big) < \infty \quad \text{for some } \beta \text{ and all } \epsilon > 0. \tag{4.2}$$

In Theorem 3 one might also seek to replace the functional T_n by other Euclidean functionals for which the analogue of the Beardwood-Halton-Hammersley Theorem is known. The analogue of Theorem 3 does hold for the Steiner tree problem and the rectilinear Steiner tree problem. The proof in these cases differs little from that of §3. It is not known at present if the full analogue of Theorem 3 is valid for the whole class of Subadditive Euclidean Functionals (Steele (1979)).

As a final point one should note that both Beardwood-Halton, and Hammersley (1959) and Steele (1979) contain results valid for random variables with nonuniform distributions. The approximation processes used in these papers to extend the uniform case can again be tried here although to do so would require considerable space. Since the algorithmic applications initiated by Karp (1976), (1977) are already ably served by Theorem 1, these nonuniform extensions are not pursued here.

References

[1] Beardwood, J., Halton, J. H. and Hammersley, J. M. (1959). The Shortest Path through Many Points. *Proc. Cambridge Philos. Soc.* **55** 299–327.

[2] Efron, B. and Stein, C. (1978). The Jackknife Estimate of Variance. Technical Report No. 120, Department of Statistics, Stanford University.

[3] Few, L. (1955). The Shortest Path and Shortest Road through n Points. *Mathematika* **2** 141–144.

[4] Hsu, P. L. and Robbins, H. (1947). Complete Convergence and the Law of Large Numbers. *Proc. Nat. Acad. Sci. U.S.A.* **33** 25–31.

[5] Karp, R. M. (1976) . The Probabilistic Analysis of Some Combinatorial Search Algorithms. *Algorithms and Complexity: New Directions and Recent Results*. J. F. Taub, ed. Academic Press, New York, pp. 1–19.

[6] ———. (1977). Probabilistic Analysis of Partitioning Algorithms for the Traveling-Salesman Problem in the Plane. *Math Oper. Res.* **2** 209–224.

[7] Lukacs, E. (1968). *Stochastic Convergence*. D. C. Heath and Co., Lexington, Mass. 48–49.

[8] Steele, J. M. (1979). Subadditive Euclidean Functionals and Non-Linear Growth in Geometric Probability. Technical Report, Department of Statistics, Stanford University (to appear in *Ann. Probability*).

[9] Weide, B. W. (1978). *Statistical Methods in Algorithm Design and Analysis*. Thesis, Department of Computer Science, Carnegie-Mellon University.

DEPARTMENT OF STATISTICS, STANFORD UNIVERSITY, STANFORD, CALIFORNIA 94305