# On the Practical Importance of Asymptotic Optimality in Certain Heuristic Algorithms

**Harilaos N. Psaraftis**
*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

This article presents an informal discussion of the issue of asymptotic optimality of heuristics from the viewpoint of the operations research practitioner. It is suggested that certain heuristics belonging to the above class are likely to perform questionably in practice, with regard both to relative error and to computational tractability. Possible explanations of this phenomenon are offered and suggestions for further research toward a better understanding of this problem are presented.

## I. INTRODUCTION

The probabilistic analysis of heuristic algorithms is an area that has received growing attention over the past several years. Research in this area has been motivated by the observation that a heuristic's worst-case performance may have little or nothing to do with how the heuristic behaves "in practice," or "on the average." To state one example, the worst-case performance of the $k$-interchange heuristic of Lin and Kernighan [5] for the Traveling Salesman Problem (TSP) can be arbitrarily poor (as demonstrated by Papadimitriou and Steiglitz in [8]). However, this heuristic is recognized to be one of the best devised for the TSP to date, in the sense that it produces very good, near-optimal, or optimal TSP tours "for typical TSP instances most of the time." Many similar examples can be found in other problems. Given the above, many researchers have become interested in the performance of heuristics from an "average case" point of view, and the probabilistic analysis of algorithms has been adopted as the main vehicle toward that goal.

It is fair to say that the main thrust of research efforts in this area over the years ended up being the design of heuristics which are "asymptotically optimal," that is, for which it can be proven that the value of the objective produced by the heuristic converges "with probability one" (or, "almost surely"–a.s.) to the optimal value of the problem when the size of the problem becomes arbitrarily large. Heuristics belonging to this class include the partitioning heuristic for the TSP suggested by Karp [4], a similar family of heuristics for the single and multivehicle Dial-A-Ride Problem suggested by Stein [10], the "honeycomb" heuristic for the Euclidean $K$-median problem developed by Papadimitriou [7], and others. The success of those and other efforts in demonstrating asymptotic optimality have naturally stimulated development of con-

NETWORKS, Vol. 14 (1984) 587-596

CCC 0028-3045/84/040587-10$04.00

ceptually similar algorithms for other difficult combinatorial optimization problems, and the literature in this area is rapidly growing.

The scope of this paper is to reexamine informally the issue of asymptotic optimality from the viewpoint of the operations research practitioner. From that specific viewpoint, it may be of little or no importance if a heuristic for the TSP (say) produces an arbitrarily small error for sizes of problems of the order of (say) $10^7$ cities (much in the same sense that it may be irrelevant if the heuristic's worst-case error is arbitrarily high). What is likely of more interest to the operations research practitioner is whether a particular heuristic performs well at more typical problem sizes and whether its computational effort at those problem sizes is tractable. In that spirit, a practitioner would likely be very interested to know whether an asymptotically optimal heuristic is *really viable* (in terms of both relative error and computational effort) for realistic problem sizes, or whether its asymptotically optimal behavior is a feature of only theoretical (or "academic") importance.

This author feels that this important issue has hitherto received far less attention in the literature than it really deserves. It is fair to say that little theoretical or empirical work has been performed to establish *the rate of convergence* of such heuristics to their asymptotic limits *if the problem size is finite.* Some researchers (e.g., Karp [4] and Papadimitriou [7]) have presented analyses concerning the *bounds* on the relative error of their heuristics, but those analyses stop short of giving an idea of how tight those bounds can be. Karp [4] conjectures that the error bounds suggested by his analysis are too pessimistic, but the evidence that supports that statement is not really conclusive. Some other researchers may content themselves with a proof of asymptotic optimality without performing an analysis of the *magnitude* of the algorithm's relative error or its rate of convergence. In such cases, the conclusions regarding the practical merit of the algorithm can be misleading.

This article attempts to shed more light into those issues as follows: Section 2 focuses on an algorithm that has been recently developed for the routing of a fleet of vehicles to serve points on the Euclidean plane [6]. Despite the fact that the algorithm is indeed asymptotically optimal, Section 2 presents analytical evidence suggesting that the algorithm is likely to perform questionably in practice (in terms of both relative error and computational effort). Section 3 generalizes the arguments of Section 2 to other problems, presents some insights explaining the disparity between "asymptotic" and "finite" performance of such algorithms and finally suggests research directions toward a better understanding of this problem.

## 2. THE ALGORITHM OF MARCHETTI SPACCAMELA, RINNOOY KAN, AND STOUGIE

### 2.1. A Brief Description

The writing of this paper has been inspired to a great extent by a recent article by Marchetti Spaccamela, Kinnooy Kan, and Stougie [6]. The article defines a vehicle routing problem for which decisions can be broken down into two hierarchical levels: *At the aggregate level*, a decision has to be made about the number $k$ of vehicles that have to be acquired at a cost $c$ each, to serve $n$ customers from a single depot. Those customers are assumed independently and uniformly distributed within a circle of

radius $r$ with the depot at its center. The decision at the aggregate level is to be made *before* the actual location of those customers becomes known, and should be such, that the expected value of vehicle acquisition plus routing costs is minimized. *At the detailed level*, the vehicles whose number was decided upon at the aggregate level have to be routed to service $n$ customers so as to minimize the maximum route length assigned to a vehicle. Of course, decisions at the lower level are made *after* the exact locations of the $n$ customers are known.

The authors of [6] develop a two-stage heuristic algorithm for solving the above problem. *At the aggregate level*, the objective function is approximated by a deterministic equivalent $Z^{\mathrm{LB}}(k)$, which is almost surely a lower bound on that objective as follows:

$$Z^{\mathrm{LB}}(k) = ck + \frac{1}{k}\beta\sqrt{n\pi r^2} \tag{1}$$

where $\beta$ is the "asymptotic constant" for the TSP (see Beardwood et al. [1]), defined by:

$$\lim_{n\to\infty} \frac{|T^0|}{\sqrt{n\pi r^2}} = \beta, \tag{2}$$

with $|T^0|$ being the length of the optimal Traveling Salesman tour through the $n$ points.

$k$ is subsequently chosen so as to minimize $Z^{\mathrm{LB}}(k)$. It is shown that the optimal value of $k$ is

$$k^{\mathrm{LB}} = \lfloor \alpha n^{1/4} \rfloor \text{ or } \lceil \alpha n^{1/4} \rceil \tag{3}$$

with

$$\alpha = (\beta\sqrt{\pi r^2}/c)^{1/2} \tag{4}$$

Decisions *at the detailed level* are made heuristically through a *partitioning* procedure similar in spirit with (but simpler than) Karp's heuristic for the TSP [4]. The procedure essentially partitions the circle into $2^d$ subregions, containing no more than $t$ customers each, with $t$ being a parameter yet to be determined and $d$ defined by

$$d = \left\lceil \log_2 \frac{n-1}{t-1} \right\rceil \tag{5}$$

An *exact* Traveling Salesman algorithm is subsequently used to create minimum length tours for each subregion. The tours are then connected in a certain way to form a set of $k^{\mathrm{LB}}$ routes. Details on the partitioning procedure and on how the tours in each subregion are connected can be found in [6] and need not be repeated here.

The intriguing features of the procedure are twofold:

1. Its running time is polynomial in $n$ provided that $t$ is chosen to depend appropriately on $n$. In fact, for some constant $\theta > 2$ the above running time is

$$O(n \log n + n\theta^t/t + t^2 n^{1/4}) \tag{6}$$

The above function is obviously polynomial in $n$ for any fixed choice of $t$. It is also polynomial in $n$ if $t$ is taken equal to $\log n$ (as the authors of [6] have assumed). In this case, the running time becomes $O(n^2/\log n)$.

2. The heuristic is *asymptotically optimal* at both levels, in the sense that "both the error that can be ascribed to a lack of perfect information at the aggregate level and the error that results from the use of a suboptimal method at the detailed level tend to 0 as $n$ increases." According to the authors, "this is the strongest possible asymptotic optimality result that can be found for such heuristics." Mathematically, the authors show that if $t = \log n$, then (Theorem 2 of [6])

$$\lim_{n \to \infty} (\underline{Z}^H/\underline{Z}^D) = 1 \qquad \text{(a.s.)} \tag{7}$$

where $\underline{Z}^H$ is a random variable representing the total actual value of the objective, *given* $k = k^{\mathrm{LB}}$, *and, given* the partitioning and connecting strategy used at the detailed level (this is almost surely an *upper* bound on the optimal value of the problem) and $\underline{Z}^D$ is another random variable, representing the minimum cost achievable with perfect forecast into customer locations (this is almost surely a *lower* bound on the optimal value of the problem).

The authors subsequently go on to show that the heuristic is *asymptotically clairvoyant* (relation (23) of [6]) and *asymptotically optimal in expectation* (relation (24) of [6]). They also show that the solution at the aggregate level itself almost surely converges to the optimal one (Theorem 4 of [6]), and then extend their results to the case each vehicle has a different cost and speed and to some other special cases.

### 2.2. Further Analysis

The above results are definitely appealing for they imply not only that a complicated, stochastic, two-level combinatorial problem can lend itself to tractable analysis, but more important, that an asymptotically optimal procedure can be devised for it. The rest of this section pursues the approach of [6] one step further, in order to investigate the rate of convergence of the left-hand side of relation (7) above to 1.0 as $n$ increases. Indeed, in their concluding remarks, the authors of [6] give some hints that due to a term proportional to $1/\sqrt{\log n}$ that appears in several right-hand sides, such a convergence is likely to be rather slow. Here we provide a quantitative investigation of this issue, that is, give an idea of what values $n$ (as well at $t$) should take so that the left-hand side of (7) is acceptably close to 1.0 almost surely.

In order to obtain a better feeling on how the results of this investigation depend on the assumed functional relationship between $t$ and $n$, we first present the general case where no relationship between $n$ and $t$ is assumed, and then examine what happens if

$t = \log n$, as assumed in [6]. Our analysis is made assuming that the reader is familiar with the notation and arguments used in [6] to arrive at relation (7) above:

From (1), (3), and (4) we can easily see that

$$Z^{\mathrm{LB}}(k^{\mathrm{LB}}) = 2n^{1/4}\sqrt{c}\,\sqrt{\beta}\,(\pi r^2)^{1/4}. \tag{8}$$

We also have (relation (21) of [6])

$$\underline{Z}^D \geqslant Z^{\mathrm{LB}}(k^{\mathrm{LB}}). \quad \text{(a.s.)} \tag{9}$$

As far as $\underline{Z}^H$ is concerned, we have to calculate the actual value of what appears only as an *order of magnitude* in (14), (17), and (18) of [6]. To do this, we have to calculate the upper bound on $\Sigma_{j=1}^{2^d}$ per $(Y_j)$ which is implied by the authors' approach (Lemma 2 of [6]).

In fact, according to (15) and (16) of [6], we can state that

$$\begin{aligned}\sum_{j=1}^{2^d} \operatorname{per}(Y_j) &\leqslant 2^{d/2}(2r) + 2\pi r + (2^{d/2} - 1)\,4\pi r \\ &= 2^{d/2}(2r + 4\pi r) - 2\pi r.\end{aligned} \tag{10}$$

Given (5), (10) can be rewritten as

$$\sum_{j=1}^{2^d} \operatorname{per}(Y_j) \leqslant \sqrt{\frac{n-1}{t-1}}\,(2r + 4\pi r) - 2\pi r. \tag{11}$$

Given (13) and (18) of [6] and (11) above, we have

$$\underline{Z}^H \leqslant ck^{\mathrm{LB}} + \frac{1}{k^{\mathrm{LB}}}\,|T^0| + \left(\frac{3}{2}\right)\frac{\sqrt{(n-1)/(t-1)}\,(2r + 4\pi r) - 2\pi r}{k^{\mathrm{LB}}}. \quad \text{(a.s.)} \tag{12}$$

For convenience purposes, and for anything but very small values of $t$ and $n$ we can safely assume that $t - 1 \simeq t$ and $n - 1 \simeq n$. Then (2), (3), (8), (9), and (12) finally combine into

$$\frac{\underline{Z}^H}{\underline{Z}^D} \leqslant 1 + \frac{3(1 + 2\pi)}{2\beta\sqrt{\pi}}\,\frac{1}{\sqrt{t}} - \frac{\sqrt{\pi}}{2\beta}\,\frac{1}{\sqrt{n}}. \quad \text{(a.s.)} \tag{13}$$

It is interesting to note that the right-hand side is independent of $c$ and $r$. Assuming that $\beta = 0.765$ (the speculated, approximate value of $\beta$) the above can be finally written as

$$\underline{Z}^H/\underline{Z}^D \leqslant 1 + 8.061/\sqrt{t} - 1.158/\sqrt{n}, \quad \text{(a.s.)} \tag{14}$$

which, in fact, reduces to (7) for $n \to \infty$ and $t = \log n$ (since $\underline{Z}^H \geqslant \underline{Z}^D$ a.s.).

It can be seen that the rate at which the right-hand side of (14) converges to 1.0 depends on the rate at which the relative error term $(8.061/\sqrt{t} - 1.158/\sqrt{n})$ goes to zero as $n \to \infty$. Since the term $8.061/\sqrt{t}$ is present, the rate of convergence depends critically on the assumed functional relationship between $t$ and $n$. It runs out that for $t = \log n$ (as was assumed in [6]) the convergence is extremely slow, as indicated by the following table:

| $n$ | $t = \log n$ | $(8.061/\sqrt{t} - 1.158/\sqrt{n})$ |
|---|---|---|
| $10^3$ | 10 | $2.55 - 0.04 = 2.51$ |
| $10^6$ | 20 | $1.80 - 1.158 \times 10^{-3} = 1.80$ |
| $10^{12}$ | 40 | $1.27 - 1.158 \times 10^{-6} = 1.27$ |
| $10^{24}$ | 80 | $0.90 - 1.158 \times 10^{-12} = 0.90$ |

(A minor observation is that in all cases the error contribution of the term $-1.158/\sqrt{n}$, although favorable, is several orders of magnitude smaller than the error contribution of the term $8.061/\sqrt{t}$, which is of course unfavorable.)

If now one is to keep the error term acceptably small (for instance, so as to guarantee *almost surely* a relative error of no more than a modest 10%) then $t$, and, a fortiori, $n$ must take on intolerably large values (in this case $t \simeq 6.500$ and $n \simeq 10^{1957}$, respectively). Of course, it is impossible that distribution problems of such size would ever occur in practice (it suffices to realize that the total number of atoms in the universe is of the order of $10^{80}$). Furthermore, such high values of $t$ would virtually prohibit the use of an *exact* Traveling Salesman algorithm for the routing in each subregion; in fact, one cannot use such exact algorithms if $t$ is more than about 100 (at best). This intractability is reflected in the procedure's running time, which, albeit $O(n^2/\log n)$, is still an exponential function of $t$ (according to (6)).

Thus, this analysis "quantitatively" confirms the "qualitative" pessimism of the authors of [6] regarding the convergence of their procedure. The practical implication of the analysis is that the error that can be ascribed to this heuristic *cannot* be guaranteed—not even in a "probabilistic" (or "almost sure") sense—to go to zero as $n$ increases, *without* increasing the overall computational effort of the procedure to a prohibitive level. Conversely, in order to maintain computational tractability in the heuristic, and for the range of problem sizes that are likely to occur in the real world, one has to accept bounds on its relative error that are too loose to serve as practical performance guarantees, however liberally the word "guarantee" is interpreted.

Of course, as the authors of [6] mention, one can try to use a *heuristic* algorithm instead of an exact one for the TSP problems that are defined in each subregion. Such substitution would alleviate (but not eliminate) the computational burden of the procedure. However, this would also nullify any claims on asymptotic optimality. Golden et al. [3], investigated such "hybrid" procedures in conjunction with Karp's TSP approach [4] and reported "reasonable" results but *for very large problem sizes* (an evidence of the slow rate of convergence of that procedure). Whether such an observation is also applicable to this algorithm is an open question at this point.

One may also argue that the bounds obtained by the above analysis are too loose and that *tighter bounds* could be obtained by a more careful investigation. However true this might be, it would be impossible to bring the term $\Sigma_{j=1}^{2^d}$ per $(Y_j)$ to a value lower

than $2^{d/2}(2r) + 2\pi r$ (Lemma 2 of [6]). And even if the second contribution $(2^{d/2} - 1)\,4\pi r$ is completely ignored in (10) (a very optimistic assumption), its only effect would be to substitute (14) with

$$\underline{Z}^H/\underline{Z}^D \leqslant 1 + 1.106/\sqrt{t} + 1.158/\sqrt{n}. \quad \text{(a.s.)} \tag{15}$$

Of course, (15) is better than (14), but still suffers from the same problems (a 10% error requires $t \simeq 120$, or $n \simeq 10^{36}$, still bad).

Finally, one might argue that better results could be obtained if the functional relationship between $t$ and $n$ were different. For instance, if $t$ grew *faster* than $\log n$, a faster convergence could be achieved. However, a faster growth is a two-edged sword, for it automatically increases the computational effort of the procedure (according to (6)). In summary, the price one has to pay (in terms of computational effort) in order to limit the algorithm's error to an acceptable level seems excessively high.

## 3. DISCUSSION

In their concluding remarks the authors of [6] suggest that the probabilistic analysis of algorithms serves not so much to cover the user against every possible situation, but rather to *explain* why certain simple heuristics almost always seem to produce much better results than their worst-case analysis would seem to suggest. We agree with this basic premise. At the same time, we feel that one should be cautious in interpreting an algorithm's asymptotic optimality as *conclusive evidence* that the performance of this algorithm in practice will be good, very good or near-optimal. In fact, we feel that this paper has demonstrated that such conclusions cannot be drawn for the algorithm of [6]. Under the current state of knowledge in the probabilistic analysis of algorithms, we feel that extensive computational experience with the procedure is about the only way to shed more light on how good it is in practice. Until such experience is obtained, or until the state of the art in analytical methods in this area is improved, this issue is likely to remain open.

So far, this investigation has been specific to the approach of [6]. To what extent can one generalize the above arguments to other similar approaches that have been presented in the literature to date? And what can be offered as an explanation for such potential disparities between, on the one hand, asymptotic optimality, and, on the other, questionable practical performance of such algorithms? The rest of this section attempts to answer those questions.

Obviously, one cannot simply use the above example to draw conclusions about other heuristics, because it may be that such behavior is specific only to the heuristic of [6] and absent elsewhere. However, there is *some* evidence that this behavior (or symptoms of it) is present in some other similar heuristic procedures: In addition to the *partial* and *empirical* evidence that Karp's partitioning TSP heuristic [4] converges slowly (as reported in [3]), this author provided *analytical* evidence that the same is true of Stein's asymptotically optimal heuristic for the single vehicle Euclidean Dial-A-Ride Problem (DARP) [10]. The full analysis is presented in section 3.2 of [9], but it might be useful to summarize its main rationale below:

The single vehicle Euclidean DARP calls for the routing of a vehicle to service $n$ customers, each of whom wishes to travel from a distinct origin to a distinct destination.

Assume all origins and destinations are independently and uniformly distributed in a Euclidean service region. Stein's asymptotically optimal heuristic divides the service region into $m$ regions of equal size. On the "first pass" through the regions the vehicle visits in region $i$ ($i = 1, 2, \ldots, m$) all origins in that region as well as all destinations whose origins are in regions $1, 2, \ldots, i-1$, with the routing in each region done *optimally* (i.e., using an exact TSP algorithm). On the "second pass" through the regions the vehicle visits the remaining destinations in each region (again optimally). Subtours are then connected arbitrarily.

Stein's heuristic is asymptotically optimal if $n \to \infty$, $m \to \infty$, *and* $n/m \to \infty$ (the latter so that one can use relation (2) for the lengths of the TSP tours in each region). If an exact, Dynamic Programming algorithm is used for the TSP's of each region, the overall complexity of the algorithm becomes $O(n^2 2^{n/m}/m)$, which can be *polynomial* with respect to $n$ for certain rates of growth of $m$ with respect to $n$. One such type of growth is $m = n/(h \log n)$ with $h > 0$ being a constant. If this is the case, the computational complexity of the algorithm becomes $O(n^{h+1} \log n)$, polynomial with respect to $n$. From a running time point of view, one would like to choose $h$ to be small. If $h = 1$ however, and if $n/m \simeq 20$ customers per region is the minimum that is adequate for using the approximation implied by (2) (see also Eilon et al. [2, p. 170]), this would imply $m = 52{,}500$ regions and $n = 1{,}000{,}000+$ customers, a problem size very unlikely to be encountered in a single-vehicle DARP in the real world. A smaller value for $h$ would increase $n$ even more. All this leads to the conjecture that if one is to maintain computational tractability in this heuristic, then one should be ready to accept a rate of convergence which is unlikely to be of any practical value. Indeed, further analysis in [9] showed that a simple $O(n^2)$ DARP heuristic devised by this author, although by no means asymptotically optimal, matched or even exceeded the performance of Stein's procedure for finite problem sizes.

Similar behaviors may be observed in other asymptotically optimal heuristics, not necessarily associated with routing problems. Indeed, recent analysis by the author and his colleagues have revealed similar symptoms in a known heuristic for the planar $K$-median problem. This investigation is still ongoing and will not be reported here. The author is also aware of sparse similar observations by other researchers with respect to other asymptotically optimal algorithms.

How can one explain such types of behavior? One can speculate on one plausible explanation at this point: It is generally recognized that the probabilistic analysis of heuristic algorithms lends itself only to heuristics whose structure is very simple, even naive, with the words "simple" and "naive" referring not to the ingenuity in the design of those algorithms, but to the coupling and interdependence of the algorithms' stages. In those terms, any "nontrivial" heuristic generally has the property of introducing strong dependencies between its stages, thus rendering its probabilistic analysis intractable (this is the main reason such analysis has not been attempted thus far for other TSP heuristics, such as that of Lin and Kernighan [5]).

Unfortunately, while the structural simplicity of such heuristic algorithms is a definite advantage in facilitating their probabilistic analysis and can be usually exploited to prove asymptotic optimality, it is exactly this same feature that is likely to cause such undesirable characteristics as slowness of convergence, looseness of error bounds and computational intractability in more typical problem sizes. There are in fact two steps

in the algorithm of [6] where such structural simplicity is manifested: The first is the suppression of the combinatorial structure of the problem at the aggregate level where the stochastic objective function is replaced by a deterministic equivalent. The second is the partitioning scheme at the detailed level, which, as the authors of [6] quite correctly emphasize, is simpler than the one proposed by Karp in [4]. Both structural simplifications are ingenious because they ultimately imply asymptotic optimality for both levels of the algorithm. However, this author conjectures that the same two features ultimately work *against* the performance and tractability of this algorithm for typical problem instances and finite problem sizes.

There seems to be no question among operations research theoreticians *and* practitioners that there is often significant disparity between a heuristic algorithm's worst-case performance and its performance in practice, and that such a disparity can lead to misleading conclusions regarding the practical merits of the algorithm in question. We believe that equally significant is often the gap between what can be proven probabilistically (or asymptotically) and what can be observed in practice, and that equally misleading conclusions can be drawn. For all the insights it may provide to intuitively explaining why a "poor" heuristic from a worst-case viewpoint may in fact perform decently in practice, probabilistic analysis may often raise more questions regarding the real-world viability of certain algorithms than it can answer.

Given the above, there is clearly a need for more research in this area. Specifically, with few exceptions (and these typically refer to easier, well-understood combinatorial problems), most analyses to date have examined the issue of asymptotic optimality only from a first-order probabilistic viewpoint. An extension of such analyses into a study of *second-order* effects (such as the *variance* of the optimal value as a function of problem size) would definitely provide a better understanding of heuristics for hard combinatorial problems. From a practitioner's viewpoint, an algorithm's *robustness* (intimately linked to the above variance) is likely to be much more important than that algorithm's ability to produce arbitrarily small errors for arbitrarily large problem sizes.

Ultimately, developing *tight* error bounds as functions of problem size, as well as precise expressions for the probabilities that such tight bounds are valid seems to be a worthwhile goal. Whether such an ambitious goal would bring about the development of a new generation of analytical methods in this area remains to be seen.

I wish to thank the Editor and the referees for their comments. During the last 3 to 4 years I have also benefited from various inputs I had on this general subject from Marshall Fisher, Bruce Golden, Tai-Up Kim, Amedeo Odoni, Jim Orlin, Christos Papadimitriou, Alexander Rinnooy Kan, and David Stein.

## References

[1] J. Beardwood, S. H. Halton, and J. M. Hammersley, The shortest path through many points. *Proc. Cambridge Phil. Soc.* 55 (1959) 229–327.

[2] S. Eilon, N. Christofides, and C. D. T. Watson-Gandy, *Distribution Management: Mathematical Modeling and Practical Analysis.* Hafner, New York (1971).

[3] B. Golden, L. D. Bodin, T. Doyle, and W. Stewart, Jr., Approximate traveling salesman algorithms. *Operations Res.* 28 (1980) 694–711.

[4] R. M. Karp, Probabilistic analysis of partitioning algorithms for the traveling salesman problem. *Math. Operations Res.* **2** (1977) 209–224.
[5] S. Lin and B. W. Kernighan, An effective heuristic algorithm for the traveling salesman problem. *Operations Res.* **21** (1973) 498–516.
[6] A. Marchetti Spaccamela, A. H. G. Rinnooy Kan, and L. Stougie, Hierarchical vehicle routing problems. *Networks* **14** (1984) 571–586.
[7] C. H. Papadimitriou, Worst-case and probabilistic analysis of a geometric location problem. *SIAM J. Comput.* **10** (1981) 542–557.
[8] C. H. Papadimitriou and K. Steiglitz, Some examples of difficult traveling salesman problems. *Operations Res.* **26** (1978) 434–443.
[9] H. N. Psaraftis, Analysis of an $O(N^2)$ heuristic for the single vehicle many-to-many euclidean dial-a-ride problem. *Transportation Res.* **17B** (1983) 133–145.
[10] D. M. Stein, An asymptotic, probabilistic analysis of a routing problem. *Math. Operations Res.* **3** (1978) 89–101.