A PATCHING ALGORITHM FOR THE NONSYMMETRIC TRAVELING-SALESMAN PROBLEM

by

Richard M. Karp

Memorandum No. UCB/ERL M78/2

4 January 1978

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

A PATCHING ALGORITHM FOR THE NONSYMMETRIC TRAVELING-SALESMAN PROBLEM†

Richard M. Karp

Computer Science Division
Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California at Berkeley

Abstract

We present an algorithm for the approximate solution of the nonsymmetric n-city traveling-salesman problem. An instance of this problem is specified by a $n \times n$ distance matrix $D = (d_{ij})$. The algorithm first solves the assignment problem for the matrix D, and then patches the cycles of the optimum assignment together to form a tour. The execution time of the algorithm is comparable to the time required to solve a $n \times n$ assignment problem.

If the distances d_{ij} are drawn independently from a uniform distribution then, with probability tending to 1, the ratio of the cost of the tour produced by the algorithm to the cost of an optimum tour is $< 1 + \epsilon(n)$, where $\epsilon(n)$ goes to zero as $n \to \infty$. Hence the method tends to give nearly optimal solutions when the number of cities is extremely large.

Key words: traveling-salesman problem, combinatorial optimization, approximation algorithms, probabilistic analysis of algorithms.

†Research supported by National Science Foundation Grant MCS74-017680-A02.

1. Introduction

Let Σ_n denote the set of all permutations of $\{1,2,\ldots,n\}$ and let Σ_n^* denote the set of all cyclic permutations of $\{1,2,\ldots,n\}$. For any $n \times n$ matrix $D = (d_{ij})$ of nonnegative real numbers and any permutation $\pi \in \Sigma_n$, define $c(\pi,D) = \sum_{i=1}^{n} d_{i,\pi(i)}$.

The (nonsymmetric) traveling-salesman problem is stated as follows: given D, find a cyclic permutation $\pi^*(D)$ (or simply π^*, when D is understood) such that $c(\pi^*,D) = \min_{\pi \in \Sigma_n^*} c(\pi,D)$. This problem typically arises in machine scheduling applications, where d_{ij} represents the set-up cost for job j upon the completion of job i, and an optimum sequence of job execution is desired. Since the directed traveling-salesman problem is NP-hard [4], it is not reasonable to expect to find a polynomial-time algorithm for its exact solution. Well-designed branch-and-bound methods are capable of efficiently solving problem instances of size up to about $n = 100$ [6].

By an approximation algorithm for the traveling-salesman problem we mean an algorithm A that, given any matrix D, produces a cyclic permutation $\hat{\pi}(D)$. The relative error associated with the execution of A on D is

$$e(D) = \frac{c(\hat{\pi}(D),D) - c(\pi^*(D),D)}{c(\pi^*(D),D)} .$$

Sahni and Gonzales have shown that, given any $\varepsilon > 0$, it is *NP*-hard to solve the traveling-salesman problem with relative error $< \epsilon$. Thus, we cannot expect to find a polynomial-time approximation algorithm with uniformly bounded relative error.

In this paper we present a polynomial-time approximation algorithm which tends to gives solutions with small relative error. The algorithm starts by solving the $n \times n$ assignment problem, which is stated as follows: given D, find a permutation $\bar{\pi}(D)$ (or simply $\bar{\pi}$) such that

$$c(\bar{\pi},D) = \min_{\pi \in \Sigma_n} c(\pi,D).$$

There are algorithms which solve the assignment problem in time $O(n^3)$ ([1], [7]). Our approximation algorithm produces a cyclic permutation $\hat{\pi}$ by patching together the cycles of the optimal assignment permutation $\bar{\pi}(D)$. The running time of the algorithm is $O(n^3)$. The algorithm also yields an upper bound on the relative error $\epsilon(D)$. Our main theorem states that, if the d_{ij} are drawn independently from the uniform distribution in $[0,1]$, then with probability tending to 1 as $n \to \infty$, this upper bound is very small.

A companion paper to the present one [5] gives similar results for the traveling-salesman problem in the plane.

It is interesting that a patching algorithm similar to ours has been proven to give strictly optimum solutions for an important special class of traveling-salesman problems [3].

2. The Patching Algorithm

We begin by stating the $m \times n$ assignment problem. Let m and n be positive integers with $m \leq n$. Let $S_{m,n}$ denote the set of single-valued one-one functions from $\{1,2,\ldots,m\}$ into $\{1,2,\ldots,n\}$. In particular, when $m = n$, $S_{m,n} = \Sigma_n$, the set of permutations of $\{1,2,\ldots,n\}$. Given a $m \times n$ matrix $A = (a_{ij})$ of real numbers, the assignment problem asks for a function $\bar{\pi} \in S_{m,n}$ such that

$$\sum_{i=1}^{m} a_{i,\bar{\pi}(i)} = \min_{\pi \in S_{m,n}} \sum a_{i,\pi(i)} .$$

There are algorithms to solve the $m \times n$ assignment problem in $O(m^2n)$ steps ([1],[7]).

Given a $n \times n$ matrix D, the patching algorithm begins by finding an optimum assignment $\bar{\pi}$. If, fortuitously, $\bar{\pi}$ is a cyclic permutation, then the traveling-salesman problem is solved. Otherwise, $\bar{\pi}$ will have two or more cycles. The algorithm patches these cycles together into a single cycle, thereby obtaining a cyclic permutation.

We next describe how the patching is done. Let $\rho_{ij} \in \Sigma_n$ be the permutation that interchanges elements i and j, leaving all other elements fixed. The transformation

$$R_{ij}: \Sigma_n \rightarrow \Sigma_n ,$$

defined by $R_{ij}(\pi) = \pi \circ \rho_{ij}$, is called the <u>i,j patching operation</u>. <u>Also</u>, define

$$\Delta_{ij}(\pi,D) = d_{i,\pi(j)} + d_{j,\pi(i)} - d_{i,\pi(i)} - d_{j,\pi(j)} \cdot$$

The following lemma is immediate.

<u>Lemma 1</u>. For all i, j,

$$c(\pi \circ \rho_{ij},D) = c(\pi,D) + \Delta_{ij}(\pi,D) .$$

Also, if i and j are in different cycles of π, then the elements in these two cycles lie in a single cycle of $\pi \circ \rho_{ij}$, and the other cycles of π remain unchanged.

Figure 1 indicates the effect of the i,j patching operation.

We next describe a <u>patching process</u> which, given a permutation α with k cycles, attempts to transform α to a cyclic permutation by applying a sequence of $k-1$ patching operations. This sequence is selected as follows. First, some cycle $\bar{C}$ of maximum length in α is selected. Let the remaining cycles be $C_1, C_2, \ldots, C_{k-1}$. Ambiguously, we let the name of a cycle also stand for the set of elements of the cycle. If $|\bar{C}| < k-1$, then the algorithm reports failure and halts. (We shall see that this event has negligible probability.) Otherwise, a $(k-1) \times |\bar{C}|$ assignment problem is set up whose solution gives an optimum way to patch all the cycles $C_1, C_2, \ldots, C_{k-1}$ into $\bar{C}$ at distinct places. The matrix A defining this problem has $k-1$ rows and a column for each $j \in \bar{C}$. The i-j entry is

$$(*) \qquad a_{ij} = \min_{\ell \in C_i} \Delta_{\ell j}(\alpha, D) .$$

Thus, a_{ij} is the least cost of a patching operation involving element $j \in \bar{C}$, and any element in C_i. Let the minimizing ℓ in (*) be denoted $\ell(i,j)$. Let the solution to this assignment problem be a 1-1 function $\theta: \{1,2,\ldots,k-1\} \to \bar{C}$. Then, for $i = 1,2,\ldots,k-1$, the patching process performs the patching operation $R_{\ell(i,\theta(i)),\theta(i)}$. These operations commute, and may thus be performed in any order.

Figure 2 indicates how a permutation α with four cycles might be converted into a tour. Let $\delta(\alpha,D,\bar{C}) = \sum_{i=1}^{k-1} \min_{\ell \in C_i} \Delta_{\ell,\theta(i)}(\alpha,D)$. Then the cyclic permutation obtained by applying the patching process to α has cost $C(\alpha,D) + \delta(\alpha,D,\bar{C})$. The time required for the patching process is $O((k-1)^2|\bar{C}|) = O(n^3)$.

The over-all patching algorithm is now easily stated:

i) find an optimal assignment $\bar{\pi}$ for the matrix D;

ii) if $\bar{\pi}$ has k cycles and no cycle is of length $\geq k$, then halt and report failure;

iii) otherwise, apply the patching process to obtain a cyclic permutation $\hat{\pi}(D)$ of cost $c(\bar{\pi},D)+\delta(\bar{\pi},D,\bar{C})$;

iv) print out the permutation $\hat{\pi}$ and the error bound

$$\epsilon(D) \le \frac{\delta(\bar{\pi},D,\bar{C})}{c(\bar{\pi},D)} .$$

The execution time of the patching algorithm is $O(n^3)$.

Section 3 analyzes the distribution of $\epsilon(D)$. Section 4 gives a heuristic error analysis for a variant of the patching algorithm.

3. Relative Error of the Patching Algorithm

Recall that, for any matrix D, $\hat{\pi}(D)$ denotes the cyclic permutation produced by the patching algorithm, and $\pi^*(D)$ denotes an optimal solution to the traveling-salesman problem for D. Thus the relative error of the patching algorithm is given by

$$\epsilon(D) = \frac{c(\hat{\pi},D) - c(\pi^*,D)}{c(\pi^*,D)} .$$

Let $U_{n\times n}$ be the uniform distribution over the set of $n\times n$ matrices whose elements lie in $[0,1]$.

Theorem 1. Let D be drawn from $U_{n\times n}$. Then with probability tending to 1 as $n\to\infty$,

(1) $$\epsilon(D) < 9\sqrt{8+2\sqrt{7}}(\ln n)^2\cdot n^{-1/2} .$$

The present section is devoted to the proof of Theorem 1. We begin with some preliminary remarks and propositions needed for the proof.

Drawing a matrix D from $U_{n\times n}$ is equivalent to drawing each element independently from the uniform distribution on [0,1]. With probability 1, a D drawn from $U_{n\times n}$ has the property that no two sets of its elements have the same sum. We assume that all distance matrices considered have this property. Thus, in particular, we assume that every matrix D presented to the algorithm has a unique optimum assignment.

The first proposition concerns the tails of the binomial distribution. It is a direct consequence of the Chernoff bound [2].

Proposition 1. For $0 \le p \le 1$ and N a positive integer

a) for all $\beta \in [0,1]$

$$\sum_{k=0}^{[(1-\beta)Np]} \binom{N}{k} p^k (1-p)^{N-k} < \exp\left(-\frac{\beta^2 Np}{2}\right) \tag{2}$$

b) for all $\beta \in [0,\infty)$,

$$\sum_{k=(1+\beta)Np}^{\infty} \binom{N}{k} p^k (1-p)^{N-k} < \exp\left(-\frac{\beta^2 Np}{3}\right) . \tag{3}$$

The second proposition concerns random permutations. Ambiguously, let Σ_n denote both the set of all permutations of $\{1,2,\dots,n\}$, and the uniform distribution over this set. The symbol $m(\alpha)$ denotes the maximum length of any cycle of the permutation α.

Proposition 2. Let α be drawn from Σ_n. Then, with probability tending to 1 as $n \to \infty$,

a) α has at most $3 \ln n$ cycles

b) $\frac{n}{3 \ln n} \le m(\alpha) \le n - n^{2/3}$

and c) α has exactly one cycle of length $m(\alpha)$.

Proof. Let α be a permutation in Σ_n. By the cycle structure of π is meant that partition of n in which an integer ℓ appears as many times as π has cycles of length ℓ. The uniform distribution over Σ_n induces a probability distribution P_n over all the possible cycle structures.

Consider the following method of selecting a random partition of n.

```
PROCEDURE PARTITION(n)
    begin;
    i ← n;
    while i > 0 do begin;
                1. select a random integer k from the uniform distribution
                   over {1,2,...,i};
                   i ← i-k
                   end
    end
```

The multiset of integers k selected during the process forms the desired partition.

We claim that executing PARTITION(n) is equivalent to sampling from the distribution P_n. This follows from two observations. First, for each k between 1 and n, the number of permutations in Σ_n such that some fixed element x lies in a cycle of length k is $(n-1)!$. Therefore, the length of the cycle containing x is uniformly distributed over $\{1,2,\ldots,n\}$. Secondly, given that x lies in a cycle of length k, all permutations of the remaining $n-k$ elements are equally likely.

We prove (a) by considering the execution of PARTITION(n). Call an execution of step 1 a success if $\lceil \log_2 i \rceil > \lceil \log_2 (i-k) \rceil$. At each step the probability of success is $\geq \frac{1}{2}$. Also, the process terminates no later than the $\lceil \log_2 n \rceil$-th success. Hence, the probability that α has more than $3 \ln n$ cycles is less than or equal to the probability that $3 \ln n$ flips of a fair coin will result in fewer than $\lceil \log_2 n \rceil$ heads; and, by Proposition 1, this probability is $o(1)$.

To prove (b), note that

$$\Pr\{m(\alpha) < \frac{n}{3\ln n}\} < \Pr\{\alpha \text{ has at least } 3\ln n \text{ cycles}\} = o(1)$$

and

$$\Pr\{m(\alpha) > n-n^{2/3}\} \le \Pr\{\text{element 1 lies in a cycle of length} < n^{2/3} \text{ or } > n-n^{2/3}\}$$
$$= O(n^{-1/3})$$

To prove (c), note that

$$\Pr\{\alpha \text{ contains two cycles of length } k\}$$
$$\le \frac{n}{2k} \cdot \left(\Pr\{\alpha \text{ contains two cycles of length } k, \text{ and element } x \text{ is in one of them}\}\right)$$
$$\le \frac{n}{2k} \cdot \frac{1}{n} \cdot \frac{1}{n-k} = \frac{1}{2k(n-k)} .$$

Hence,

$$\Pr\{\text{for some } k \ge \frac{n}{3\ln n}, \alpha \text{ contains two cycles of length } k\}$$
$$\le \sum_{k=\frac{n}{\ln n}}^{n/2} \frac{1}{2k(n-k)} = \frac{1}{2n} \sum_{k=\frac{n}{\ln n}}^{n/2} \frac{1}{k} + \frac{1}{n-k} = O\left(\frac{\ln n}{n}\right) .$$

The third proposition concerns properties of permutations, given an upper bound on the lengths of their cycles.

<u>Proposition 3</u>. Let m and n be integers. Let Σ_n^m denote $\{\alpha | \alpha \in \Sigma_n$ and $m(\alpha) \le m\}$. Let A_k^n denote the expected number of elements occurring in cycles of length k in a permutation drawn at random from Σ_n^m. Then

a) $\sum_{k=1}^{m} A_k^n = n$ and

b) $A_1^n \le A_2^n \le \cdots \le A_m^n$.

Proof. Part (a) is immediate, since every element is in exactly one cycle. To prove (b) let $F_n = |\Sigma_n^m|$. Then

$$F_\ell = \sum_{k=1}^{\min(m,n)} (\ell-1)(\ell-2)\cdots(\ell-k+1)F_{\ell-k} \; . \tag{4}$$

This follows from the observation that $(\ell-1)(\ell-2)\cdots(\ell-k+1)F_{\ell-k}$ gives the number of permutations in Σ_ℓ^m such that a fixed element x lies in a cycle of length k. From (4) it follows by induction on ℓ that

$$\text{for all } \ell, \quad F_\ell \leq \ell F_{\ell-1} \; . \tag{5}$$

Also,

$$A_k^n = \frac{n(n-1)\cdots(n-k+1)F_{n-k}}{F_n} \; , \tag{6}$$

since A_k^n is n times the probability that element x lies in a cycle of length k. Hence

$$\frac{A_k^n}{A_{k-1}^n} = \frac{(n-k+1)F_{n-k}}{F_{n-k+1}} \geq 1 \; .$$

□

The fourth proposition concerns matrices drawn from $U_{n\times n}$.

Proposition 4. Let D be drawn from $U_{n\times n}$. Let $\bar{\pi}$ be the optimal assignment for D. Then, with probability tending to 1 as $n \to \infty$, $\frac{1}{3} < c(\bar{\pi},D) < 3$.

Proof. To prove the lower bound on $c(\bar{\pi},D)$, note that

$$\Pr\{c(\bar{\pi},D) > \tfrac{1}{3}\} \geq \Pr\{\sum_i \min_j d_{ij} > \tfrac{1}{3}\} \geq \Pr\{|\{i \mid \min_j d_{ij} > \tfrac{2}{3n}\}| \geq \tfrac{n}{2}\} \; .$$

But, for any fixed i,

$$\Pr\{\min_j d_{ij} > \frac{2}{3n}\} = (1 - \frac{2}{3n})^n \xrightarrow[n\to\infty]{} e^{-2/3} .$$

Applying Proposition 1a (with $N = n$, $\frac{1}{2} < p < e^{-2/3}$, $(1-\beta) = \frac{1}{2p}$), it follows that $\Pr\{|\{i|\min_j d_{ij} > \frac{2}{3n}\}| \geq \frac{n}{2}\}$ tends to 1.

The upper bound on $c(\bar{\pi},D)$ is an unpublished result due to David Walkup [8]. □

Now we embark on the proof of Theorem 1. Let $D = (d_{ij})$ be drawn from $U_{n\times n}$. Call D <u>exceptional</u> if any of the following are violated:

i) $\bar{\pi}(D)$ has at most $3 \ln n$ cycles;

ii) $\frac{n}{3 \ln n} \leq m(\bar{\pi}(D)) \leq n - n^{2/3}$;

iii) $\bar{\pi}(D)$ has a unique longest cycle;

iv) $\frac{1}{3} < c(\bar{\pi},D) < 3$.

By the above propositions, the probability that D is exceptional tends to 0 as $n \to \infty$. For any $\sigma \in \Sigma_n$, define the matrix D^σ by $(D^\sigma)_{ij} = d_{i,\sigma(j)}$. Thus D^σ is obtained by permuting the columns of D. Let $[D]$ denote the set $\{D^\sigma|\sigma\in\Sigma_n\}$. The following lemma is the basis of our proof.

<u>Lemma 2</u>. For any permutation α, $c(\alpha,D) = c(\sigma^{-1}\alpha,D^\sigma)$. Hence, $\bar{\pi}$ is an optimal assignment for D if and only if $\sigma^{-1}\bar{\pi}$ is an optimal assignment for D^σ. Also, $\Delta_{ij}(\alpha,D) = \Delta_{ij}(\sigma^{-1}\alpha,D^\sigma)$.

Given any set $T \subseteq \{1,2,\ldots,n\}$, let Σ_n^T denote the set of all permutations α in Σ_n such that

a) T is the set of elements in a cycle of α

and b) the cycle containing these elements is a longest cycle of α; i.e., $m(\alpha) = |T|$.

Let $\mathcal{S} = \{T\subseteq\{1,2,\ldots,n\}|\frac{n}{3 \ln n}\leq |T| \leq n-n^{2/3}\}$. For any $T \in \mathcal{S}$ let $[D,T]$ denote the set of matrices in $[D]$ whose optimum assignment is in Σ_n^T.

Note that, unless D^σ is exceptional, it lies in exactly one set $[D,T]$.

In the next four lemmas, let S be a fixed set in $\mathcal{S}$. For any permutation α, we construct a patching matrix $\Delta^\alpha(D,S)$ of dimension $(n-|S|) \times |S|$. The rows of this matrix correspond to the elements of $\{1,2,\ldots,n\}$ not in S, and the columns, to the elements of S. The i,j entry gives the patching cost $\Delta_{ij}(\alpha,D)$, which, by Lemma 2, is equal to $\Delta_{ij}(\sigma^{-1}\alpha,D^\sigma)$.

A bad element of $\Delta^\alpha(D,S)$ is one which is $> \sqrt{\frac{(8+2\sqrt{7})\ln n}{m}}$. An element which is not bad is a good element. A bad row of $\Delta^\alpha(D,S)$ is one that contains fewer than $3\ln n$ good elements. The matrix $\Delta^\alpha(D,S)$ is a bad matrix if it contains more than $\sqrt{3n\ln 4}+1$ bad rows; otherwise $\Delta^\alpha(D,S)$ is a good matrix.

Lemma 3. If D is drawn at random from $U_{n\times n}$ and α is drawn at random from Σ_n, then $\Pr\{\Delta^\alpha(D,S) \text{ is bad}\} < 4^{-n}$.

Proof. Define a matrix $\Omega^\alpha(D,S)$ (or, briefly, Ω^α) with the same rows and columns as $\Delta^\alpha(D,S)$, such that $(\Omega^\alpha)_{ij} = d_{i,\alpha(j)} + d_{j,\alpha(i)}$. Then Ω^α is element-by-element greater than or equal to $\Delta^\alpha(D,S)$, and it remains only to prove that $\Pr\{\Omega^\alpha \text{ is bad}\} \le 4^{-n}$. The elements of Ω^α are independent, and each is the sum of two independent samples from the uniform distribution on $[0,1]$. Thus, independently for each pair i,j, $\operatorname{pr}\{\Omega^\alpha_{ij} \text{ is good}\} \ge (4+\sqrt{7})\frac{\ln(n-m)}{m}$. Thus, applying (2) with $N = m$, $p = \frac{(4+\sqrt{7})\ln(n-m)}{m}$ and $\beta = \frac{1+\sqrt{7}}{4+\sqrt{7}}$, we obtain:

$$\Pr\{\text{row } i \text{ has} \le 3\ln n \text{ good elements}\} < \exp\left(\frac{(1+\sqrt{7})^2}{(4+\sqrt{7})^2}\cdot\frac{(4+\sqrt{7})}{2}\right)\ln(n-m)$$
$$= e^{-\ln(n-m)} = \frac{1}{n-m}\ .$$

Thus each row has probability $< \frac{1}{n-m}$ of being bad. Then, the probability that there are more than $1+\sqrt{3n \ln 4}$ bad rows is bounded above by substituting $N = n-m$, $p = \frac{1}{n-m}$, $\beta = \sqrt{3n \ln 4}$ in (3). The resulting upper bound is 4^{-n}. □

Lemma 4. Let D be drawn from $U_{n\times n}$, let D^{σ} be drawn at random from $[D,S]$, and let $\bar{\pi} = \bar{\pi}(D^{\sigma})$. Then $\Pr\{D^{\sigma}$ is not exceptional and $\Delta^{\bar{\pi}}(D^{\sigma},S)$ is bad$\} = O((\frac{3}{4})^n)$.

Proof. $\Pr\{D^{\sigma}$ is not exceptional and $\Delta^{\bar{\pi}}(D^{\sigma},S)$ is bad$\}$

$$\le \Pr\{c(\bar{\pi},D^{\sigma}) < 3 \text{ and } \Delta^{\bar{\pi}}(D^{\sigma},S) \text{ is bad}\}$$

$$\le \Pr\{\exists\alpha \mid c(\alpha,D^{\sigma}) < 3 \text{ and } \Delta^{\alpha}(D^{\sigma},S) \text{ is bad}\}$$

$$\le E\ |\{\alpha \mid c(\alpha,D^{\sigma}) < 3 \text{ and } \Delta^{\alpha}(D^{\sigma},S) \text{ is bad}\}|$$

$$(*) \qquad = n!\Pr\{c(\alpha,D^{\sigma}) < 3 \text{ and } \Delta^{\alpha}(D^{\sigma},S) \text{ is bad}\}\ ,$$

where α is a random element of Σ_n. But (*) is equal to

$$n!\Pr\{\Delta^{\beta}(D,S) \text{ is bad and } c(\beta,D) < 3\}$$

where $\beta = \sigma^{-1}\alpha$ is a random permutation. The two events "$c(\beta,D) < 3$" and "$\Delta^{\beta}(D,S)$ is bad" are independent, since the first depends only on matrix entries d_{ij} such that $j \ne \beta(i)$, and the second depends only on $\{d_{i,\beta(i)}\}$. By Lemma 3 the first event has probability $< 4^{-n}$. The probability of the second event is

$$\underset{\substack{x_1+\cdots+x_n \le 3 \\ x_1,\ldots,x_n \ge 0 \\ x_1,\ldots,x_n \le 1}}{\int\cdots\int} dx_1 dx_2 \cdots dx_n \le \frac{3^n}{n!}\ .$$

Thus $(*) \le n!4^{-n}\frac{3^n}{n!} = O((\frac{3}{4})^n)$. □

Lemma 5. If $\Delta^{\alpha}(D,S)$ is a good matrix and D^{σ} is drawn at random from $\{D^{\theta}\in[D]|\theta^{-1}\alpha\in\Sigma_n^S\}$ then, $\Pr\{\theta^{-1}\alpha$ has a cycle, all of whose elements correspond to bad rows of $\Delta^{\alpha}\} = O(\ln n\, n^{-1/6})$.

Proof. The permutation $\theta^{-1}\alpha$ is a random element of Σ_n^S. Thus, restricting $\theta^{-1}\alpha$ to the domain $\{1,2,\ldots,n\}-S$ gives a random permutation ϕ from Σ_{n-m}^m, where $m = |S|$. If the number of bad rows in $\Delta^{\alpha}(D,S)$ is t, then the expected number of cycles of ϕ with all rows bad is

$$\sum_{k=1}^{\min(m,t)} \frac{1}{k} A_k^{n-m} \frac{\binom{t}{k}}{\binom{n-m}{k}}$$

where $\frac{1}{k}A_k^{n-m}$ gives the expected number of cycles of length k in a random permutation from Σ_{n-m}^m, and the ratio $\binom{t}{k}/\binom{n-m}{k}$ is the probability that all the rows of a cycle of length k are bad. Using the facts that

a) $\frac{\binom{t}{k}}{\binom{n-m}{k}}$ is a decreasing function of k;

b) A_k^{n-m} is an increasing function of k (cf. Proposition 3); and

c) $\sum_{k=1}^{m} A_k^{n-m} = n - m$,

we conclude that

$$\sum_{k=1}^{\min(m,t)} \frac{1}{k} A_k^{n-m} \frac{\binom{t}{k}}{\binom{n-m}{k}} \le \sum_{k=1}^{\min(m,t)} \frac{1}{k}\frac{n}{m} \frac{\binom{t}{k}}{\binom{n-m}{k}}$$

$$\le \frac{n}{m}\sum_{i=1}^{\infty}\left(\frac{1}{k}\right)\left(\frac{t}{n-m}\right)^k$$

$$= \frac{n}{m}\left(-\ln\left(1-\frac{t}{n-m}\right)\right) .$$

Using the inequalities

$$\frac{n}{\ln n} \le m \le n - n^{2/3} \quad \text{and} \quad t \le \sqrt{3n \ln 4} + 1 ,$$

pr{ϕ has a cycle with all rows bad}

$$\leq \ln n\left(-\ln\left(1 - \frac{\sqrt{3n \ln 4} + 1}{n^{2/3}}\right)\right) = O(\ln n \cdot n^{-1/6}) \ . \qquad \square$$

Lemma 6. If $\theta^{-1}\alpha \in \Sigma_n^S$ has at most $3 \ln n + 1$ cycles, and has no cycle whose elements all correspond to bad rows of $\Delta^{\alpha}(D,S)$, then

$$\delta(\theta^{-1}\alpha, D^{\theta}, S) \leq 3\sqrt{8 + 2\sqrt{7}}(\ln n)^2 \cdot n^{-1/2} \ .$$

Proof. Under the stated assumptions it is possible to carry out the patching process so that each patch has cost $\leq \sqrt{\frac{(8+2\sqrt{7})\ln n}{m}}$. Hence $\delta(\theta^{-1}\alpha, D^{\theta}) \leq 3 \ln n \sqrt{\frac{(8 + 2\sqrt{7})\ln n}{m}}$. Using the inequality $m \geq \frac{n}{\log n}$, the result follows. $\square$

Lemma 7. Let D be drawn from $U_{n \times n}$. Let $\bar{\pi}$ be the optimal assignment for D. Let $\delta(\bar{\pi}, D)$ denote the cost of applying the patching algorithm to D. Then

$$\Pr\{\delta(\bar{\pi}, D) > 3\sqrt{(8+2\sqrt{7})}(\ln n)^2 n^{-1/2}\} = o(1) \ .$$

Proof. All elements of $[D]$ are equally likely to be drawn. Hence, the desired probability is equal to

$$\frac{1}{n!}E(|\{D^{\sigma} \in [D] | \delta(\sigma^{-1}\bar{\pi}, D^{\sigma}) > 3\sqrt{8+2\sqrt{7}}(\ln n)^2 n^{-1/2}\} \ .$$

In order that $\delta(\sigma^{-1}\bar{\pi}, D^{\sigma})$ be greater than this bound one of three events must occur:

a) D^{σ} is exceptional;

b) D^{σ} is not exceptional and, for some $T \in \mathcal{S}$, $D^{\sigma} \in [D,T]$ and $\Delta^{\bar{\pi}(D^{\sigma})}(D,T)$ is bad;

c) D^{σ} is not exceptional and, for some $T \in \mathcal{S}$, $D^{\sigma} \in [D,T]$, $\Delta^{\bar{\pi}(D^{\sigma})}(D,T)$ is good, and $\delta(\bar{\pi}(D^{\sigma}), D^{\sigma}, T) > 3\sqrt{8 + 2\sqrt{7}}(\ln n)^2 \cdot n^{-1/2}$.

The expected number of matrices for which the first event is true is $o(n!)$. The expected number of matrices for which the second event is true is $\sum_{T \in S} |[D,T]| O(\frac{3}{4})^n$. Here, $O(\frac{3}{4})^n$ is an upper bound on the probability that $\Delta^{\bar{\pi}}(D,T)$ is bad (cf. Lemma 4). By Lemmas 5 and 6 the expected number of matrices for which the third event is true is $\sum_{T \in S} |[D,T]| O(\ln n \cdot n^{-1/6})$. Finally, recalling that $\sum_{T \in S} |[D,T]| \leq 3 \ln n \cdot n!$, since no matrix D^{σ} can occur in more than $3 \ln n$ of the classes $[D,T]$, the result follows. □

<u>Proof of Theorem 1</u>. The inequality (1) can fail only if $c(\bar{\pi}(D),D) < \frac{1}{3}$ or $\delta(\bar{\pi}(D),D) > 3\sqrt{8+2\sqrt{7}}(\ln n)^2 n^{-1/2}$. By Proposition 4 and Lemma 7, the probability of each of these events tends to zero. □

Results analogous to Theorem 1 hold whenever the d_{ij} are drawn from a distribution over $[0,\infty]$ having a bounded density function continuous at 0^+.

4. Heuristic Analysis of a Modified Patching Algorithm

Theorem 1 shows that, when the number of cities is sufficiently large, the patching algorithm tends to give nearly optimal solutions to random nonsymmetric traveling-salesman problems. The result is not entirely satisfying, however, because the upper bound on $\epsilon(D)$ given in the theorem tends to zero very slowly, and is acceptably small only when n is astronomically large.

In this section we present a modified patching algorithm and offer a heuristic argument indicating that its expected patching cost is less than $2n^{-1/2}$.

In the modified patching algorithm, all entries d_{ii} are set to $+\infty$. This ensures that the optimal assignment permutation will have no fixed points; i.e., no cycles of length 1. All permutations without fixed points remain equally likely to occur as the optimum assignment.

Having constructed the optimal assignment $\bar{\pi}$, the algorithm converts it to a tour as follows:

MODIFIED PATCHING PROCESS

```
     σ ← π̄
while σ has more than one cycle do;
        begin;
        let C be a shortest cycle of σ;
        let R_ij be a minimum-cost patching operation such that i ∈ C, j ∉ C,
        and neither i nor j has been involved in a previous patching operation;
        σ ← R_ij(σ)
        end.
```

Thus, we no longer restrict attention to patching operations that join the short cycles of $\bar{\pi}$ directly into the longest cycle of $\bar{\pi}$. Figure 3 indicates how the modified patching process converts a permutation to a tour.

Next we study the behavior of the modified patching algorithm on a special class of matrices, and argue heuristically that the algorithm should have similar behavior when applied to matrices from $U_{n\times n}$.

We denote the special class of matrices by M. A matrix D is in the class M if the row minima in D lie in distinct columns, and hence determine the optimal assignment for D. Formally, $D \in M$ if there is a permutation $\bar{\pi}$ such that, for all i and j, $d_{i,\bar{\pi}(i)} \leq d_{ij}$.

The following theorem states that, when a matrix is drawn at random from M, the patching costs Δ_{ij} tend to be at least as small as they would be if the Δ_{ij} were independent, and each were distributed as the sum of two independent samples from the uniform distribution over $[0,1]$.

To frame the theorem precisely, we introduce the concept of stochastic dominance. Let $X = (x_1,x_2,\ldots,x_m)$ and $Y = (y_1,y_2,\ldots,y_n)$ be two random variables over R^m, where R denotes the reals. We say $X \prec Y$ (X is stochastically smaller than Y) if, for every $A = (a_1, a_2, \ldots, a_m) \in R^n$, $\Pr\{X < A\} \geq \Pr\{Y < A\}$. Here $X < A$ if, for all i, $x_i < a_i$.

Let $\bar{\pi}$ be a fixed permutation of $\{1,2,\ldots,n\}$ without fixed points. Let $X = \{x_{ij} | 1 \leq i < j \leq n,\ j \neq \bar{\pi}(i) \text{ and } i \neq \bar{\pi}(j)\}$ be the random variable over $R^{\binom{n}{2}-n}$ determined by the following experiment: draw a matrix D from the set of matrices in M having $\bar{\pi}$ as their optimal assignment; then let $x_{ij} = \Delta_{ij}(\bar{\pi}, D)$. Let $Y = \{y_{ij} | 1 \leq i < j < n,\ j \neq \bar{\pi}(i) \text{ and } i \neq \bar{\pi}(j)\}$ be the random variable over $R^{\binom{n}{2}-n}$ determined as follows: the y_{ij} are independent, and each is the sum of two independent samples from the uniform distribution over $[0,1]$.

<u>Theorem 2</u>. $X \prec Y$

<u>Proof</u>. We condition on arbitrary fixed values for the entries $d_{i,\bar{\pi}(i)}$. Then the d_{ij}, $j \neq \bar{\pi}(i)$, are independent, with d_{ij} distributed according to a uniform distribution over $[d_{i,\bar{\pi}(i)}, 1]$. Hence the differences $d_{ij} - d_{i,\bar{\pi}(i)}$ are independent, and each such difference is drawn from a uniform distribution over $[0, 1-d_{i,\bar{\pi}(i)}]$. Hence the Δ_{ij} are independent of one another, and each particular patching cost $\Delta_{ij} = (d_{ij} - d_{i,\bar{\pi}(i)}) + (d_{ji} - d_{j,\bar{\pi}(j)})$ is the sum of two independent random variables; one drawn from the uniform distribution over $[0, 1-d_{i,\bar{\pi}(i)}]$, and the other from the uniform distribution over $[0, 1-d_{j,\bar{\pi}(j)}]$. The result now follows, since a random variable uniformly distributed over $[0, 1-d_{i,\bar{\pi}(i)}]$ stochastically dominates a random variable uniformly distributed over $[0,1]$. □

We conjecture that an analogous property holds when matrices drawn from $U_{n \times n}$, rather than M, are considered. More precisely, let $\bar{\pi}$ be a fixed permutation without fixed points. Let $Z = \{z_{ij} | 1 \leq i < j \leq n,\ j \neq \bar{\pi}(i),\ i \neq \bar{\pi}(j)\}$

be a random variable over $R^{\binom{n}{2}-n}$ determined by the following experiment: draw a matrix D from the set of matrices in $U_{n\times n}$ having $\bar{\pi}$ as their optimal assignment; then let $z_{ij} = \Delta_{ij}(\bar{\pi},D)$.

<u>Conjecture</u>. $Z \prec Y$

As a heuristic argument in support of the conjecture, we define a mapping $\tau: U_{n\times n} \to M$ as follows. Let $D \in U_{n\times n}$ have $\bar{\pi}$ as its optimal assignment. Then

$$(\tau(D))_{i,\bar{\pi}(i)} = \min_j d_{ij}$$

for $k \neq \pi(i)$

$$(\tau(D))_{ik} = \begin{cases} d_{ik} & \text{if } d_{ik} \neq \min_j d_{ij} \\ d_{i,\bar{\pi}(i)} & \text{if } d_{ik} = \min_j d_{ij} \end{cases}.$$

Thus, $\tau(D)$ is obtained by interchanging the minimum element in each row with the element of that row which occurs in the optimum assignment. The following facts are immediate.

a) The matrices D and $\tau(D)$ have the same optimal assignment $\bar{\pi}$;

b) For all i and j, $\Delta_{ij}(\bar{\pi},D) \leq \Delta_{ij}(\bar{\pi},\tau(D))$.

Theorem 2, coupled with condition (b), which asserts that the patching costs aasociated with $\tau(D)$ are at least as great as those associated with D, tends to support the conjecture. To prove the conjecture, it would be necessary to show that, when D is uniformly distributed over $U_{n\times n}$, its image $\tau(D)$ is approximately uniformly distributed over M.

In view of Theorem 2 and Conjecture 1, it is of interest to elucidate the behavior of the modified patching algorithm when $\bar{\pi}$ is a random permutation without fixed points, the Δ_{ij} are independent, and each is the sum of two independent random variables uniformly distributed over $[0,1]$. We do so briefly, omitting details. Let the random variable γ_n denote the cost of the modified patching process under these assumptions.

<u>Theorem 3</u>. $\lim_{n\to\infty} n^{1/2}E(\gamma_n) \leq 2$

The underlying ideas of the proof are as follows:

a) the expected number of cycles of length k in a random permutation without fixed points is $\frac{1}{k}(1+O(n^{-1}))$;

b) given a cycle C of length k, $E(\min\{\Delta_{ij} | i \in C, j \notin C\}) \leq \sqrt{\frac{\pi}{k(n-k)}}$.

These facts suggest that the expected patching cost is bounded above by $\sum_{k=2}^{n} \frac{1}{k}(1+O(n^{-1}))\sqrt{\frac{\pi}{2k(n-k)}} \sim 2n^{-1/2}$. The proof becomes more complicated than this sketch because of the possibility that the short cycles of $\bar{\pi}$ may become joined as the patching process takes place. We omit further details.

A Monte Carlo simulation was conducted to further determine the behavior of the random variable γ_n. The simulation was equivalent to determining γ_n at each of 100 random choices of $\bar{\pi}$ and $\{\Delta_{ij}\}$, for each of the values $n = 100$, $n = 1000$, and $n = 10{,}000$. The simulation avoided explicit generation of random permutations and random patching costs; instead, it conducted a probabilistically equivalent experiment using theoretical properties of the cycle structure of a random permutation, and of the distribution of the minimum of a given number of independent patching costs. The results were as follows.

n	100	1000	10,000
sample size	100	100	100
sample mean	.18	.067	.018
sample mean × $\sqrt{n}$	1.8	2.1	1.8
sample median × $\sqrt{n}$	1.6	2.0	1.7
sample maximum × $\sqrt{n}$	5.4	4.9	4.0

Table 1 - Simulated Behavior of the Random Variable γ_n

References

1. Edmonds, J. and R.M. Karp, "Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems," Journal of ACM 19, 2, 248-264 (April, 1972).

2. Erdös, P. and J. Spencer, Probabilistic Methods in Combinatorics, Academic Press, New York (1974).

3. Gilmore, P. and R. Gomory, "Sequencing a One State Variable Machine: A Solvable Case of the Traveling Salesman Problem," Operations Research 12, 655-679 (1964).

4. Karp, R.M., "Reducibility Among Combinatorial Problems," in Complexity of Computer Computations (edited by R.E. Miller and J. Thatcher), Plenum Press (1972).

5. Karp, R.M., "Probabilistic Analysis of Partitioning Algorithms for the Traveling-Salesman Problem in the Plane," Mathematics of Operations Research (August, 1977).

6. Thompson, G.L., "Algorithmic and Computational Methods for Solving Symmetric and Asymmetric Traveling-Salesman Problems," Working Paper presented at the Workshop on Integer Programming, Bonn (1975).

7. Tomizawa, N., "On Some Techniques Useful for Solution of Transportation Problems," Networks 1, 173-194 (1972).

8. Walkup, D., "On the Expected Value of a Random Assignment Problem," submitted to SIAM J. Comput. (1977).

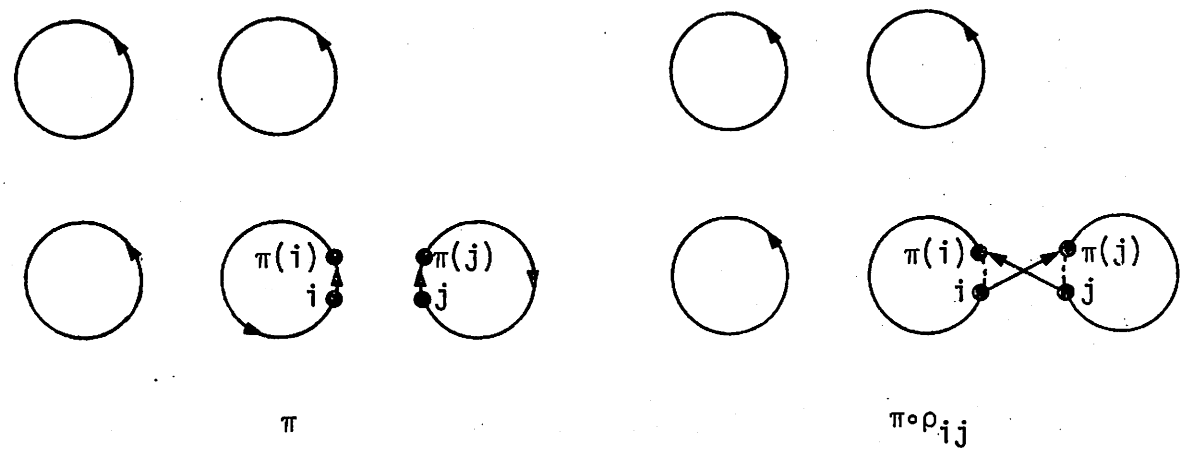

Figure 1: Effect of the i,j Patching Operation

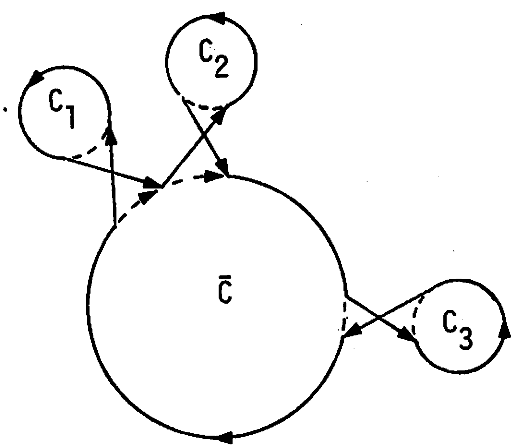

Figure 2: Application of the Patching Process

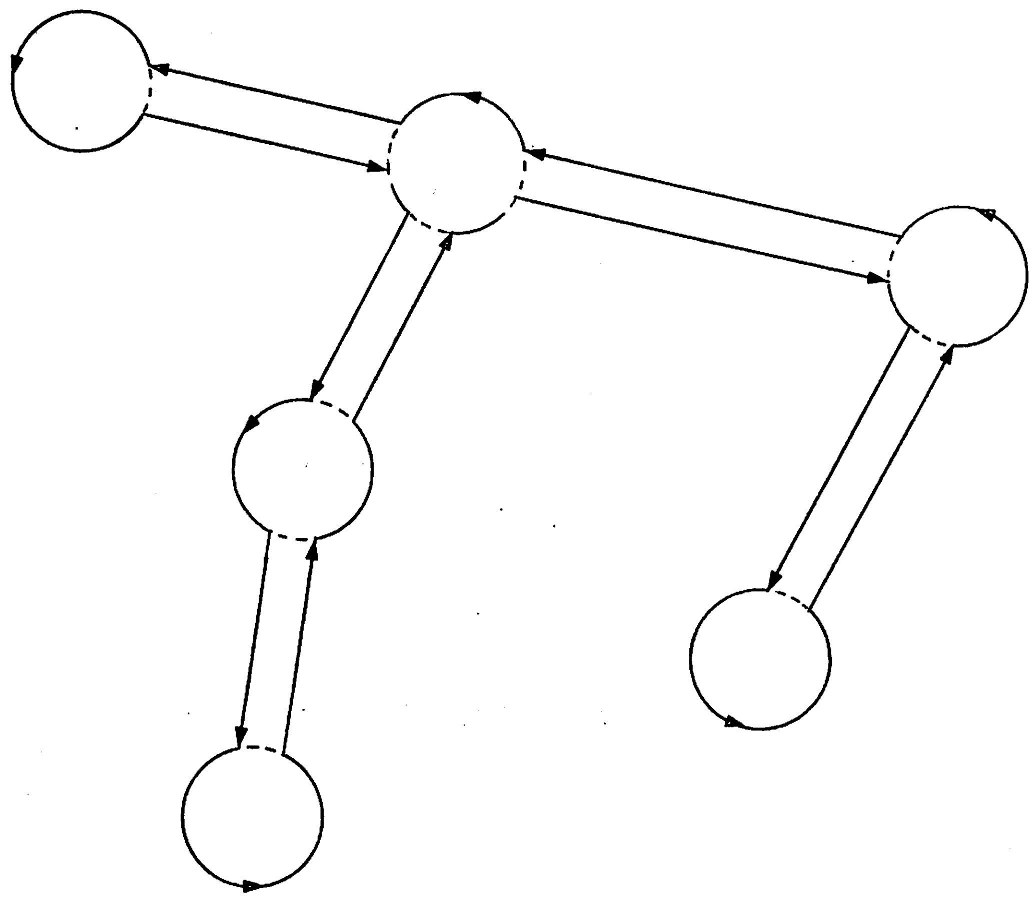

Figure 3: Application of the Modified Patching Process