

# A double chromosome structure to solve TSP in GA

遺伝的アルゴリズムにおける巡回セールスマン問題を解くための二重染色体構造

Prof. Matayoshi, Mitsukuni

## Abstract

In this paper, we represent a solution for TSP by a double chromosome structure. And we try to solve some TSP benchmark tests by it with procedures of GA. The GA procedures consist of an elite selection strategy for selection, two-change (exchange)/insertion mutation and a subtour exchange crossover. We also use a local search method using insertion method.

In experimental results, we show that some small benchmark tests can be efficiently solved by the double chromosome structure GA compared with a conventional single chromosome structure GA.

## 1. Introduction

Genetic Algorithms (GAs) imitates the organic evolution and is known as a method for solving the combinatorial optimization problems.

To solve target problems, GAs use a reproduction processes contains crossover, mutation and selection. In many cases, even target problems have multiple features, the conventional GAs use a single chromosome representation of solutions.

Bagchi[1] represented multiple

features of a Job Shop Scheduling Problem within a single chromosome. Job order, process plans and resources were all represented in a single chromosome in order that the GA could search over the entire search space.

In Evolutionary computation, many researchers use a single chromosome representation and obtain good results.

On the other hand, Chromosome structure of the creature has the double helix is the well-known fact. It is the structure of DNA (Deoxyribonucleic Acid) as first published by James Watson and Francis Crick in 1953. Helmut [2] investigated the potential benefits of multi-chromosomal representations for genotypes processed by an Evolutionary Algorithm in 2003. Kate[3][4] used "multi-chromosome" in 1993. Kate applied multi-chromosome representation to a combinatorial optimization constraint satisfaction problem of pallet loading. Robert[5] extended self-adaptation to nonnumeric problems in GA by using a multi-chromosome representation. Markus[6] presented experiments evolving the membership functions of a fuzzy controller for the inverted pendulum problem with multiple chromosomes in an evolutionary

algorithm.

Matayoshi[7][8] also used a multi-chromosome structure and succeeded in to identify some kinds of functions. The system could obtain some target functions in a few minutes on 350MHz Pentium machine.

We propose a chromosome structure to solve Traveling Salesman Problem (TSP) is known as NP-hard. The chromosome structure has two chromosomes such the double helix of the creature. We show that our proposed structure can solve some TSP benchmark tests quickly but some tests are not easy to solve compared with a conventional single chromosome structure. We give some consideration in the last section.

## 2. TSP and Double chromosome Structure

### 2.1 TSP

We solve Symmetric Traveling Salesman Problem (STSP). We represent an assignment by the vector  $\pi$ , which is a permutation of the numbers (cities){1, 2, ..., n}.  $\pi(i)$  is the location to which city j is assigned. With these definitions, STSP can be written as

$$\min C = \sum_{i=1}^N d_{\pi(i)\pi(i+1)} \cdots (1)$$

where N is the number of cities,  $d_{ij}$  is the Euclidean 2-dimensional distance between city i and city j. The distance from city i to city j is the same as from city j to city i ( $d_{ij}=d_{ji}$ ). The goal of solving STSP is to find a

round trip of minimum total length visiting each city exactly once.

Many methods by GAs make a single chromosome structure along the definition. The encoding of TSP solution in conventional implementation is straightforward. Figure 1 shows a conventional genotype chromosome structure for TSP when the array number is the round order and the array element is the city. The representation shown in Figure 1 is called path representation. The round order is called the tour and the part of tour is called subtour. It only has not to be solution that cannot be executed for making a Hamilton cycle. If there is overlapping value in the chromosome then it becomes fatal chromosome.

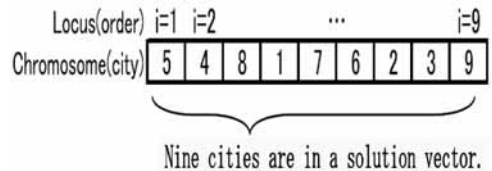


Figure 1: Conventional Single Chromosome Structure (9 cities).

### 2.2 The round order by double chromosome structure

Figure 2 is a simple tour to explain our double chromosome structure representation. We find out city 1 enter from city 2, and goes out of city 3.

We make the broken line route and the solid line route in Figure 2 represents to the double chromosome structure in Figure 3. In Figure 3, the above one is for the implement of the

solid route, and the below one is for the broken route. The number of array means two aspects that are for the locus of chromosome and current city. The elements of the Out-chromosome array mean the destination city of the current city. The elements of the In-chromosome array mean the departure city to the current city. To solve TSP, our chromosome is composed of the two chromosomes (double chromosome structure).

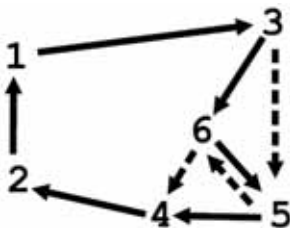


Figure 2: Solid route and Broken route.

	Solid route					
Locus(current city)	1	2	3	4	5	6
Out-Chromosome(destination)	3	1	6	2	4	5
In-Chromosome(departure)	2	4	1	5	6	3

	Broken route					
Locus(current city)	1	2	3	4	5	6
Out-Chromosome(destination)	3	1	5	2	6	4
In-Chromosome(departure)	2	4	1	6	3	5

Figure 3: Two chromosomes for Solid route and Broken route.

The double chromosome structure consists of Out-chromosome and In-chromosome that gives a path representation. Where we call our representation "double-chromosome" instead of "multi-chromosome" because our representation did not use no more than two chromosomes per individual to represent TSP and also chromosome's

gene represents city.

### 3. Genetic Algorithms.

GA is well-known methods for solving combinatorial optimization problems in the realm of Evolutionary Computation. GAs have three basic operations called selection, crossover and mutation, and repeat the operations to reproduce offspring until the termination conditions of search is satisfied.

Thereinafter we show that GA's three basic operations is applicable to double chromosome structure.

#### 3.1 Selection

The elite selection strategy, whereby the most suitable individual succeeds to the next generation, is used in this paper. More than one elite could exist, in which case the one to survive to the next generation is selected at random. This operation relates to not the chromosome structure but the evaluation of individual. Therefore selection operation operates without problems in GA with double chromosome structure.

#### 3.2 Crossover

We use the subtour exchange crossover(SXX)[9]. SXX can exchange the set of cities when the subtour of each parent is consisted by same set of cities.

In TSP, genetic character, which should be inherited to offspring, is subtour. A subtour is provided by two loci on the chromosome in conventional

chromosome such as Figure 1. Our proposed double chromosome structure also provides a subtour by two loci. For example, there are 4 cities {1,2,3,4} or {3,4,5,6} between the city 4 and the city 3 in Figure 2. These cities are found by tracing destination genes from the city 4 to the city 3 in the Out-chromosome or backward process in the In-chromosome. These set of cities, which are the subtours of Solid(line) route and Broken(line) route, are able to exchange each other.

However we think that SXX does not make our double chromosome structure bring out the best. Because Yanagiura[10] showed that the expected number of common subtours for two random tours is at most  $4 + O(n^{-1})$ . Therefore, to invent new crossover method for double chromosome structure is one of the important our future works.

### 3.3 Mutation -Exchange-

Conventional mutation operation in single chromosome structure is exchange two genes (see Figure 4). Now we call the conventional mutation the exchange mutation.

In this subsection, we indicate that exchange mutation is possible in double chromosome structure. The steps of exchange process are constant 3 steps in single chromosome structure but double chromosome structure requires 12 steps or less. Broken tour in Figure 5 is the changed tour by

exchanging 8 shadowed genes (see Figure 6).

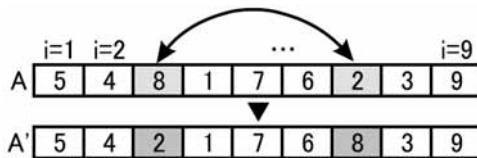


Figure 4: Exchange mutation genes in single chromosome structure.

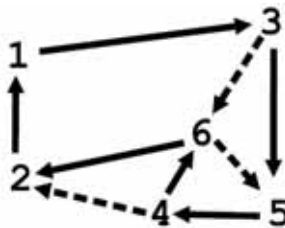


Figure 5: Mutation by exchange genes by Figure 6.

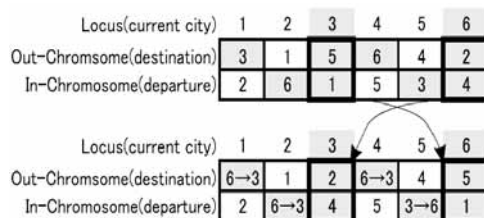


Figure 6: Mutation by exchange genes in double chromosome.

Although double chromosome executes exchange mutation in 12 steps or less, we think that exchange mutation may destroy the original tour more than in single chromosome. In Figure 4, the two subtour of {1,7,6} and {3,9,5,4} are kept. However there is one subtour {5,4} in Figure 5. Even if we take into account the fact that the original tour of Figure 5 is different from in Figure 4, the exchange mutation is not good for double chromosome from inheritance of blueprint point of view.

Therefore we experimented with exchange mutation in double chromosome for the sake of comparison against single chromosome in later Tests and Results section.

### 3.4 Mutation -Insertion-

In above crossover section, we said that genetic character that should be inherited for offspring is subtour in TSP. And we also suppose that exchange mutation in double chromosome structure does not work well in Mutation -Exchange- section.

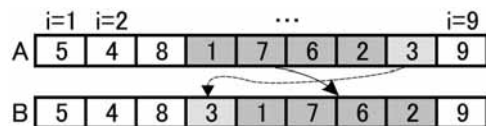


Figure 7: Mutation by insertion in single chromosome structure.

Accordingly we think about a insertion mutation as alternated exchange mutation. Insertion mutation method keeps subtour compare with exchange mutation by two genes. Figure 7 provides of the insertion in Figure 1. In this case, the insertion mutation makes the shadowed genes shifted after the insertion locus ( $i=4$ ) without breaking subtours but requires the same steps of the number of genes.

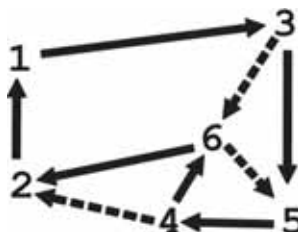


Figure 8: The route change by Insertion.

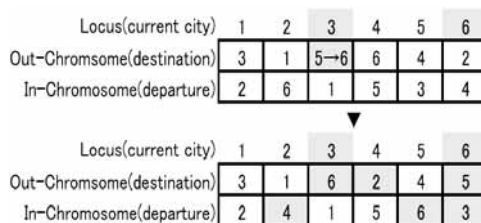


Figure 9: Mutation by insertion in double chromosome structure.

Our proposed double chromosome structure changes no more than 6 genes on In-chromosome and Out-chromosome for insertion mutation.

It is 12 steps or less regardless of the length of chromosome.

In Figure 8, the city 6 is inserted into between city 3 and city 5 as the new tour. We indicate in Figure 9 how to change the genes on the double chromosome. Note the shadowed genes are just 6.

Because we would like to know the effects by keeping subtour structures and also by increasing / decreasing the calculation steps, we tried to experiment with the exchange mutation and the insertion mutation separately. It is also in order to make clear the ability of double chromosome structure against the single chromosome structure.

### 3.5 Local Search

We also use insertion method for a local search method with Fast Admissible Move Strategy (FAMS) in GA. FAMS searches neighbors of the solution, which is as initial solution,

at random and if FAMS finds improving solution by exchange two genes then the next initial solution becomes improving one. FAMS executes the above-mentioned movement until the improving solution is not gotten any longer.

It has been told that insertion neighborhood is worse than exchange neighborhood[11]. It is clear and natural that if use straightforward conventional single chromosome structure (see Figure 7). But we think that character which should be inherited is the subtour in TSP. Therefore it is considered that LS with insertion neighborhood is a good search method to solve TSP. However, The insertion method does not get shortening of the search time from viewpoint of calculation costs in using single chromosome structure.

On the other hand, our double chromosome structure might be able to achieve the desired effect, which is shortening search time, because the steps of calculation are constant steps regardless of chromosome length (number of genes).

The disadvantage of insertion method using insertion neighborhood is for conventional single chromosome that is represented as a permuted string. If we use double chromosome structure then there is not always disadvantage for calculation cost.

### 3.6 GA Process

We describe an outline of GA

process, which is used in our experiments. GA starts from initial population of which genes made at random without generating lethal gene(s), and iterates a reproduction process consists of crossover, mutation and selection until a termination condition is satisfied. The local search method gives effect to the operated individuals that are selected from population at random and are operated by mutation or crossover. In mutation case, the previous individual changes into the offspring disregard the evaluation value. In a word, the operated descendant remains in the population without fail. In crossover case, the evaluation value of offspring is compared with the parents' one. If either of parents is worse than offspring then one of worse parents is exchanged for offspring. The size of population has kept since the iteration started.

## 4. Tests and Results

We tested two chromosome structures. One of them is conventional single chromosome structure and other one is proposed double chromosome structure.

### 4.1 Environment and Settings

The experimental environment was a AMD Sempron 2600+ machine with Turbo Linux 10 and the programming language used was Java (ver.1.4.1). Basic parameters of genetic operation of GA at GLS did not change throughout

the experiment. The population size of each benchmark tests and the termination condition were shown in Table 3 and Table 1. Ratio of mutations among the population (rate of mutation) was 10%, rate of mutations among the length of chromosome (rate of mutational length) was 10%, and crossover rate was 50%. These values were decided through preliminary experiments based on experience.

The number of experiments is 100 times in each benchmark tests which optimal evaluation value is known already.

Because the number of experimental runs was 100, we selected 10 TSP benchmark tests[12] in 30 minutes in our environment, and not more than 724 in tour size.

## 4.2 Results

Table 1 shows the improvement rate results of average route length by

conventional Single Chromosome Structure (SCS) or Double Chromosome Structure (DCS) at the finish time of search, and by using an exchange mutation and an insertion LS (E&I), and by using an insertion mutation and an insertion LS (I&I).

In Table 1, there is approximately slight difference of two structures in benchmark tests size 50. Proposed double chromosome structure obtained good results in st70~a280. Single-chromosome structure got good results in pcb442~u724 comparison with DCS. DCS could not surpass SCS in the evaluation value in the problem size is more than 442.

As for a280, DCS obtained good results in 94 points that are in 140 sampling points in the total search time. We can see these phenomena in Figure 10, and we also find out that double chromosome structure preserved a diversity of solution. The two circles

Table 1: Improvement rates and Setting values

TSP	Opt.	E&I(%)		I&I(%)		POP	TC
		SCS	DCS	SCS	DCS		
ei151	426	0.23	0.14	0.02	0.17	50	600
berlin	7542	0.00	0	0.00	0.06	50	600
st70	675	2.56	0.95	1.48	0.52	70	600
eil101	629	5.64	3.75	4.83	2.96	100	1200
lin105	14379	6.03	2.66	5.08	2.25	110	1200
ch150	6528	10.60	5.83	9.86	4.94	150	1200
tsp225	3916	13.43	9.85	13.21	9.94	230	1800
a280	2579	23.15	23.01	24.47	23.34	280	1800
pcb442	50778	25.68	38.19	27.78	39.04	440	1800
u724	41910	48.06	65.19	50.22	66.21	730	1800

E&I(%)=100(Opt.-E&I of DCS or SCS)/Opt.

I&I(%)=100(Opt.-I&I of DCS or SCS)/Opt.

POP: Population size.

TC: Termination Condition(Sec).

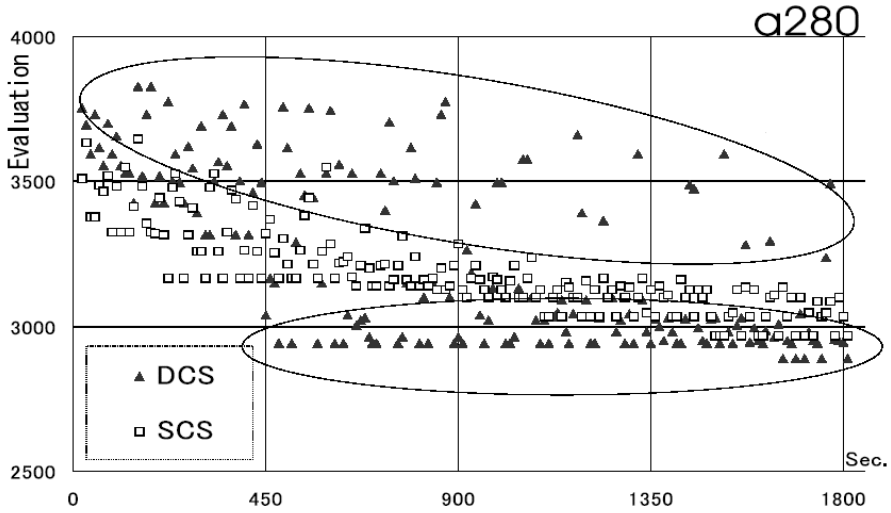


Figure 10: The best tour's cost of a280 with I&I in each sampling interval.

Table 2: Wilcoxon rank-sum test's difference between SCS and DCS by E&I(average)

ei151	114	713**	789**	824**
berlin52	158	292**	510**	660**
st70	-698++	-884++	-772++	-798++
eil101	-1989++	-3111++	-3395++	-3482++
lin105	-2018++	-3504++	-3480++	-3452++
ch150	-176	-3516++	-3548++	-3600++
tsp225	1403**	5	-1803++	-1983++
a280	1691**	1855**	1495**	-123
lin318	1253**	1749**	1183**	773**
pcb442	1849**	2025**	2025**	2025**
att532	1703**	2025**	2025**	2025**
u724	1627**	2025**	2025**	2025**

++, \*\*: 1% significant difference,  $p < 0.01/2$ .

+, \*: 5% significant difference,  $p < 0.05/2$ .

in Figure 10 mean the two sets of dispersion of individuals of DCS in 140 sampling points. The individuals in upper circle are not good evaluation value group. The individuals in lower circle are good evaluation value group. Figure 10 also shows that double chromosome structure obtained a lot of better solutions than single chromosome structure since 40 seconds. We guess that DCS with insertion strategy

may obtain good performances up to city size about 300.

We examined the results that had been divided into four (-1-, -2-, -3-, and -4-) time regions. We can see the divisions on Figure 11.

If we see the line charts of SCS and DCS in Figure 11, we understand the reason soon. The two line charts show different aspects at initial, middle term, and the end of the search.



Table 3: Wilcoxon rank-sum test's difference between SCS and DCS by I&I(average)

ei151	-37	478**	390**	300**
berlin52	184*	30	0	120*
st70	-707++	-755++	-530++	-438**
eil101	-2509++	-2485++	-2828++	-2580**
lin105	-2740++	-2926++	-2874++	-3002**
ch150	-878+	-2269++	-2886++	-3234**
tsp225	487*	-1005++	-1107++	-1783**
a280	1225**	-173	-697++	-1015**
lin318	889**	297	-697++	-1115**
pcb442	1543**	1607**	1855**	1913**
att532	1565**	1695**	1889**	2025**
u724	1583**	1951**	2025**	1925**

++, \*: 1% significant difference,  $p < 0.01/2$ .

+, \*: 5% significant difference,  $p < 0.05/2$ .

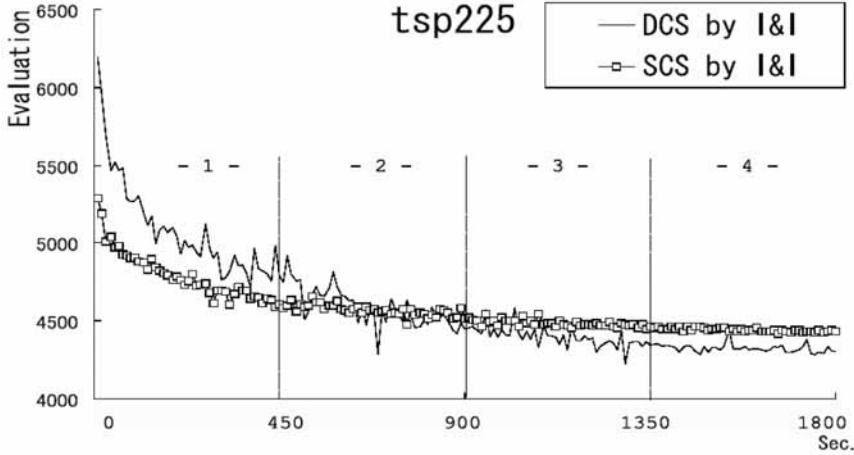


Figure 11: Average tour cost of tsp225 in each sampling interval.

Table 2 and Table 3 show the difference of non-parametric Wilcoxon rank-sum, which is a statistical technique, between SCS and DCS. The symbols in Table 2 and Table 3 are the examined results of non-parametric Wilcoxon rank-sum test. {++,\*\*} mean that there is 1% significant difference,  $p < 0.01/2$ . {+,\*} mean that there is 5% significant difference,  $p < 0.05/2$ . For brevity, benchmark tests marked by "\*\*", DCS required longer search time, but "+" did not require in comparison

with SCS.

Table 2 and Table 3 give the validity of consideration to Table 1.

## 5. Conclusion

We proposed a double chromosome structure for TSP representation and solved some TSP benchmark tests using GA approach. That is SXX, an exchange mutation and an insertion LS(E&I) or an insertion mutation and an insertion LS(I&I). The advantage of proposed double chromosome struc-

ture is that needs constant steps in the mutation and the local search.

The results showed that our proposed structure obtained good results until problem size 280, but did not run pcp442~u724 in comparison with conventional single chromosome structure.

Therefore, as future works, we have to study that why our proposed structure does not work more than problem size 442. In addition, we have to study that is to device a new crossover for double chromosome structure to educe its ability.

Because of in spite of taking a lot of search time, the improvement rate is not good in each chromosome structure SCS and DCS (see Table 1). The reason results from SXX, which calculation cost to find a subtour is very large[10]. We hope it is surmountable by a new crossover for double chromosome structure.

## REFERENCES

[ 1 ] S.Bagchi, S.Uckun, and Y.Miyabe, "Exploring problem-specific recombination operators for job shop scheduling," *Proceedings of the Fourth International Conference on Genetic Algorithms*, pp.13-17, 1991.

[ 2 ] H. A. Mayer and M. Spitzlinger, Multi-chromosomal representations and chromosome shuffling in evolutionary algorithms," *Electronic Proceedings of the 2003 Congress*

*on Evolutionary Computation*, pp. 1145-1149, December 2003.

- [ 3 ] K.Juliff, "A multi-chromosome genetic algorithm for pallet loading," *Proceedings of the Fifth International Conference on Genetic Algorithms*, pp.467-473, 1993.
- [ 4 ] R. Hinterding and L. Khan, "Genetic algorithms for cutting stock problems: With and without contiguity," *Technical Report 40 COMPI2 Department of Computer and Mathematical Sciences, Victoria University of Technology, Victoria Australia*, pp.166-186, August 1994.
- [ 5 ] R.Hinterding, "Self-adaptation using multi-chromosomes," *Proceedings of the 4th IEEE International Conference on Evolutionary Computation*, pp.87-91, April 1997.
- [ 6 ] M. Spitzlinger and H. A. Mayer, "Evolution of fuzzy controllers with multi-chromosomal representation of membership functions," *Proceedings of the 10th International Fuzzy Systems Association World Congress (IFSA 2003)*, pp.417-420, June 30-July 2, 2003.
- [ 7 ] M. MATAYOSHI, "Tree chromosome structure in a genetic algorithm to identify functions," *Systems and Computers in Japan*, vol. 31, No.10, pp.32-40, 2000.(Japanese).
- [ 8 ] M. MATAYOSHI, M. NAKAMURA,

- and H. MIYAGI, "Local search methods for tree chromosome structure in a genetic algorithm to identify functions", *IEEJ Trans. EIS*, vol.126, pp.123-131, January 2006. (Japanese).
- [ 9 ] M. Yamamura, T. Ono, and S. Kobayashi, "Character-preserving genetic algorithms for traveling salesman problem," *Journal of the Japanese Society for Artificial Intelligence*, vol.7, pp.1049-1059, November 1992. (Japanese).
- [10] M. Yanagiura, H. Nagamochi, and T. Ibaraki, "Two comments on the subtour exchange crossover operator," *Journal of the Japanese Society for Artificial Intelligence*, pp.464-467, 1995. (Japanese).
- [11] M. Yangiura and T. Ibaraki, "On metaheuristic algorithms for combinatorial optimization problems," *Systems and Computers in Japan*, vol.32, no.3, pp.33-55, 2001.
- [12] G. Reinelt, "Tsp-lib," *Updated Version: <http://www.iwr.uni-heidelberg.de/groups/comopt/soft/TSPLIB95/TSPLIB.html>*, November 2004.