Linköping University

Metaheuristics

Sections 2.1 – 2.2

Mina Niknafs

LINKÖPING UNIVERSITY

Outline

- Common concepts for single-solution based metaheuristics
  - Neighborhood
  - Very large neighborhoods
    - Heuristic search in large neighborhood
    - Exact search in large neighborhood
  - Initial solution
- Fitness landscape analysis
  - Distance in search space
  - Landscape properties

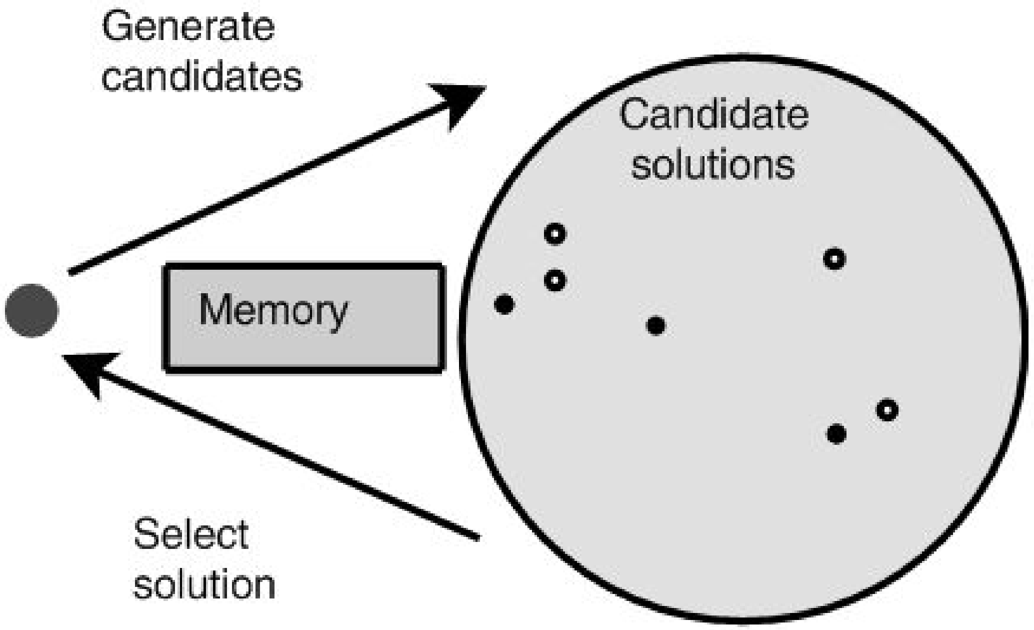
Single-Solution Based Metaheuristics (S-metaheuristics)

- Improve a single solution
  - walk though neighborhoods by performing iterative procedures that move from the current solution to another one in the search space
- Have two iterative phases:
  - Generation phase
  - Replacement phase

Main Principles of S-metaheuristics

A set of candidate solutions are generated from the current solution s

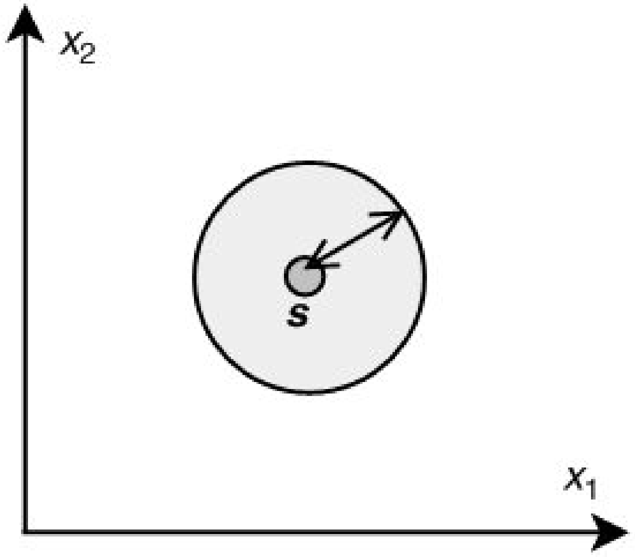
A selection is performed from the candidate solution set C(s)



Neighborhood

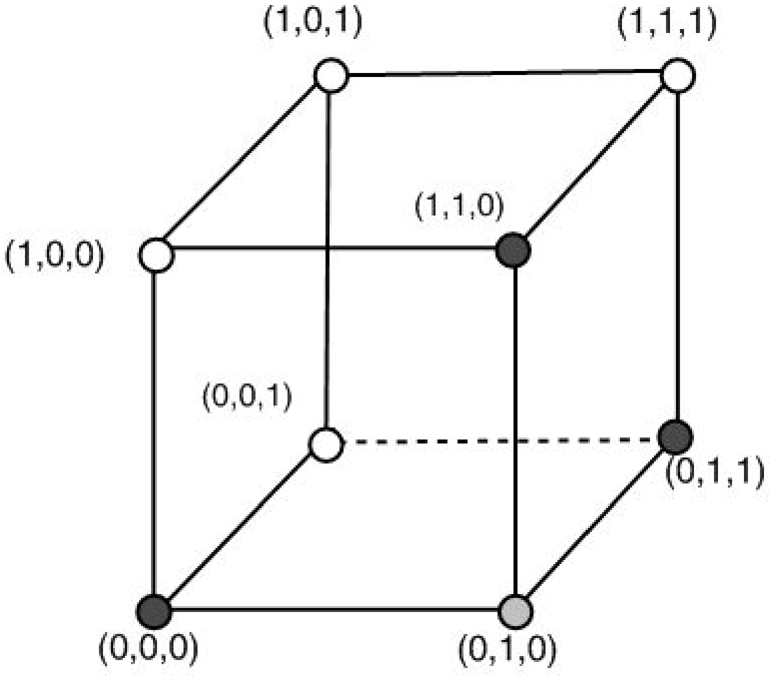
- **Neighborhood**
  - A neighborhood function *N* is a mapping $N : S \rightarrow 2^S$ that assigns to each solution s of S a set of solutions $N(s) \subset S$.
  - *Move operator m*
  - *Depends strongly on the representation*
    - *Continuous or discrete space*
- **Locality:** the effect on the solution when performing the move in the representation
  - *Strong locality*
  - *Weak locality*

Neighborhood of a continuous and a discrete binary problem



The circle represents
the neighborhood of *s*
in a continous problem
with two dimensions.

Euclidean distance

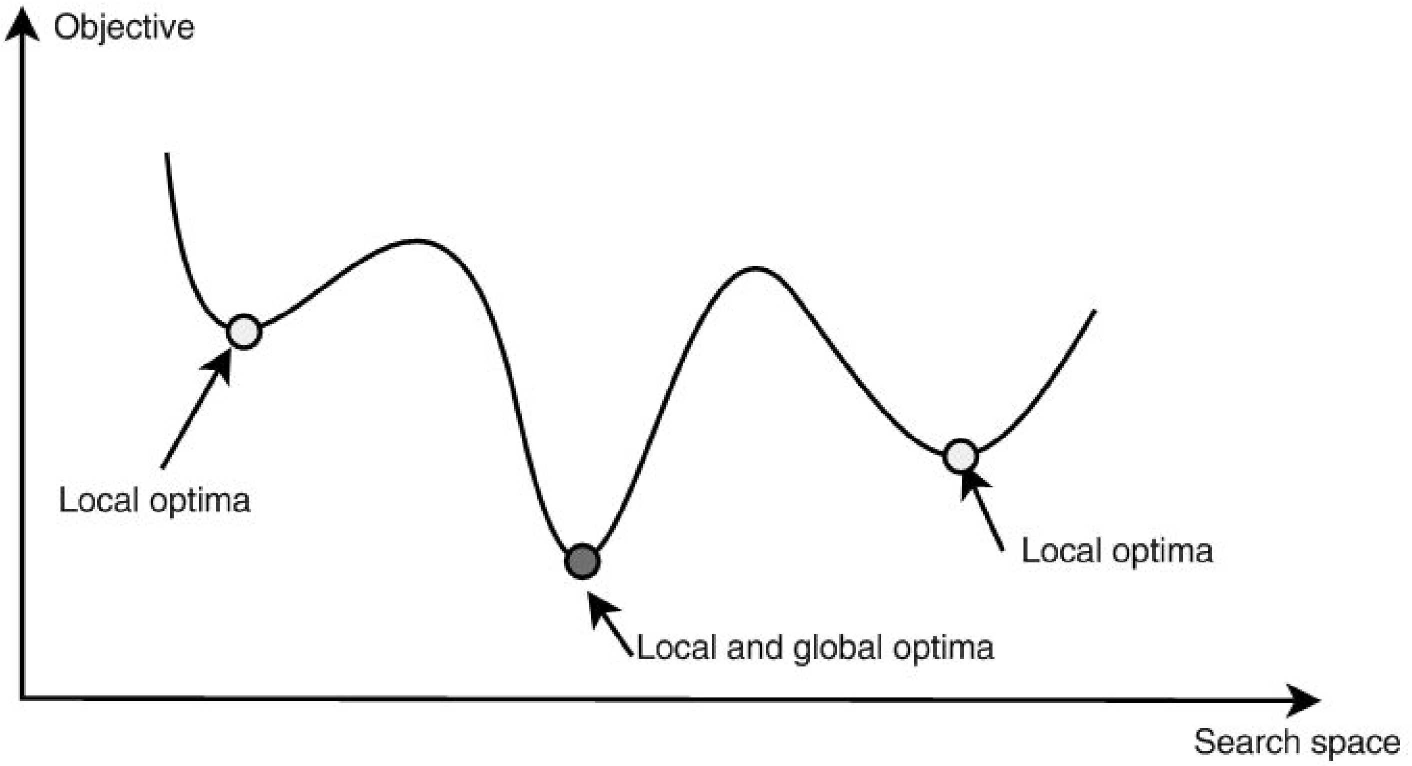


- Nodes of the hypercube represent solutions of the problem.
- The neighbors of a solution (e.g., (0,1,0)) are the adjacent nodes in the graph.

Hamming distance

Local optimum

Definition 2.4 Local optimum. *Relatively to a given neighboring function N, a solution $s \in S$ is a local optimum if it has a better quality than all its neighbors; that is, $f(s) \leq f(s')$ for all $s' \in N(s)$ (Fig. 2.4).*



k-distance neighborhood vs. k-exchange neighborhood

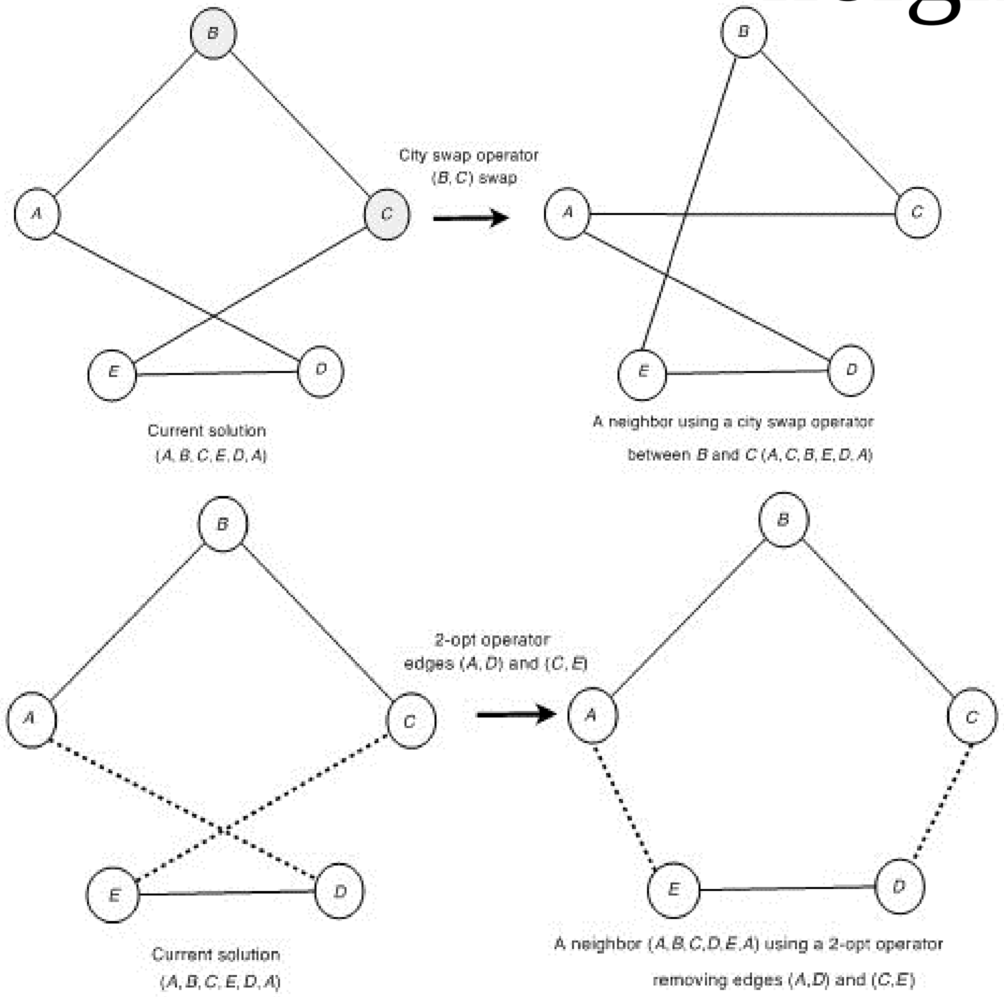


FIGURE 2.5 City swap operator and 2-opt operator for the TSP.

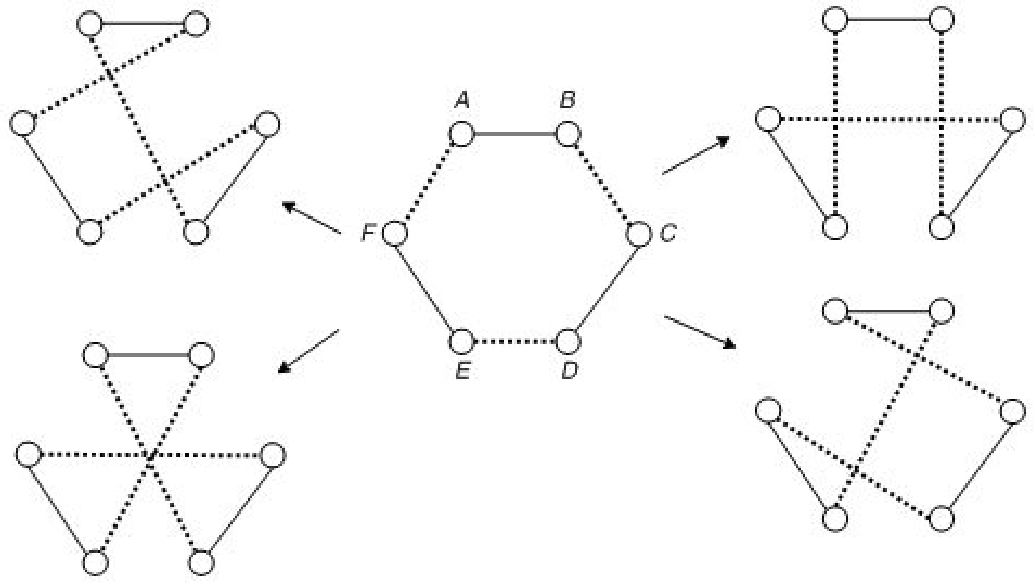


FIGURE 2.6 3-opt operator for the TSP. The neighbors of the solution (*A*,*B*,*C*,*D*,*E*,*F*) are (*A*,*B*,*F*,*E*,*C*,*D*), (*A*,*B*,*D*,*C*,*F*,*E*), (*A*,*B*,*E*,*F*,*C*,*D*), and (*A*,*B*,*E*,*F*,*D*,*C*).

Not good for scheduling problems: 2-opt operator will generate a very large variation (weak locality)

Neighborhood for permutation scheduling problems

- Position-based

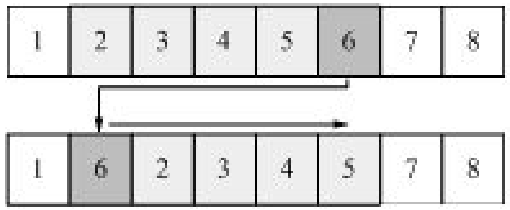


FIGURE 2.7 Insertion operator.

- Order-based

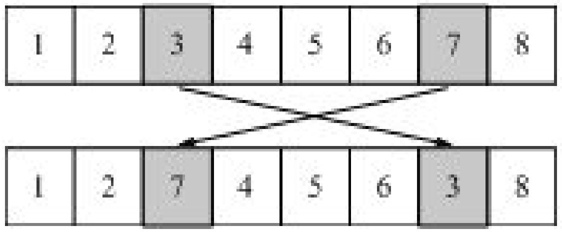


FIGURE 2.8 Exchange operator.

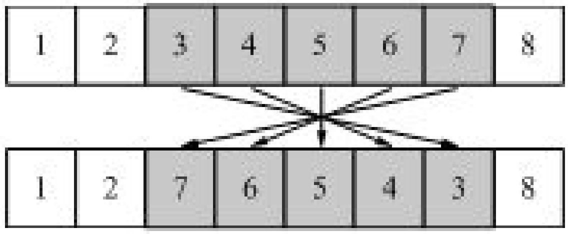


FIGURE 2.9 Inversion operator.

Very large neighborhoods

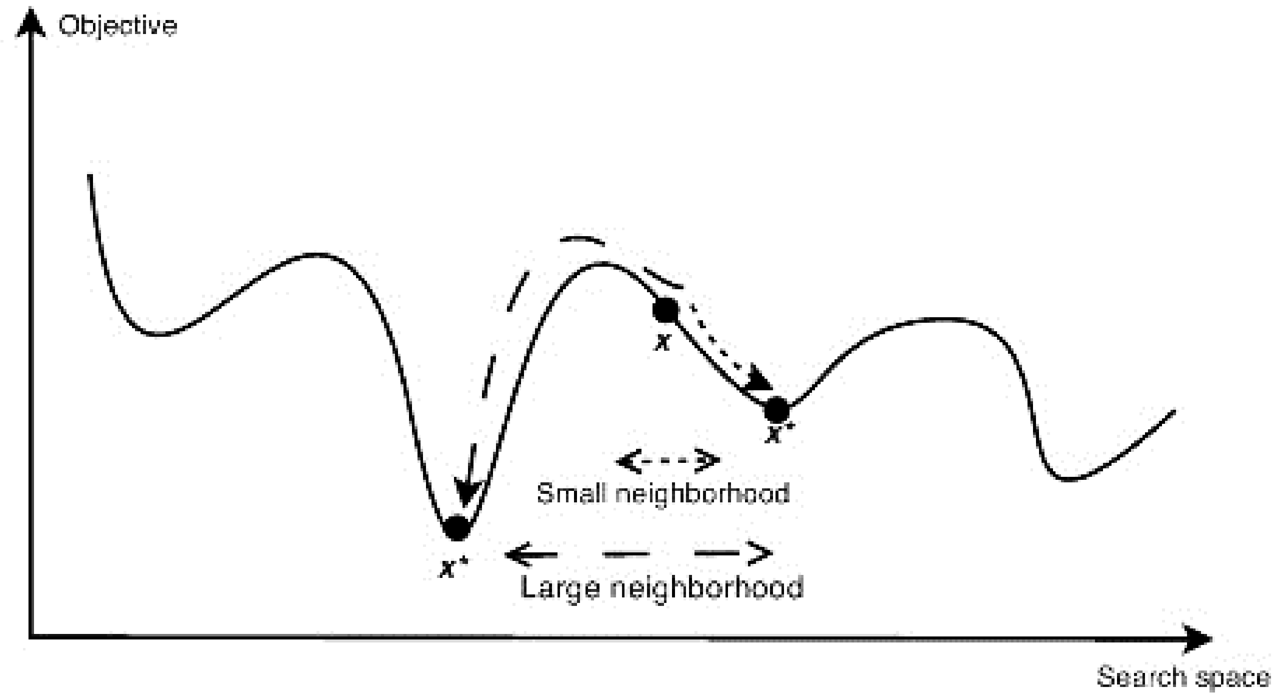
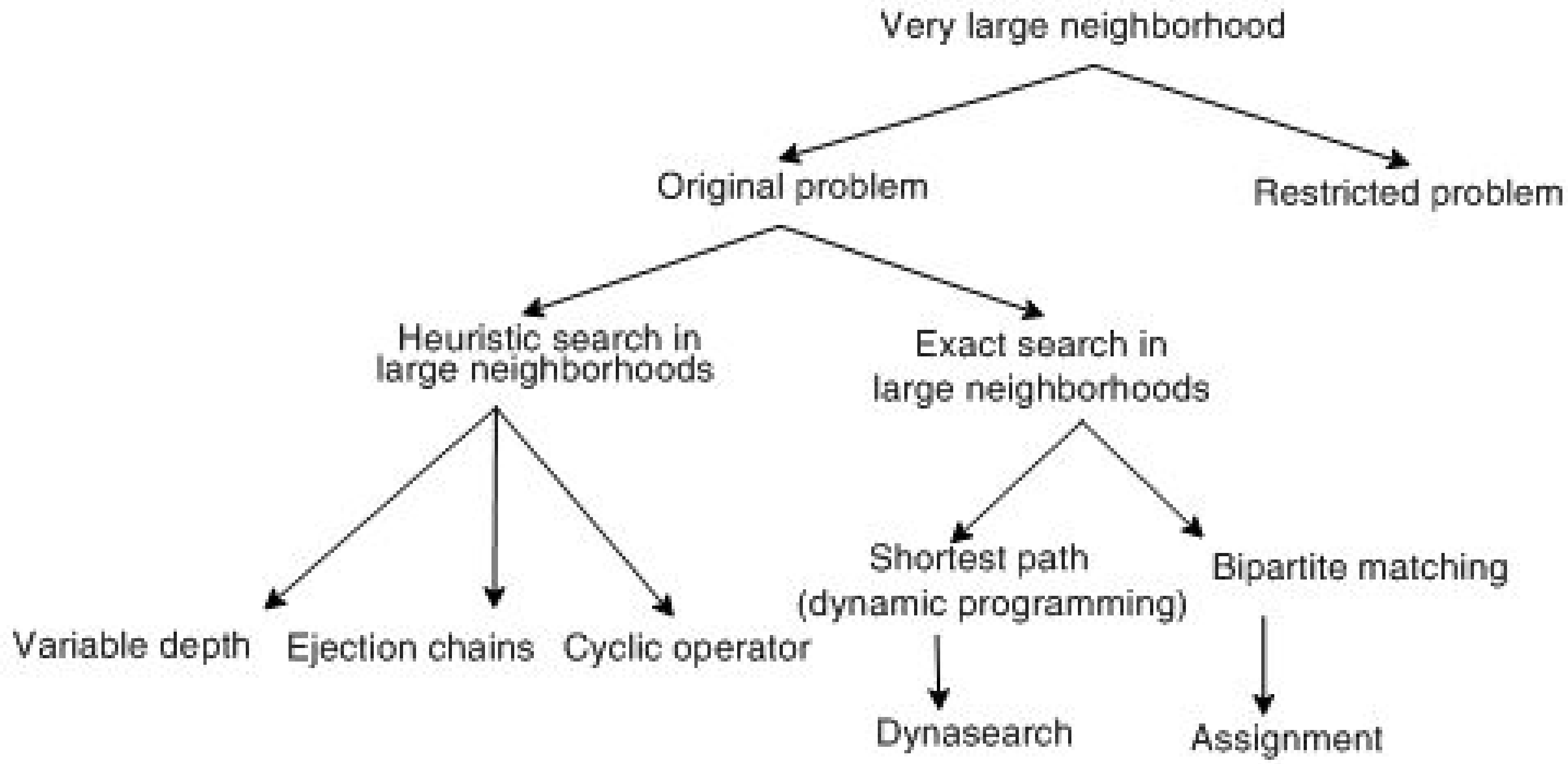


FIGURE 2.10 Impact of the size of the neighborhood in local search. Large neighborhoods improving the quality of the search with an expense of a higher computational time.

Very large neighborhood strategies



Heuristic search in large neighborhoods

- A partial set of the large neighborhood is generated → Finding the best neighbor is not guaranteed
- **Variable depth methods:** k-distance or k-exchange
- **Ejection chains:** a sequence of coordinated moves, alternating paths methods that is alternating sequence of addition and deletion
- **Cyclic exchange:** for partitioning problems

Ejection chain

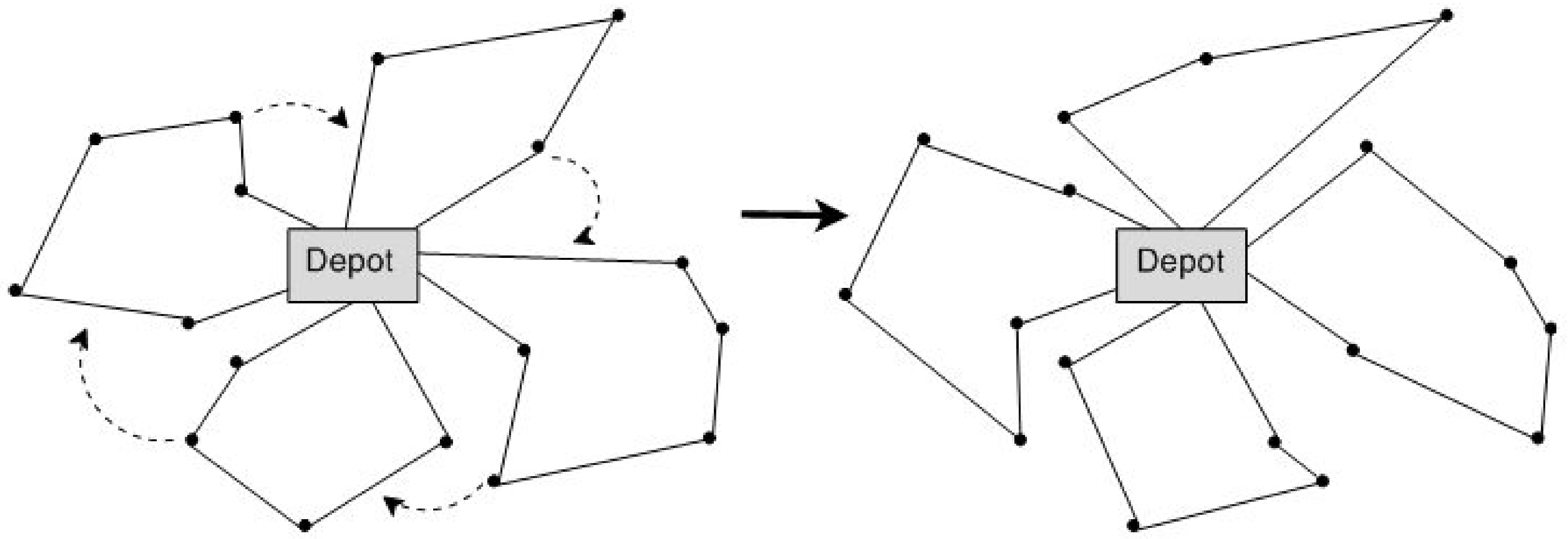


FIGURE 2.12 A four-level ejection chain for vehicle routing problems. Here, the ejection chain is based on a multinode insertion process.

Cyclic exchange

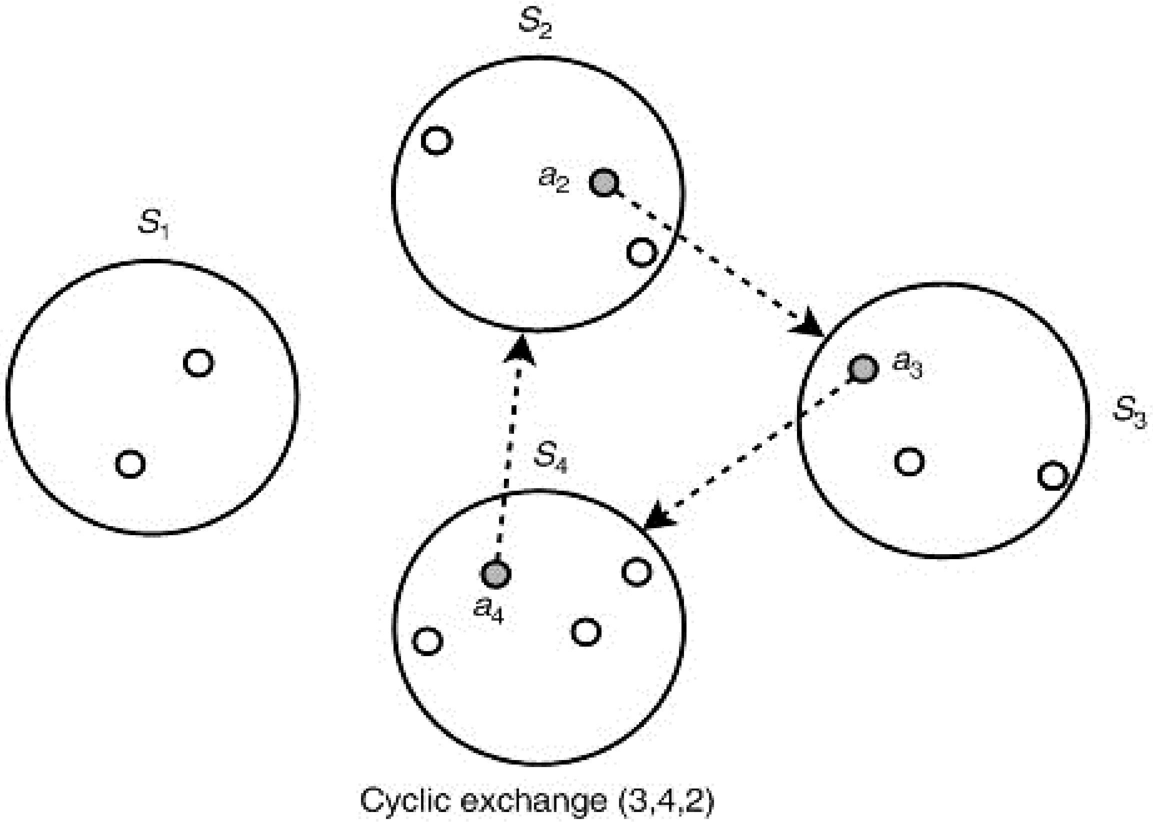
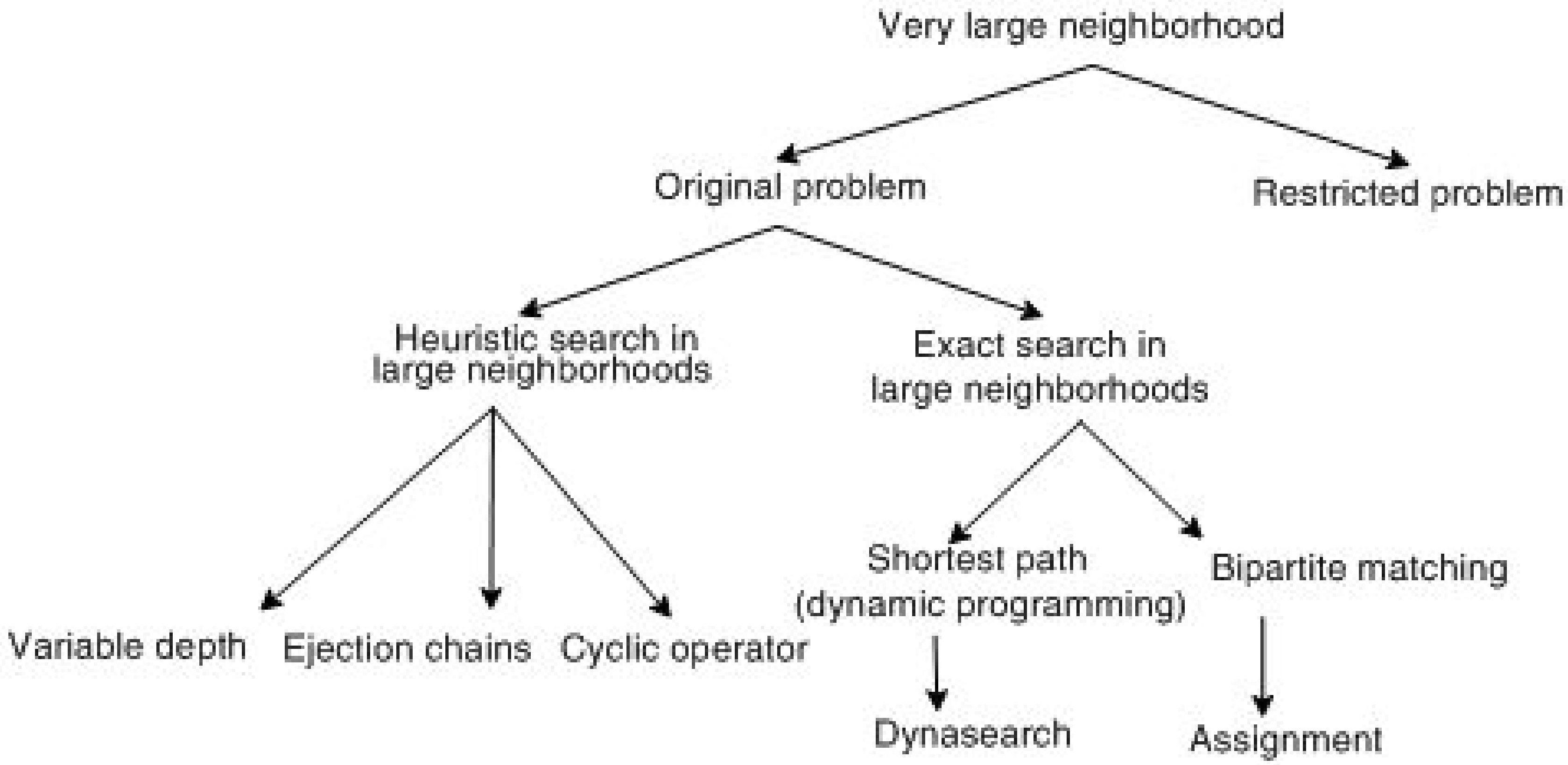


FIGURE 2.13 Very large neighborhood for partitioning problems: the cyclic exchange operator. Node a_2 is moved from subset S_2 to subset S_3, node a_3 is moved from subset S_3 to subset S_4, and node a_4 is moved from subset S_4 to subset S_2.

Very large neighborhood strategies



Exact search in large neighborhoods

- The main goal is to find an improving neighbor, search large neighborhood in a polynomial time
  - **Path finding:** shortest path and dynamic programming
  - **Matching:** well-known polynomial time matching
- **Definition 2.6 Independent swaps.** *Given a permutation* $\pi = \{\pi_1, \pi_2, \ldots, \pi_n\}$, *a swap move* (i, j) *consists in exchanging the two elements* π_i *and* π_j *of the permutation* π. *Two swap moves* (i, j) *and* (k, l) *are independent if* $(max\{i, j\} < min\{k, l\})$ *or* $(min\{i, j\} > max\{k, l\})$.

Exact search in large neighborhoods (cont.)

- **Dynasearch**
  - Polynomial exploration of exponentially large neighborhood
  - Where solutions are encoded by permutation
  - Two-exchange move, based on a Hamiltonian path between $\pi(1)\ and\ \pi(n)$

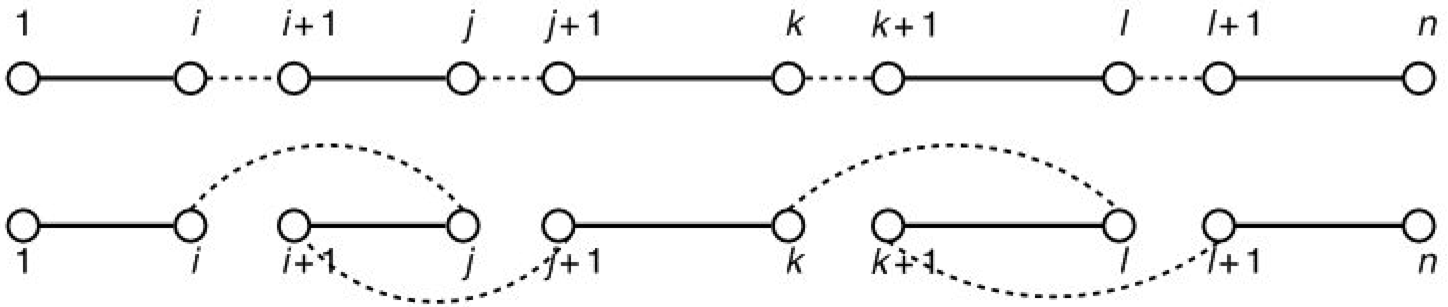


FIGURE 2.14 Dynasearch using two independent two-exchange moves: polynomial exploration of exponentially large neighborhoods.

Polynomial-specific neighborhood

- Some NP hard problems maybe solved in polynomial time for some restricted input instances
- Halin graph: TSP

If formed by embedding a tree having no degree-2 vertices and connecting it leaves by a cycle that crosses none of its edges

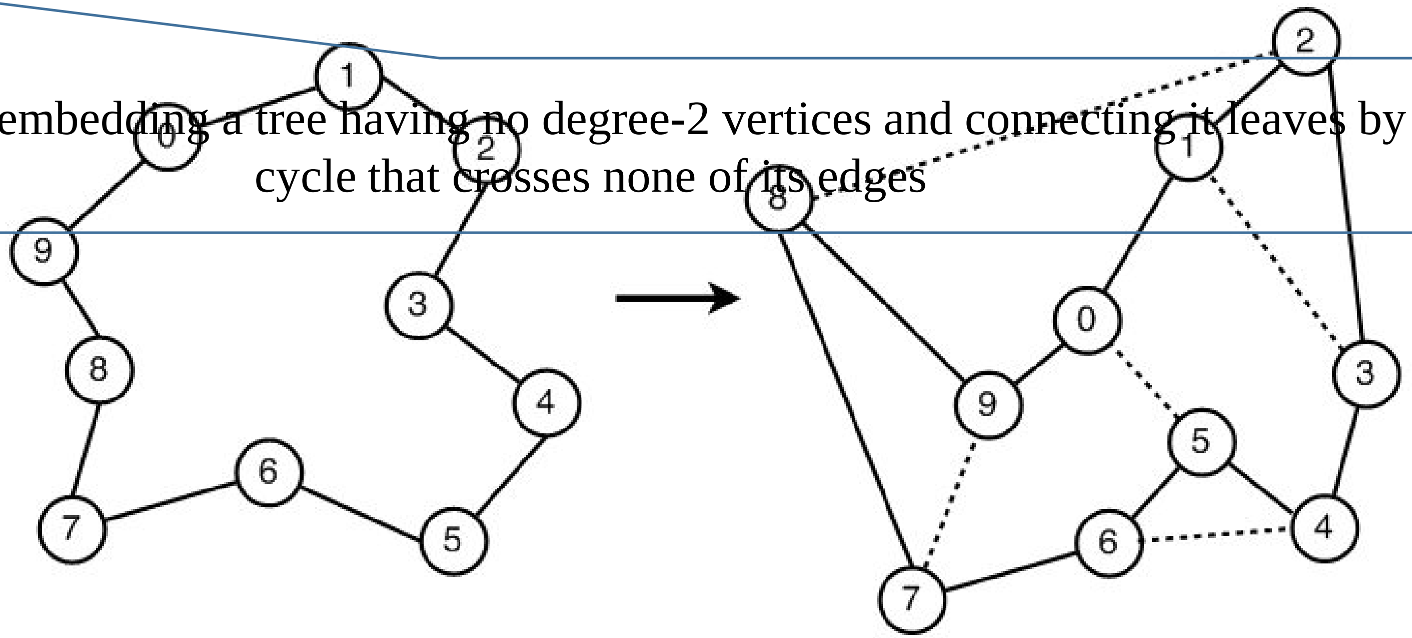


FIGURE 2.17 Extending a given input instance to a Halin graph.

Initial solution

- Strategies:
  - Random: quick but might take much larger number of iterations to converge
  - Greedy: faster but not always is better
  - Hybrid: combining both random and greedy approaches
- Trade off between quality of the solutions and computational time

Incremental Evaluation of the Neighborhood

- The evaluation of objective function is expensive

✖ A complete evaluation of the objective function

✓ Incremental evaluation: evaluation $\Delta(s, m)$ of the objective function (s: the current solution, m: the applied move)

$$f(s') = f(s \oplus m)$$

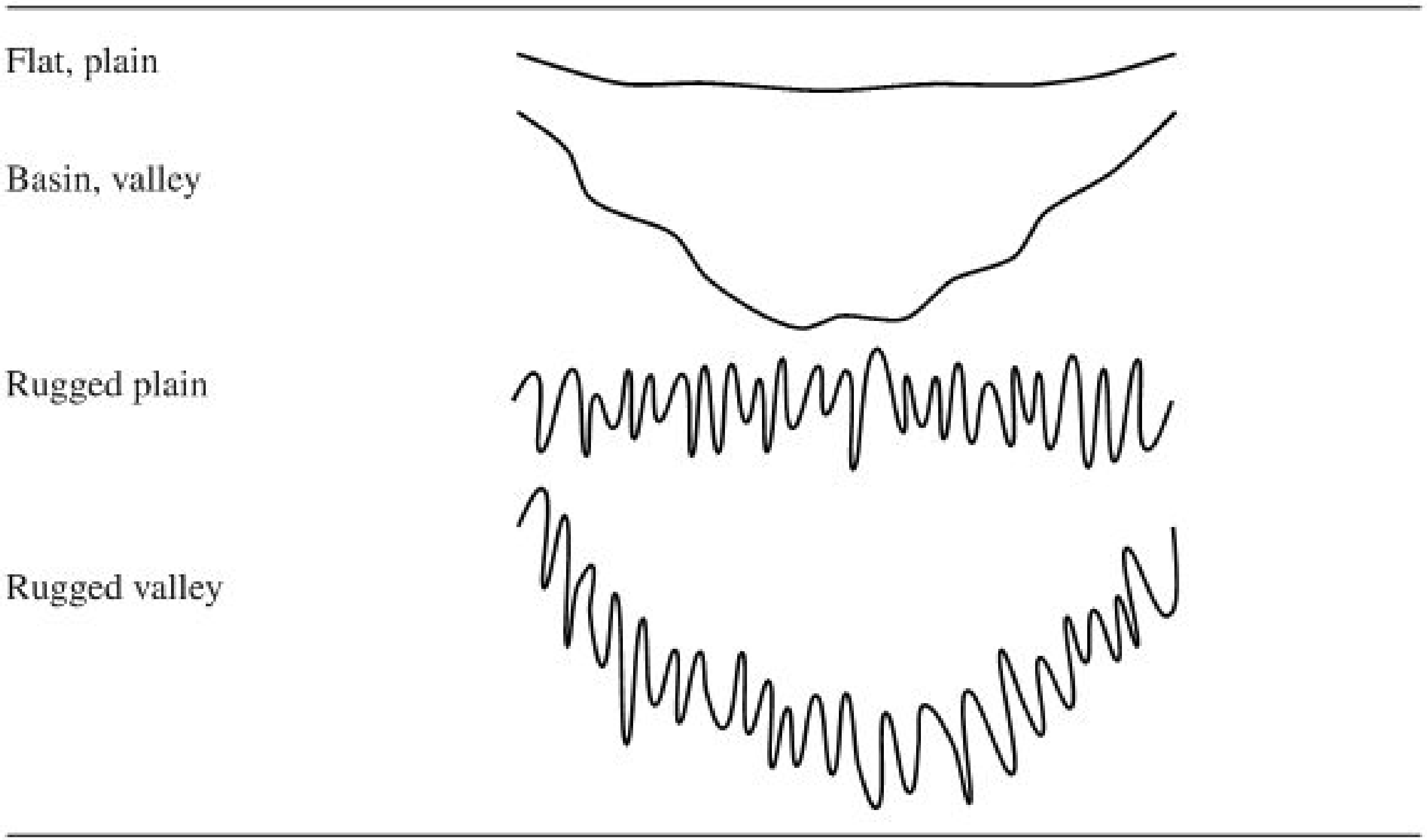
Fitness landscape analysis

- Superiority of algorithms: No algorithm is always the best
- Effectiveness of metaheuristics depends on:
  - Properties of the landscape(roughness, convexity,etc)
  - Instances to solve
- Landscape is defined by :
  - Representation
  - Neighborhood
  - Objective function
- Is performed in the hope to predict the behavior of different search components (representation, search operators, and objective function) of a metaheuristic

Definitions

- **Search space:** A directed graph *G = (S,E)*, where set of vertices S corresponds to the solutions of the problem, and E corresponds to the move operators
- **Fitness landscape:** The fitness landscape *l* may be defined by the tuple *(G,f)*, where f represents objective function that guides the search

Representation of landscape using the geographical metaphor



Conexity of the search space

- For any solutions s_i and s_j, there should be a path from s_i to s_j → Form any initial solution s_i there will a path to the S*

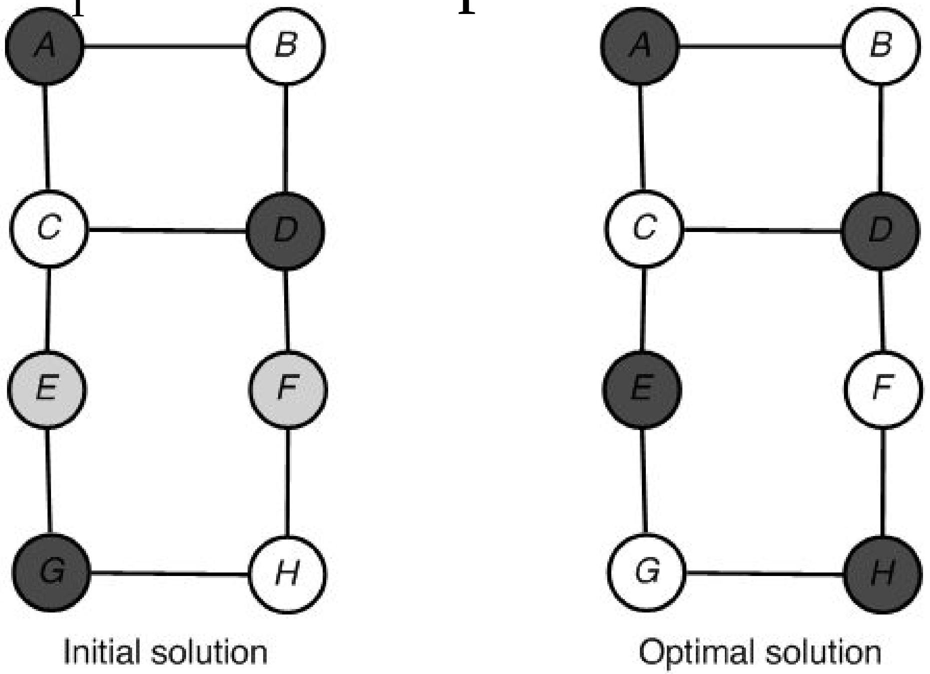


FIGURE 2.18 Connexity of the search space related to the graph coloring problem. The optimal solution cannot be reached from the given initial solution.

Distance is search space

- The minimum number of applications of the move operator to obtain solution S_j from solution S_i
- Properties: separative, symmetrical, triangular
- Distance in usual search spaces
  - Binary representations and flip move aperator (Hamming distance), Size of Search space = 2^n , Diameter = n
  - Permutation representations and the exchange move operator, Size of Search space = n! , Diameter = n-1
- Coherent Distance: must be related to the search operator

Landscape properties

- Landscape properties indicators
  - Global
  - Local
- Two different statistical measures:
  - Distribution measures : study the topology of local optima solutions
  - Correlation measures: analyze the rugosity of the landscape and the correlation between the quality of solutions and their relative distance

Distribution measures

- Objective: distribution analysis of the local optimal solutions in the landscape projected both in the search space G and in the objective space f
- Distribution indicators
  - Distribution in the search space
  - Entropy in the search space
  - Distribution in the objective space

Distribution in the search space

- For a population *P* of *S:*
  - Average distance

$$\mathrm{dmm}(P) = \frac{\sum_{s \in P} \sum_{t \in P, t \neq s} \mathrm{dist}(s, t)}{|P| \cdot (|P| - 1)}$$

  - Normalized average distance

$$\mathrm{Dmm}(P) = \frac{\mathrm{dmm}(P)}{\mathrm{diam}(S)}$$

  - Diameter of a population

$$\mathrm{diam}(P) = \max_{s,t \in P} \mathrm{dist}(s, t)$$

A weak distance: solutions belonging to the population P are clustered in a small region of the search spcae

Entropy

- To measure diversity of a given population in the search spcae
  - Different mathematical formulation
  - Weak: reveals a concentration of solutions
  - High: shows an important dispersion of the solution in the search space

Distribution in the objective space

- The amplitude of an arbitary population P of solutions is the relative difference between the best quality of the population P and the worst one:

$$\mathrm{Amp}(P) = \frac{|P| \cdot (\max_{s \in P} f(s) - \min_{s \in P} f(s))}{\sum_{s \in P} f(s)}$$

- Relative variation of Amp between a starting random population and the final population:

$$\Delta_{\mathrm{Amp}} = \frac{\mathrm{Amp}(U) - \mathrm{Amp}(O)}{\mathrm{Amp}(U)}$$

- The average gap of the relative gaps between the cost of the population of the local optima and the global optima solutions:

$$\mathrm{Gap}(O) = \frac{\sum_{s \in O} (f(s) - f(s^*))}{|O| \cdot f(s^*)}$$

Correlation measures

- Objective: estimate the ruggedness of the landscape along with the correlation between the quality of solutions and their distance to a global optimal solution
- Correlation indicators
  - Length of the walks
  - Autocorrelation function
  - Fitness distance correlation
  - Deception
  - Epistasis
  - Multimodality
  - Neutrality
  - Fractal

Length of the walks

$$\mathrm{Lmm}(P) = \frac{\sum_{p \in P} l(p)}{|P|}$$

l(P): The length of the walk starting with the solution $p \in P$

- Information about ruggedness
- More number of optima and short walks: rugged
- Few number of optima and long walks: smooth

Autocorrelation function

- Measures the ruggedness
- Correlation of solutions in the search space with distance d

$$\rho(d) = \frac{\sum_{s,t \in S \times S, \text{dist}(s,t)=d} (f(s) - \bar{f})(f(t) - \bar{f})}{n.\sigma_f^2}$$

P(1) considers only neighboring solutions

- A low value: the variation of fitness between two neighbors is equal on average to the variation between any two solutions and the landscape is rugged

Autocorrelation function(cont.)

- Random walk:

$$r(s) \approx \frac{1}{\sigma_f^2(m-s)} \sum_{t=1}^{m-s} (f(x_t) - \bar{f})(f(x_{t+s}) - \bar{f})$$

m: size of random walk
s: distance between solutions

- Correlation length

$$l = \frac{1}{\ln(|r(1)|)} = -\frac{1}{\ln(|\rho(1)|)}$$

- The smaller is the correlation length, the more rugged is the associated landscape and harder is the search.

Fitness distance correlation

- Measures how much the fitness of a solution correlates with the distance to the global optimum

$$F = \{f_1, f_2, \dots, f_n\} \qquad D = \{d_1, d_2, \dots, d_n\}$$

$$r = \frac{\mathrm{cov}(F, D)}{\sigma_f \sigma_d} \qquad \mathrm{cov}(F, D) = \frac{1}{n}\sum_{i=1}^{n}(f_i - \bar{f})(d_i - \bar{d})$$

Fitness distance correlation (cont.)

- FCD
  - Straightforward
    - large positive
    - easy to solve
    - as the fitness decreases, the distance to the global optimum also decreases
  - Misleading
    - Large negative
    - The move operator will guide the search away from the global optimum
  - Difficult
    - Near-zero
    - There is no correlation between fitness and distance

Fitness distance correlation (cont.)

- Fitness distance plot: fitness of the solutions against their distance to the global optima

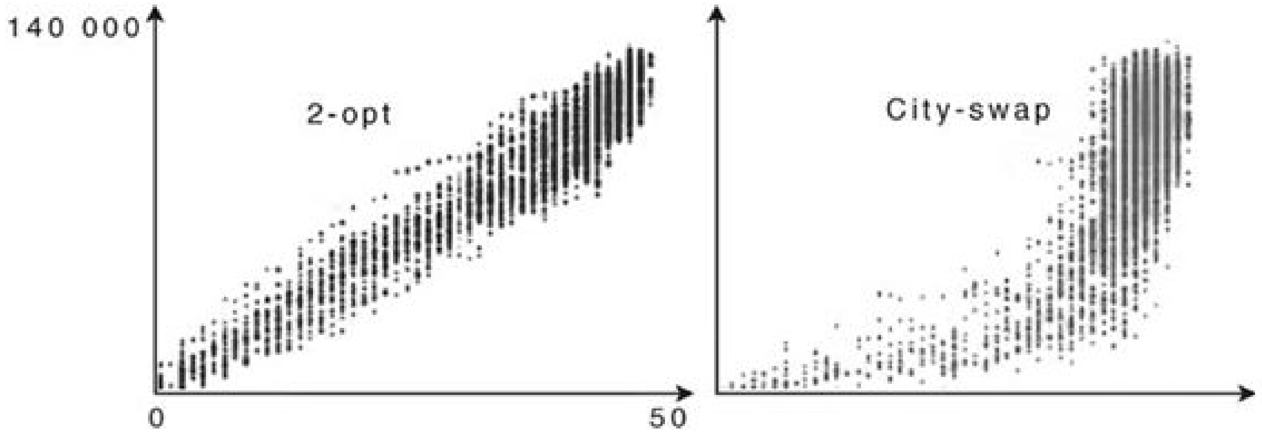


FIGURE 2.19 Using the FDC analysis, the figure shows the fitness distance plot of the instance *att48* of the TSP using the 2-opt and the city-swap neighborhood structures. The left figure for the 2-opt shows a high FDC (0.94) and the right figure shows a less important FDC for the city-swap operator (0.86). For the problem instance *tsp225*, the FDC is 0.99 for the

Breaking plateaus in a flat landscape

Definition 2.10 Plateau. *Given a point s in the search space S and a v value taken in the range of values of the criterion f. Given $N(s)$, the set of points s' in the neighborhood of the solution s. Considering X a subset of $N(s)$ defined by $s'' \in X$ iff $f(s'') = v$, X is a plateau iff it contains at least two elements (i.e., $|X| \geq 2$).*

Breaking plateaus in a flat landscape (cont.)

- A metaheuristic has difficulties to be guided in the neighborhood of the current solution
- Changing objective function (embedding more information to have a significant improvement in quality of the related solutions)
  - Discrimination criterion f', can discriminate points that have the same value for the main criterion f
  - Solutions with the same value in the objective space but with different value with regarded to decision space

$$f''(x) = k_1 \times f(x) + f'(x)$$