N. M. M. de Abreu
T. M. Querido
P. O. Boaventura-Netto

RedInv-SA: a simulated annealing for the quadratic assignment problem

RAIRO. Recherche opérationnelle, tome 33, no 3 (1999), p. 249-273

<http://www.numdam.org/item?id=RO_1999__33_3_249_0>

© AFCET, 1999, tous droits réservés.

L’accès aux archives de la revue « RAIRO. Recherche opérationnelle » implique l’accord avec les conditions générales d’utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d’une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

Numdam

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/
REDINV-SA: A SIMULATED ANNEALING FOR THE QUADRATIC ASSIGNMENT PROBLEM (*)

by N.M.M. de ABREU (1), T.M. QUERIDO (2) and P.O. BOAVENTURA-NETTO (1)

Communicated by Catherine ROUCAROL

Abstract. - An algebraic and combinatorial approach to the study of the Quadratic Assignment Problem produced theoretical results that can be applied to (meta) heuristics to give them information about the problem structure, allowing the construction of algorithms. In this paper those results were applied to inform a Simulated Annealing-type heuristic (which we called RedInv-SA). Some results from tests with known literature instances are presented.

Keywords: Quadratic assignment problem, meta-heuristics, simulated annealing, and combinatorial programming.

Résumé. - Une étude algébrique et combinatoire du Problème d'Affectation Quadratique a produit des résultats théoriques qui peuvent être appliqués à des (méta)-heuristiques. Donnant des informations sur la structure du problème, ils conduisent de cette façon à la formulation d'algorithmes. Dans cet article, ces résultats ont été appliqués à l'information d'une heuristique de type Recuit Simulé, appelée RedInv-SA. Des résultats de l'application de l'algorithme à des problèmes classiques de la littérature sont présentés à la fin de l'article.

Mots clés : Problèmes d'affectation quadratique, métheuristiques, recuit simulé et programmation combinatoire.

1. INTRODUCTION

We consider the Quadratic Assignment Problem (QAP) in the classical form defined by Koopmans and Beckmann (KB57) as that of minimizing a sum of flow-distance products, which is the objective function of an allocation problem of n “machines” to n “positions” with pairwise cost dependency. This description has been applied to a variety of practical situations, from hospital plants through university campi to integrated circuits and keyboards.

(*) Received January 1996.
(1) COPPE/UFRJ, CP 68.507, 21.945-970 Rio de Janeiro, RJ, Brazil.
(2) CEFET/RJ, Av. Maracanã n° 229, 20.271-110 Rio de Janeiro, RJ, Brazil.
The purposes of this paper are to present the essential features of an algebraic-combinatorial approach for QAP study and to apply some of these features in the building of a Simulated Annealing (SA) heuristic algorithm, as well as to allow for a solution quality evaluation. The theoretical basis, following Abreu (Ab84) involves a number of results concerning the permutation set. Among the results there presented it is important, for what follows, a condition for partial ordering of the permutations corresponding to the solutions of QAP instances and the connection between the permutation inversion number and the solution costs.

The K-B formulation is that of equations (1.1).

$$\min Z = \sum_{i,j,p,q=1}^{n} f_{ij}d_{pq}x_{ip}x_{jq}$$

subject to

$$\sum_{i=1}^{n} x_{ij} = 1 \quad (j = 1, \ldots, n) \quad (1.1)$$

$$\sum_{j=1}^{n} x_{ij} = 1 \quad (i = 1, \ldots, n)$$

$$x_{ij} = 0 \text{ or } 1$$

where the $f_{ij}$ and the $d_{pq}$ are non-negative integers defining matrices $F$ and $D$ respectively. In this work we will consider these matrices as being symmetric, which allows us to substitute the Gavett-Plyter vector description (GP66) for them.

The QAP has some amazing properties, as it is extremely well structured and, on the other hand, it is a NP-hard problem for whose resolution the present tendency is the design of efficient heuristics. Such algorithms require instance banks for comparative testing and, in fact, a QAP instance library, QAPLIB [BKR94] has been constantly utilized and amplified. Techniques for generating instances with known optima have also been developed by Palubetskis [Pa88] and Li and Pardalos [LP92], allowing researchers to design their own test instances. An important reference on QAP is [PW94ed] where, among several important articles, a good review by the same authors [PW94] can be found.
2. SOME CONCEPTS AND DEFINITIONS

The graph theory basic concepts and definitions used in this work follow Berge [Be73].

The Koopmans-Beckmann description is equivalent to the Burkard and Stratman form [BS78], considering the assignment of facilities to locations as a permutation $\varphi$ on $\{1,2,\ldots,n\}$ by setting $\varphi(i) = j$ if facility $i$ is assigned to location $j$, such that

$$\min Z = \sum_{i,j=1}^{n} f_{ij}d_{\varphi(i)\varphi(j)}.$$  \hfill (2.1)

At this point it is convenient to recall some basic definitions.

**Definition 2.1:** A *permutation* (or $t$-permutation) is a bijective function $\xi : T \to T$ (where $T$ is the set of the first $t$ naturals) represented as

$$\xi = \begin{pmatrix} 1 & 2 & \ldots & t \\ \xi(1) & \xi(2) & \ldots & \xi(t) \end{pmatrix}.$$

The natural $t$ is known as the *order* of the permutation.

Let $S_t$ be the set of all permutations over the elements in $T$. The *identity permutation* is the function $\iota_t \in S_t$ such that $\iota_t(j) = j$, for $j = 1, \ldots, t$. Two elements $x, y \in T$ are *adjacent* in $\xi \in S_t$ when there exists $i \in T - \{t\}$ such that $x = \xi(i)$ and $y = \xi(i + 1)$ or $y = \xi(i)$ and $x = \xi(i + 1)$. We say that $x$ precedes $y$ in the image of $\xi$ when there exists $i, j \in T$ such that $i < j$, $x = \xi(i)$ and $y = \xi(j)$. An ordered pair $(x,y) \in T \times T$ such that $x$ precedes $y$ in the image of $\xi$ is an *inversion* in $\xi$ when $x > y$.

It is frequently useful to shorthand a permutation by its image $\xi = (\xi(1)\xi(2)\ldots\xi(t))$.

**Definition 2.2:** Let $\xi$ be a permutation. Then

$$\vartheta(\xi) = \{(x,y) \in T \times T | (x,y) \text{ is an inversion in } \xi \}$$

where $\vartheta(\xi)$ is the *inversion set* in $\xi$. The cardinality $|\vartheta(\xi)|$ is the number of inversions in $\xi$.

**Definition 2.3:** A *cycle* ($k$-cycle) in $S_t$ is a $t$-permutation $\mu$ with $(t - k)$ *fixed elements* $\mu(j) = j$, the remaining $k$ elements being given...
by $\mu(x_i) = x_{i+1}$ ($i = 1, \ldots, k-1$), $\mu(x_k) = x_1$. A cycle is usually represented as $(x_1, \ldots, x_k)$. A 2-cycle is also called a transposition.

**Definition 2.4:** A QAP instance of order $n$ is the association of two $n$-cliques $K_f$ and $K_d$, respectively valued by the material flows between machines and by the distances between positions (Fig. 2.1).

![Fig. 2.1. - Cliques $K_f$ and $K_d$.](image)

This example is the Gavett-Plyter instance [GP66]. These authors developed a QAP relaxation (which is a Linear Assignment Problem of size $N = C_{n,2}$) defined by two column vectors, $F = [f_r]$ and $D = [d_s]$ ($r, s = 1, \ldots, N$) where the components are respectively the edges of $K_f$ and $K_d$ taken in lexicographic order.

**Definition 2.5:** We define a flow-distance edge-edge matrix $\Gamma$, as

$$\Gamma = F D^T = [\gamma_{rs}].$$

(2.4)

Every QAP feasible solution is a solution for the relaxed problem defined by $\Gamma$. The reciprocal, however, is not true, because the relaxed problem has $N!$ solutions, of which only $n!$ are feasible.

The original and the relaxed problems can be associated to each other by the bijection between lexicographic and numerical orders; for a pair $(i, j)$ ($1 \leq i < j \leq n$) the numerical ordering $\psi_{ij}$ is

$$\psi_{ij} = (i - 1)n - i(i + 1)/2 + j.$$  

(2.5)

**Definition 2.6:** By sorting $F$ components in non-decreasing order and $D$ components in non-increasing order and calling $F$ and $D$ the vectors thus obtained, we define a standard matrix $\overline{\Gamma}$ as

$$\overline{\Gamma} = \overline{F} \overline{D}^T.$$  

(2.6)
and $\phi_F$ and $\phi_D$ as the permutations taking respectively $F$ into $\overline{F}$ and $D$ into $\overline{D}$.

Example matrices are shown in Figures 2.2 and 2.3, where lines and columns are identified by both orderings and by the edge values.

$$
\begin{array}{ccc|ccc}
\Psi(i,j) & i,j & F & \Psi(i,j) & i,j & D \\
1 & 12 & 28 & 1 & 2 & 3 & 4 & 5 & 6 \\
& 12 & 13 & 14 & 23 & 24 & 34 \\
6 & 7 & 2 & 5 & 6 & 1 & 168 & 196 & 56 & 140 & 168 & 28 \\
2 & 13 & 25 & 150 & 175 & 50 & 125 & 150 & 25 \\
3 & 14 & 13 & 78 & 91 & 26 & 65 & 78 & 13 \\
4 & 23 & 15 & 90 & 105 & 30 & 75 & 90 & 15 \\
5 & 24 & 4 & 24 & 28 & 8 & 20 & 24 & 4 \\
6 & 34 & 23 & 138 & 161 & 46 & 115 & 138 & 23 \\
\end{array}
$$

**Fig. 2.2.** - $\Gamma$ matrix - [GP66] instance.

$$
\begin{array}{cc|ccc|ccccc}
\phi_F^{-1} & \psi^{-1} & \overline{F} & \phi_D^{-1} & \psi^{-1} & \overline{D} \\
5 & 24 & 4 & 28 & 24 & 24 & 20 & 8 & 4 \\
3 & 14 & 13 & 91 & 78 & 78 & 65 & 26 & 13 \\
4 & 23 & 15 & 105 & 90 & 90 & 75 & 30 & 15 \\
6 & 34 & 23 & 161 & 138 & 138 & 115 & 46 & 23 \\
2 & 13 & 25 & 175 & 150 & 150 & 125 & 50 & 25 \\
1 & 12 & 28 & 196 & 168 & 168 & 140 & 56 & 28 \\
\end{array}
$$

**Fig. 2.3.** - Standard $\Gamma$ matrix for the [GP66] instance.

The feasible solution corresponding to [2 1 4 3] in the graphs of Figure 2.1, is shown in both matrices as underlined characters (with a cost equal to 464).

As a consequence of $\overline{\Gamma}$ ordered structure we can obtain a lower bound $\text{tr}(\overline{\Gamma})$ for the optimal QAP solution; Hardy *et al.* [HLP52] and Wimmert [Wi58] show it is as being equal to the trace of $\overline{\Gamma}$, that is, the sum of its diagonal elements. (For this instance we have $\text{tr}(\overline{\Gamma}) = 389$, the optimal solution being [4 1 3 2] with cost 403.)

**Definition 2.7:** To a given matrix $\overline{\Gamma}$ we associate a set $QA(\overline{\Gamma})$ of all instances $I_{\overline{\Gamma}}$ which can be obtained by different $\overline{F}$ and $\overline{D}$ orderings. Each of these instances will be defined by a given $(F, D)$ pair.
3. THE FEASIBILITY PROBLEM

**DEFINITION 3.1:** Let us now consider a given matrix $\Gamma$ and a given $N$-permutation $\rho_\Gamma$. We say that the $N$-vector $X = (f_1 d_{\rho_\Gamma(1)}, \ldots, f_N d_{\rho_\Gamma(N)})$ is *feasible* for a given instance $P \in QA(\Gamma)$ if we can find a vertex permutation $\varphi$ such that

$$f_r d_{\rho_\Gamma(r)} = f_{\varphi^{-1}(r)} d_{\varphi^{-1}(\rho_\Gamma(r))} = f_{\Psi_{i_j}} d_{\Psi_{\varphi(i_j)\varphi(j)}} \quad (r = 1, \ldots, N). \quad (3.1)$$

The feasibility is guaranteed by the coherent application of $K_d$ edges over $K_f$ edges, that is, preserving vertex over vertex.

For a given $\rho_{\Gamma} \in QA(\Gamma)$ we consider the permutations $\rho_\Gamma$ and $\rho_{\Gamma}$ as acting respectively on $\Gamma$ and $\Gamma$ columns. It is easy to see that

$$\rho_{\Gamma} = \phi_F \circ \rho_{\Gamma} \circ \phi_D^{-1}. \quad (3.2)$$

In order to clarify the composition operation shown in the example that follows, we will use here the notation presented in (2.2). This composition is done from left to right.

**Example:** We take again the [GP66] instance, with the solution defined in $\Gamma$ as $\rho_\Gamma = [1 \ 5 \ 4 \ 3 \ 2 \ 6]$ and in $\Gamma$ as $\rho_{\Gamma} = [1 \ 4 \ 5 \ 6 \ 3 \ 2]$.

$$\rho_{\Gamma} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 5 & 4 & 3 & 2 & 6 \end{pmatrix}.$$  

Then

$$\phi_F \circ \rho_{\Gamma} \circ \phi_D^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 2 & 3 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 5 & 6 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 5 & 4 & 3 & 6 \end{pmatrix} = \rho_{\Gamma}.$$  

To identify whether an edge-permutation $\rho_\Gamma$ corresponds or not to a feasible solution for a given instance $\Gamma$, we have to verify if there is a vertex-permutation $\varphi \in S_n$ leading to an edge allocation in $\Gamma$ corresponding to

$$\rho_{\Gamma} = \begin{pmatrix} \Psi(1,2) = 1 & \Psi(1,3) = 2 & \ldots & \Psi(n-1,n) = N \\ \Psi(\varphi(1),\varphi(2)) & \Psi(\varphi(1),\varphi(3)) & \ldots & \Psi(\varphi(n-1),\varphi(n)) \end{pmatrix}.$$  

Recherche opérationnelle/Operations Research
For the [GP66] instance we have, as an example (see Figs. 2.2 and 2.3),

<table>
<thead>
<tr>
<th>Matrix</th>
<th>( \bar{\Gamma} )</th>
<th>( \Gamma )</th>
<th>( \bar{\Gamma} )</th>
<th>( \Gamma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>(6,2)</td>
<td>(1,1)</td>
<td>(5,3)</td>
<td>(2,5)</td>
</tr>
<tr>
<td>Entry value</td>
<td>168</td>
<td>150.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. INVERSIONS AND COSTS

Through this item we will consider every \( N \)-permutation as a permutation of \( \bar{\Gamma} \) columns; for sake of simplicity we will not use the \( \bar{\Gamma} \) index.

The translation of (1.1) into edge permutations associated to permutations of \( \bar{\Gamma} \) columns allows us to write, for a given \( N \)-permutation \( \xi \),

\[
Z_\xi = \sum_{i=1}^{N} f_i^* d_{\xi(i)} \quad (4.1)
\]

**Definition 4.1:** The *increase* associated to a permutation \( \xi \in S_N \) is

\[
\Delta_\xi = Z_\xi - Z_I \quad (4.2)
\]

where \( Z_I = \text{tr}(\bar{\Gamma}) \) is the problem lower bound associated to the identity permutation \( \iota_N \).

**Theorem 4.1:** The increase associated to a cycle \( \mu = (x_1, \ldots, x_k) \in S_N \) is given by

\[
\Delta(\mu) = \sum_{j=1}^{k-1} \bar{f}(x_{j+1})d(x_j) - \bar{d}(x_j)) + f(x_k)d(x_1) - \bar{d}(x_k)).
\]

**Proof:** By considering (4.1) and (4.2). \( \square \)

We know that every permutation can be expressed as a composition of disjoint cycles (see for instance Berge [Be68]). The next theorem allows the increase of a permutation to be defined as a function of the increases associated to its cycles.

vol. 33, n° 3, 1999
THEOREM 4.2: Let $\xi \in S_N$ be a permutation expressed as a composition of $r$ disjoint cycles $\xi = \mu_1 \circ \mu_2 \circ \ldots \circ \mu_r$. The increase associated to $\xi$ is

$$\Delta \xi = \sum_{j=1}^{r} \Delta(\mu_j).$$

Proof: As the cycles are disjoint, we can write

$$Z_\xi = \sum_{j=1}^{r} \sum_{y \in \mu_j} \bar{f}(y)\bar{d}(\mu_j(y))$$

and therefore

$$\Delta \xi = Z_\xi - Z_I = \sum_{j=1}^{r} \left[ \sum_{y \in \mu_j} \bar{f}(y)[\bar{d}(\mu_j(y)) - \bar{d}(y)] \right] = \sum_{j=1}^{r} \Delta(\mu_j). \quad (4.4)$$

This rather simple result has a strong consequence; as every parcel in the sum is non-negative, we can say that, for every $j$, $j = 1, \ldots, r$,

$$\Delta(\mu_j) \leq \Delta(\xi). \quad (4.5)$$

Then we can have permutation pairs $(\xi, \pi)$ such that $\Delta \xi \leq \Delta \pi$ or $\Delta \xi \geq \Delta \pi$, regardless of the instance values.

DEFINITION 4.2: Two increases $\Delta \xi_1$ and $\Delta \xi_2$ are freely comparable if it is possible to verify either $\Delta \xi_1 \leq \Delta \xi_2$ or $\Delta \xi_1 \geq \Delta \xi_2$, regardless of $\Gamma$ entry values.

THEOREM 4.3: If a permutation $\xi_2$ is obtained from a permutation $\xi_1$ by applying a transposition $\theta$, then $\xi_1$ and $\xi_2$ are freely comparable.

Proof: Let us consider $\theta = (x, y)$, $x < y$ and $\xi_2 = \xi_1 \circ \theta$. We will have $\xi_2(x) = \xi_1(y)$, $\xi_2(y) = \xi_1(x)$ and $\xi_2(i) = \xi_1(i)$ for $i \neq x, y$. Then

$$\Delta(\xi_2) - \Delta(\xi_1) = (\bar{f}(x) - \bar{f}(y))[\bar{d}(\xi_1(y)) - \bar{d}(\xi_1(x))]. \quad (4.6)$$

As $\bar{F}$ has a non-descending ordering the first factor of the right-hand member will always be non-positive. For the same ordering reason, the second factor will always be non-negative if $\xi_1(y) < \xi_1(x)$ and non-positive if $\xi_1(y) > \xi_1(x)$.
According to this proof the comparison direction (either $\Delta(\xi_2) \leq \Delta(\xi_1)$ or $\Delta(\xi_2) \geq \Delta(\xi_1)$) depends on whether $\xi_1(x)$ is lesser or greater than $\xi_1(y)$. This comparison corresponds to an order of scalar products, an approach used by [Re85]. The next theorem relates this result to the inversion number of $\xi_1$ and $\xi_2$.

**Theorem 4.4:** Let $\xi_1, \xi_2, \theta \in S_N$, where $\theta$ is a transposition. If $\xi_2 = \xi_1 \circ \theta$ and $|\vartheta(\xi_1)| < |\vartheta(\xi_2)|$, then $\xi_1$ and $\xi_2$ are freely comparable and $Z_{\xi_1} \leq Z_{\xi_2}$.

**Proof:** Consider $\theta = (x, y)$ where $1 \leq x < y \leq N$. Then the permutations $\xi_1$ and $\xi_2$ are respectively of the form

$$\xi_1 = \left(\ldots, x, \ldots, y, \ldots\right) \quad \text{and} \quad \xi_2 = \left(\ldots, \xi_1(x), \ldots, \xi_1(y), \ldots\right).$$

Taking into account that $|\vartheta(\xi_1)| < |\vartheta(\xi_2)|$ implies $\xi_1(x) < \xi_1(y)$ we have, from (4.6),

$$\Delta_{\xi_1} - \Delta_{\xi_2} = (\overline{f}(x) - \overline{f}(y))[\overline{d}(\xi_1(y)) - \overline{d}(\xi_1(x))].$$

The orderings defined for $\overline{F}$ and $\overline{D}$ imply respectively $\overline{f}(x) \leq \overline{f}(y)$ and $\overline{d}(\xi_1(y)) - \overline{d}(\xi_1(x))$, from which we obtain $\Delta_{\xi_1} \leq \Delta_{\xi_2}$ and, consequently,

$$Z_{\xi_1} \leq Z_{\xi_2}.$$ 

**Definition 4.5:** Theorem 4.4 allows us to build a directed graph $G_{IN} = (S_N, W)$ where $W$ is the set of pairs $(\xi_i, \xi_j)$ such that $\vartheta(\xi_i)$ has exactly one more inversion than $\vartheta(\xi_j)$. This graph is known as the inversion graph.

$G_{IN}$ is a circuitless graph whose level ordinal function is the number of inversions. Level 0 has exactly one element (which is $\vartheta_N$, the identity permutation corresponding to the lower bound $\text{tr}(\overline{F})$). It has a lattice structure [Be68, AV95].

Figure 4.1 is a scheme of $G_{IN}$ vertex distribution for $N = 6$; the graph is regular of degree $N - 1$, so the single vertex in Level 0 is adjacent to every vertex in Level 1; from now on, the adjacency relations depend on which positions of each $N$-permutation are affected by the possible inversions. For $N = 6$ we have $C_{6,2} + 1 = 16$ levels and their cardinalities $N_i$ can
be easily found, through a generating function, to be 1, 5, 14, 29, ..., 29, 14, 5, 1 [Ab84].

Figure 4.2 shows the first three levels of GI_6 with the corresponding permutations. We consider the edges to be directed from Level \( k + 1 \) to Level \( k \), according to Définition 4.5.

**Theorem 4.5:** If there exists a directed path between \( \xi_i \) and \( \xi_j \) in GI_N, then \( \Delta(\xi_i) \) and \( \Delta(\xi_j) \) are freely comparable, with \( |\vartheta(\xi_i)| < |\vartheta(\xi_j)| \) and consequently \( Z_{\xi_i} \leq Z_{\xi_j} \).

**Proof:** It follows from Définition 4.5 and Theorem 4.4.

Theorem 4.5 allows us to identify a class of freely comparable permutation pairs where the respective number of inversions grow with the increases and consequently with the costs.

The freely comparable increases are studied in more detail in [Ab84, AbBo89]. One of the most important results there presented is a Maximal Partial Ordering Theorem (MPOT), which gives necessary and sufficient conditions for free comparability between permutations.

A result equivalent to Theorem 4.5 is not available for all MPOT cases. To discuss the application of this theory with regard to heuristic applications we present a conjecture (the Inversion Conjecture) [Ab84, AB89] which
is reinforced, from a practical point of view, by Figures 6.1a and b and 6.2a and b.

**Conjecture 4.1:** The cost function within freely comparable solutions increases with their number of inversions in the corresponding permutations on \( \vec{F} \) columns.

5. THE REDINV-SA HEURISTIC

### 5.1. A brief discussion of Simulated Annealing meta-heuristic

Some exact algorithms have been formulated to work with the QAP [Ro87, CB89, MR94a, RRP96, PRR97] but its NP-hardness stimulates the use of heuristic techniques and meta-heuristic-based algorithms [BR84, WW87, SK89, Con89, Con90, Fl90, Ta91, MM92, MLP92, Wh93, BM93, LPR94, FF94, CSK94, SK94, MM95].

The Simulated Annealing (SA) meta-heuristics has been used with the QAP [BR84, WW87, Con89, and Con90]. These references will also be considered for a general SA description, allowing us, after a short recall, to deal only with our approach specificities.

Let us consider a minimization problem, where we examine at each iteration the increase \( \delta \) on the objective function. If \( \delta \) is negative the new solution will always be accepted, while for worse solutions (\( \delta > 0 \)) we evaluate the expression

\[
p = e^{-\delta / T(t)}
\]

where \( T(t) \) is a monotonically non-increasing **temperature function** and \( p \) is the probability of acceptance of a new (worse) solution. We then have a strategy to avoid the algorithm path to be stuck to local optima. The probability \( p \) will be greater at the beginning as we initialize \( T \) with a comparatively high value, a reduction ("cooling") strategy being applied during the algorithm execution, for which a given number of iterations is specified to use each new \( T \) value. The initial \( T \) value selection and the possible strategies for its reduction has been extensively discussed in the literature. The general SA form is as follows:
The stopping criterion has to do with the algorithm's efficiency in obtaining new better solutions. Eglese [Eg90] does a discussion on this point.

The criteria and the parameters used in the SA version here presented are supported by the theory discussed in Items 2 through 4.

5.2. The formulation of the RedInv-SA heuristic

We want to look for the best possible solution for the function

\[ Z_{\rho_{\overline{T}}} = \sum_{i=1}^{n} \overline{f_i d_{\rho_{\overline{T}}(i)}} \]  

(5.2)

where \( Z(\rho_{\overline{T}}) \) is the cost associated to the permutation \( \rho_{\overline{T}} \) on \( \overline{T} \) columns. The corresponding \( \Gamma \) column permutation, which defines the instance edge changing, can be obtained by the composition shown in (3.2).

We use Conjecture 4.1 to support the definition of a neighboring criterion for the algorithm to work. The strategy here presented is then a representative of a family of possible heuristic techniques using the decreasing number of inversions as a criterion for choosing elements to be exchanged [Que94].

The permutations on \( \overline{T} \) columns are related to those defined on \( \Gamma \) ones by the expression already presented,

\[ \rho_{\overline{T}} = \phi_{\overline{F}}^{-1} \circ \rho_{\Gamma} \circ \phi_{D}. \]  

(5.3)
We look for a position \( m \) in \( \rho_{\Gamma}^j \) which maximizes \(|i - \rho_{\Gamma}^j(i)|\). The exchange will be made with a position \( j \) which minimizes \(|m - \rho_{\Gamma}^j(i)| + |i - \rho_{\Gamma}^j(m)|\), so the decrease in the number of inversions will be maximum. The permutation \( \rho'_{\Gamma} \) thus obtained does not correspond to a feasible solution, as only two edges have been exchanged and a vertex exchange would deal with \( 2(n - 2) \) edges. We will then look for vertex exchanges whose realization implies exchanging edges \( i \) and \( m \), thus defining a neighborhood of feasible permutations. At this point, better quality solutions cannot be guaranteed, as the neighborhood could be that of a local optimum. The SA, nevertheless, was designed exactly to deal with such situations.

To find the neighborhood we use (3.2) to return to a \( \Gamma \) column permutation \( \rho_{\Gamma}^j \) and then look for the two positions in \( \rho_{\Gamma}^j \) which had their images exchanged. From an algorithmic point of view it is not necessary to execute the whole of (3.2) to obtain this result: we simply obtain these two positions, which will be

\[
p = \phi_F^{-1}(m) \quad \text{and} \quad q = \phi_F^{-1}(j),
\]

as well as the corresponding images \( t = \rho_F^{-1}(p) \) and \( u = \rho_F^{-1}(q) \).

We finish by applying the inverse function \( \psi^{-1} \) of (2.5),

\[
\psi^{-1}(t) = (x_t, y_t) \quad \text{and} \quad \psi^{-1}(u) = (x_u, y_u)
\]

where we can have either four different vertices, if the edges \( m \) and \( j \) are not adjacent, or, if they are adjacent, only three vertices. In the first case the vertex exchanges will be \( x_t \leftrightarrow x_u, x_t \leftrightarrow y_u, y_t \leftrightarrow x_u \) and \( y_t \leftrightarrow y_u \). The second case will give only three exchanges, the fourth one being void.

### 5.3. An exchange strategy example

The Nugent 5 instance [NVR68] is defined by

\[
F = [5, 2, 4, 1, 3, 0, 2, 0, 0, 5] \quad \text{and} \quad D = [1, 1, 2, 3, 2, 1, 2, 1, 2, 1]
\]

respectively for \( K_f \) and \( K_d \) edges in lexicographic order. Then the corresponding bijections defining the orderings for \( \overline{F} \) and \( \overline{D} \) will be

\[
\phi_F = [9 5 8 4 7 1 6 2 3 10] \quad \text{and} \quad \phi_D = [6 7 2 1 3 8 4 9 5 10].
\]

Taking as an example a vertex permutation

\[
\varphi = [4 1 3 5 2]
\]
we have, as \( \rho_\Gamma = [\psi(\varphi_l, \varphi_j)], 1 \leq i < j \leq n, \)
\[
\rho_\Gamma = [3 8 1 0 6 2 4 1 9 5 7]
\]
with cost 32. By (5.3) the corresponding permutation in \( \bar{\Gamma} \) is given by
\[
\rho_{\bar{\Gamma}} = [1 5 3 8 9 6 7 1 0 2 4]
\]
with 18 inversions.

We can easily find that \( m = 9, \) as \( \max |i - \rho_{\bar{\Gamma}}(i)| = 7 \) and \( j = 4, \) then we have \( \min \{|9 - \rho_{\bar{\Gamma}}(i)| + |i - \rho_{\bar{\Gamma}}(9)|\} = 3. \) The modified permutation \( \rho_{\bar{\Gamma}}' \) is
\[
\rho_{\bar{\Gamma}}' = [1 5 3 2 9 6 7 1 0 8 4].
\]
When returning to \( \Gamma \) we have, at first,
\[
\phi^{-1}_F(9) = 1; \rho_{\Gamma}(1) = 3 \quad \text{and} \quad \phi^{-1}_F(4) = 4; \rho_{\Gamma}(4) = 6.
\]
From where we obtain
\[
\psi^{-1}(3) = (1, 4) \quad \text{and} \quad \psi^{-1}(6) = (2, 4).
\]
The exchangings will then be done between the vertices defining the edges (1,4) and (2,4), in every possible arrangement. In this case the resulting neighborhood is then (1,2), (1,4) and (2,4). The vertex exchanges being defined on the image, we obtain respectively \([4 2 3 5 1], [1 4 3 5 2] \) and \([2 1 3 5 4] \). This last one has cost 30 and 17 inversions in its \( \bar{\Gamma} \) column permutation, which is \([1 5 9 8 3 2 7 4 6 10] \). We have \( \Delta z = -2 \), the objective function will decrease and the new solution will be accepted. If \( \Delta z \) were positive, we would use the SA proper criterion, according to (5.1).

5.4. The RedInv-SA algorithm

Given \( n \) and \( \Gamma \), we determine \( \psi, \phi_F, \phi_D \) and \( \bar{\Gamma} \) as previously defined.

We make
\[
T_0 = \sum_{i=1}^{N} f_i d_{N-i+1}
\]
(5.6)
and we begin with a randomly chosen vertex permutation from \( K_d \) vertices on \( K_f \) ones.
The initial temperature value will be taken as the problem upper bound (the element sum from the secondary diagonal of $\mathbf{F}$) reduced by an appropriate factor $K$. We used the function $T(t+1) = 0.9T(t)$, as proposed by Connolly [Con89]. The temperature function loop size $M$ ranged between one and nine (in one case forty) times the problem order $n$. The stopping criterion was $t < \text{tr}(\mathbf{F})/K$, but limited to ten (in two cases five) temperature values. As the edge exchanges are deterministic, a temporary hindrance was put on the last $n$ accepted solutions in order to avoid cycling.

The algorithm is as follows:

begin 
$T \leftarrow T_0$;
while ($T > \text{tr}(\mathbf{F})$ or $\text{ncyc} = \max\text{ncyc}$) do 
begin 
$i \leftarrow 0$;
while $i \leq n$ do 
begin 
procedure change; < return $(\mathbf{Z}, \varphi')$ >
$\Delta \mathbf{Z} \leftarrow \mathbf{Z} - \mathbf{Z}$;
if $\Delta \mathbf{Z} < 0$ then $\varphi \leftarrow \varphi'$;
else if $\exp(-\Delta \mathbf{Z}/T) < \text{random (0,1)}$ then $\varphi \leftarrow \varphi'$;
$i \leftarrow i + 1$;
end;
$T \leftarrow 0.9T$;
end;
end. < $\varphi'$: new solution; $Z_{\varphi'}$: cost >

The operation sequence within the procedure change is as follows, according with what was discussed before:

The procedure receives a solution $\varphi'$ with cost $Z_{\varphi'}$.

We determine

$$\rho_{\mathbf{F}} = \begin{pmatrix} 12 & 13 & \ldots & N - 1 & N \\ \varphi(1)\varphi(2) & \varphi(1)\varphi(3) & \ldots & \varphi(N - 1)\varphi(N) \end{pmatrix}$$

and

$$\rho_{\mathbf{F}} = \phi_{F}^{-1} \circ \rho_{\mathbf{T}} \circ \varphi_{D}.$$
We then find the positions for changing in $\rho_F$:

$$
m \leftarrow \text{position corresponding to } \max |i - \rho(i)|;
$$

$$
j \leftarrow \text{position corresponding to } \min \{|m - \rho(i)| + |i - \rho(m)|\}.
$$

We return to a $\Gamma$ column based permutation: $p = \phi_F^{-1}(m)$ and $q = \phi_F^{-1}(j)$.

We find the images $t = \rho_\Gamma(p)$ and $u = \rho_\Gamma(q)$.

We find $\psi^{-1}(t) = (a, b)$ and $\psi^{-1}(u) = (r, s)$.

We generate the neighbor set $V_\varphi'$:

$$
V_\varphi = \{\varphi'| \varphi'(i) = \varphi(i); i \neq a, b; \varphi'(a) = r \text{ or } s; \varphi'(b) = s \text{ or } r\}
$$

$$
\overline{Z} \leftarrow \min_{\varphi' \in V_\varphi} Z_{\varphi'};
$$

$$
\varphi \leftarrow \varphi';
$$

Return: new solution $\varphi'$ with value $Z_{\varphi'}$.

6. DISCUSSION OF RESULTS

The algorithm was programmed in Fortran-77. The computational tests were conducted in two equipments:

- a Pentium-166 microcomputer;
- a DEC workstation with a 170 Mhz Alpha processor.

The presented processes times are in every case those corresponding to P-166.

A sample of 35 symmetric instances from QAPLIB was processed. From these, 21 were lower-order ($12 \leq n \leq 18$) and 14 were higher-order ones ($n > 20$). In every case but the two larger instances, a number between 9 and 18 fifty-run tests were executed, combining better initial temperature values with different number of solutions per temperature value. The two larger instances [Tai60a and Sko81] used 25-run tests.

The cooling schedule was based on a constant reduction factor equal to 0.9 for all tests. The initial value was equal to the problem upper limit, reduced by a factor $K$:

$$
T_0 = \frac{\bar{F}^T \bar{D}_*}{K}
$$

(6.1)
where $\overline{D}^*$ is the vector $\overline{D}$ taken in inverse order. The schedule end was indicated by whichever the following constraints apply first:

- the last temperature going below the problem reduced lower bound, $\text{tr}(\overline{\Gamma})/K$;
- ten temperatures having already been obtained (five for Tai60a and Sko81).

The initial temperature range was estimated, for each instance, by running single runs and observing the behavior of the accepted/refused (A/R) solution relation for the set of temperature values, an initial temperature being considered appropriate when giving a fairly initial high A/R (say 4 to 8) and being able to decay to a fairly low A/R (say 0.2 to 0.4), thus showing good sensitivity to problem structure. The values for $K$ were chosen looking for circumscribing this sensible zone.

The number $\text{nsol}$ of solutions per temperature value is an indicator for the instance difficulty, as some instances required higher values of it to converge while for others the algorithm showed good results with much lesser values.

For each test, the minimum obtained value ($Z_{\text{min}}$), the average solution value ($\overline{Z}$) and the average P-166 processing time per run were obtained. The optimal (or better-known) value $Z_{\text{opt}}$ for the instance was utilized to calculate a percentile increment of the average on it,

$$ p_{av} = 100 \frac{\overline{Z}}{Z_{\text{opt}}} \quad (6.2) $$

and also a percentual increment of the test minimum obtained value on it,

$$ p_{\text{min}} = 100 \frac{Z_{\text{min}}}{Z_{\text{opt}}}. $$

In Tables 6.1 and 6.2 below we specify, for each instance:

- the number $\text{n\text{test}}$ of tests;
- the range of $\text{nsol}$ values;
- $p_{av}$, the average value of $p_{av}$ for the test set;
- the next column has a double meaning:
  - where $Z_{\text{opt}}$ was not obtained, this column shows $p_{\text{ms}}$, the minimum $p_{\text{min}}$ value for the whole set of tests.
  - where $Z_{\text{opt}}$ was obtained (so $p_{\text{ms}} = 0$) the cell contains the number $n_{\text{opt}}$ of this event for the whole test set.
- the average run time (in seconds) for the test set.
### Table 6.1

RedInv-SA results with lower-order instances ($12 \leq n \leq 18$).

<table>
<thead>
<tr>
<th>Instance</th>
<th>ntest</th>
<th>nsol</th>
<th>$p_{av}$</th>
<th>$p_{ms-n_{opt}}$</th>
<th>Av. time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rou12</td>
<td>9</td>
<td>30-50</td>
<td>2.04</td>
<td>(23)</td>
<td>0.55</td>
</tr>
<tr>
<td>Had12</td>
<td>9</td>
<td>12-36</td>
<td>2.23</td>
<td>(1)</td>
<td>0.15</td>
</tr>
<tr>
<td>Nug12</td>
<td>10</td>
<td>10-80</td>
<td>5.01</td>
<td>(39)</td>
<td>0.38</td>
</tr>
<tr>
<td>Tai12a</td>
<td>12</td>
<td>12-48</td>
<td>7.36</td>
<td>(17)</td>
<td>0.36</td>
</tr>
<tr>
<td>Scr12</td>
<td>12</td>
<td>12-36</td>
<td>9.82</td>
<td>(39)</td>
<td>0.51</td>
</tr>
<tr>
<td>Chr12b</td>
<td>9</td>
<td>36-108</td>
<td>34.42</td>
<td>(16)</td>
<td>1.69</td>
</tr>
<tr>
<td>Chr12a</td>
<td>16</td>
<td>48-480</td>
<td>40.91</td>
<td>5.73</td>
<td>4.10</td>
</tr>
<tr>
<td>Had14</td>
<td>9</td>
<td>20-70</td>
<td>1.64</td>
<td>(4)</td>
<td>0.65</td>
</tr>
<tr>
<td>Tai15a</td>
<td>16</td>
<td>15-120</td>
<td>4.78</td>
<td>0.20</td>
<td>1.40</td>
</tr>
<tr>
<td>Nug15</td>
<td>10</td>
<td>10-45</td>
<td>4.81</td>
<td>(12)</td>
<td>0.83</td>
</tr>
<tr>
<td>Rou15</td>
<td>9</td>
<td>30-90</td>
<td>5.38</td>
<td>(4)</td>
<td>1.77</td>
</tr>
<tr>
<td>Scr15</td>
<td>16</td>
<td>30-240</td>
<td>15.90</td>
<td>(2)</td>
<td>1.44</td>
</tr>
<tr>
<td>Chr15a</td>
<td>12</td>
<td>135-225</td>
<td>58.61</td>
<td>6.20</td>
<td>2.60</td>
</tr>
<tr>
<td>Chr15b</td>
<td>16</td>
<td>45-225</td>
<td>83.40</td>
<td>9.01</td>
<td>1.98</td>
</tr>
<tr>
<td>Esc16b</td>
<td>9</td>
<td>16-48</td>
<td>1.89</td>
<td>(89)</td>
<td>0.95</td>
</tr>
<tr>
<td>Nug16a</td>
<td>14</td>
<td>32-128</td>
<td>4.31</td>
<td>(2)</td>
<td>3.29</td>
</tr>
<tr>
<td>Esc16c</td>
<td>12</td>
<td>16-48</td>
<td>11.19</td>
<td>(29)</td>
<td>2.24</td>
</tr>
<tr>
<td>Had16</td>
<td>16</td>
<td>32-112</td>
<td>11.57</td>
<td>0.05</td>
<td>1.55</td>
</tr>
<tr>
<td>Esc16d</td>
<td>9</td>
<td>8-32</td>
<td>27.12</td>
<td>(69)</td>
<td>1.38</td>
</tr>
<tr>
<td>Had18</td>
<td>16</td>
<td>36-126</td>
<td>1.66</td>
<td>0.15</td>
<td>2.72</td>
</tr>
<tr>
<td>Nug18</td>
<td>16</td>
<td>36-144</td>
<td>4.94</td>
<td>(1)</td>
<td>5.86</td>
</tr>
</tbody>
</table>

### Table 6.2

RedInv-SA results with higher-order instances ($20 \leq n \leq 100$).

<table>
<thead>
<tr>
<th>Instance</th>
<th>ntest</th>
<th>nsol</th>
<th>$p_{av}$</th>
<th>$p_{ms-n_{opt}}$</th>
<th>Av. time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Had20</td>
<td>16</td>
<td>40-160</td>
<td>1.64</td>
<td>0.15</td>
<td>5.36</td>
</tr>
<tr>
<td>Nug20</td>
<td>9</td>
<td>20-300</td>
<td>3.39</td>
<td>0.15</td>
<td>16.55</td>
</tr>
<tr>
<td>Rou20</td>
<td>16</td>
<td>80-300</td>
<td>3.84</td>
<td>0.38</td>
<td>13.72</td>
</tr>
<tr>
<td>Tai20a</td>
<td>15</td>
<td>40-160</td>
<td>5.59</td>
<td>1.87</td>
<td>7.68</td>
</tr>
<tr>
<td>Scr20</td>
<td>15</td>
<td>60-240</td>
<td>21.29</td>
<td>0.02</td>
<td>17.56</td>
</tr>
<tr>
<td>Tai25a</td>
<td>18</td>
<td>50-250</td>
<td>5.98</td>
<td>1.71</td>
<td>25.55</td>
</tr>
<tr>
<td>Chr25a</td>
<td>12</td>
<td>100-400</td>
<td>110.00</td>
<td>50.31</td>
<td>61.67</td>
</tr>
<tr>
<td>Nug30</td>
<td>12</td>
<td>60-300</td>
<td>8.53</td>
<td>2.68</td>
<td>88.95</td>
</tr>
<tr>
<td>Esc32a</td>
<td>10</td>
<td>32-96</td>
<td>51.33</td>
<td>25.69</td>
<td>60.86</td>
</tr>
<tr>
<td>Esc32b</td>
<td>12</td>
<td>64-256</td>
<td>49.75</td>
<td>27.97</td>
<td>101.56</td>
</tr>
<tr>
<td>Ste36a</td>
<td>12</td>
<td>36-144</td>
<td>52.82</td>
<td>17.23</td>
<td>89.44</td>
</tr>
<tr>
<td>Sko42</td>
<td>18</td>
<td>84-672</td>
<td>7.84</td>
<td>3.74</td>
<td>598.16</td>
</tr>
<tr>
<td>Tai60a</td>
<td>9</td>
<td>120-360</td>
<td>7.58</td>
<td>5.43</td>
<td>1456.33</td>
</tr>
<tr>
<td>Sko81</td>
<td>6</td>
<td>81-324</td>
<td>8.57</td>
<td>6.47</td>
<td>2856.00</td>
</tr>
</tbody>
</table>

Recherche opérationnelle/Operations Research
The table was ordered by $p_{av}$ values in crescent order for each problem size. These values can be seen as a proximity measure (between $Z_{opt}$ and the algorithm path) which can be used as a problem difficulty criterion. We can see, for example, that the algorithm is not efficient for some tree-structured Christofides instances: it looks like it cannot easily find a path along that structure. Nevertheless in one case (Chr12b) the $p_{av}$ criterion does not seem to apply, the average differences staying high whereas a number of optimal values were found. As with any other technique, the problem order has a strong influence: the Nugent series show it clearly. There appears a curious question concerning the instances Esc16 and 32: in the former case we have an important number of different optimal solutions, which allowed the algorithm to present very good results. This is possibly not true for the latter case, where the results were worse than, for example, those shown by Ste36a. We did not consider this question of instance difficulty to be within the scope of this paper; some discussion of it can be found in [MR94b].

Some examples of the algorithm convergence are given in Table 6.3 below. Here we can observe for some sample instances how $p_{av}$ (which we define as the average of the $p_{av}$ values corresponding to a given $nsol$ value) goes to lesser values as $nsol$ grows.

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Had 18</td>
<td>Nug 18</td>
<td>Tai 25a</td>
<td>Sko 42</td>
<td>Sko 81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nsol</td>
<td>$p_{av}$</td>
<td>nsol</td>
<td>$p_{av}$</td>
<td>nsol</td>
<td>$p_{av}$</td>
<td>nsol</td>
<td>$p_{av}$</td>
<td>nsol</td>
</tr>
<tr>
<td>36</td>
<td>1,90</td>
<td>36</td>
<td>5,89</td>
<td>50</td>
<td>6,75</td>
<td>84</td>
<td>8,56</td>
<td>81</td>
</tr>
<tr>
<td>54</td>
<td>1,78</td>
<td>72</td>
<td>5,01</td>
<td>100</td>
<td>5,98</td>
<td>168</td>
<td>8,09</td>
<td>162</td>
</tr>
<tr>
<td>90</td>
<td>1,53</td>
<td>108</td>
<td>4,48</td>
<td>150</td>
<td>5,76</td>
<td>336</td>
<td>7,38</td>
<td>324</td>
</tr>
<tr>
<td>126</td>
<td>1,42</td>
<td>144</td>
<td>4,39</td>
<td>200</td>
<td>5,41</td>
<td>672</td>
<td>6,96</td>
<td></td>
</tr>
</tbody>
</table>

Another interesting point to observe in the tests is the comparison between costs and inversion numbers along the algorithm path. The graphs from Figure 6.1a show the costs and the inversion numbers for the solutions examined by the algorithm during a run with the Nugent 12 instance. The graph segments within the rectangles are zoomed in Figure 6.1b, where the solutions are individually shown. The same is shown for Nugent 15 in Figures 6.2a and 6.2b. This type of graph can be found for any instance.
A detailed analysis would require the determination of free comparability occurrence along the solution trail, but even without that we can see the close relationship between these two solution properties.

6. FINAL REMARKS AND CONCLUSIONS

Although we give information about time processing we do not discuss it, the variety of hardware resources quoted in the literature as supports for computer tests making any time comparison very difficult.
On the other hand, the number of solution changing made by previously known SA schemes is known to be between

- 1,000 and 4,000 for \( n \) between 12 and 15;
- 4,000 and 16,000 for \( n \) near 20;
- 5,000 and 30,000 for \( n = 30 \).

As a comparison, the maximum \texttt{nsol} values used in the study, multiplied by 10 (maximum number of different temperatures) gave, for the attained convergence levels, values between

- 360 and 2,250 for \( n \) between 12 and 15 (exception: Chr12a with 4,800);
• 320 and 3,840 for $n$ between 16 and 18;
• 960 and 6,720 for $n$ equal or greater than 20.

These values will evidently be pushed up when going further into convergence but we can see there is a fairly good slack, when relating them with previous results, for the work involved with the majority of the tested instances.

For problem order 18 or less the algorithm found the optimal (or best-known) solution one or more times in the majority of the tested instances. With order 20 it went near 0.4% or less in every but one instance. As it could be expected the difficulty arise with the size; nevertheless the convergence can be observed even in higher-order instances.

The second important point concerns Conjecture 4.1. By looking at Figures 6.1a to 6.2b we can see that this conjecture is reinforced by the results, similar graphs being easy to build for any instance. As the inversion number détermination is an elementary procedure we think we have in it an efficient parameter for result analysis when using metaheuristics to solve QAP instances.

ACKNOWLEDGEMENTS

The authors are grateful to the referees by their very useful improvement suggestions, which strongly contributed to give more clarity to the text. They are also indebted to the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), whose financial support was essential to the realization of this work. The co-author Tania Maia Querido is also grateful for its support to her work on QAP at Florida University.

REFERENCES


Recherche opérationnelle/Operations Research


