Travelling Salesman Problem

MIGUEL A. S. CASQUILHO
Technical University of Lisbon, Ave. Rovisco Pais, 1049-001 Lisboa, Portugal

The "Travelling Salesman Problem" is briefly presented, with reference to problems that can be assimilated to it and solved by the same technique. Examples are shown and solved.

Key words: *travelling salesman problem; supply; demand; optimization.*

1. Fundamentals and scope

With optimization in mind, the "travelling salesman problem", frequently denoted by the initials TSP[1], is a fundamental subject related to travelling and transportation, with several generalizations and with insertion in more complex situations, and also akin to others apparently unrelated, resoluble by the techniques used for the typical case. The TSP is known for the striking contrast between the simplicity of its formulation and the difficulty of its resolution, some even saying that it still does not have a solution. It is a so-called NP-hard[2] problem (its difficulty increasing more than polynomially with its size). Anyway, something substantial can be presented about the problem.

The problem arises from the typical situation of a salesman who wants to visit his clients in a given set of cities and return to his own city, thus performing a cycle. The problem can be envisaged in this large scale, but also exists in any other scales, such as within a factory or on a microchip. An asymmetrical TSP can also be the search for the optimum ordering of paint manufacturing or the preparation of fruit juices in a common plant, because, in these cases, setup costs (washing, etc.) depends significantly on the "vicinity" of the colours or of the flavours.

The mathematical formulation of the problem can be as in Eq. {1}.

$$[\min] z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij}$$

subject to

$$\begin{aligned} \sum_{j=1}^{n} x_{ij} &= 1 \qquad i = 1..m \\ \sum_{i=1}^{m} x_{ij} &= 1 \qquad j = 1..n \\ x_{ij} &= 0 \text{ or } 1 \qquad \forall i, j \end{aligned} \qquad \{1\}$$

Solution must be a cycle.

As in the Assignment Problem (AP), c_{ij}, with i, j= 1..n, is the cost (or distance, etc.) to go from city i to city j, and x_{ij} will be 1 if the arc from i to j is used and 0 otherwise. The problem would be an AP if the last condition (one and only one

[1] Also "traveling salesman problem" (American English) or sometimes "salesperson".
[2] See, *e.g.*, http://en.wikipedia.org/wiki/NP-hard .

Miguel Casquilho is Assistant Professor (retired) in the Departament of Chemical Engineering, Instituto Superior Técnico, Universidade Técnica de Lisboa (Technical University of Lisbon), Lisbon, Portugal. E-mail address: *mcasquilho@ist.utl.pt*.

cycle) —a fundamental difference— did not exist. In order to use the above formulation the costs c_{ii}, to go from one city to itself must be made prohibitive, otherwise the (useless) solution would be $x_{ii} = 1$, others zero (not a cycle).

A branch-and-bound (B&B) exact algorithm based on AP relaxation will be presented with some examples. A language such as Mathematica has functions *TravelingSalesman* and *FindShortestTour* that solve the problem, even in 3D, but the objective here is to clarify the method itself. An advanced algorithm by Carpaneto *et al.* [1995CAR] is used and made available at the author's website [2012CAS]. Several other exact algorithms, all detailed by Lawler *et al.* [1995LAW], have been constructed, as well as heuristic ones (i.e., approximate), namely for large size problems or more complex problems.

2. Examples

EXAMPLE: 5 CITIES (WINSTON)

A travelling salesman has to cover a set of 5 cities (his own included) periodically (say, once per week) and return home. The distances between the cities are given in Table 1, as could have been read on a map. Determine the most economical cycle, i.e., with minimum length (example from Winston [2003WIN], p 530 ff).

Table 1 — Cost matrix (distances in km) from one city to another

	1	2	3	4	5
1	—	132	217	164	58
2	132	—	290	201	79
3	217	290	—	113	303
4	164	201	113	—	196
5	58	79	303	196	—

In this problem the cost matrix is symmetric, which appears obvious. The cost matrix can, however, be asymmetric, as in the case of air travel because of predominant wind or in one-way urban streets. No advantage is taken from symmetry in the present text, but symmetry may be important in many variants of the solution methods.

RESOLUTION

The strategy of AP relaxation is to solve the problem as an AP, leaving aside the condition of "one cycle". The solution to the AP can be obtained, *e.g.*, with Excel or Excel/Cplex, and is the one in Table 2, with total cost $z = 495$ km.

Table 2 — Solution to the AP relaxation

	1	2	3	4	5
1	—	0	0	0	1
2	1	—	0	0	0
3	0	0	—	1	0
4	0	0	1	—	0
5	0	1	0	0	—

This is, however, not a solution to the TSP, because there are subtours: $x_{15} = x_{21} = x_{34} = x_{43} = x_{52} = 1$, i.e., two subtours, 1–5–2–1 and 3–4–3. The B&B technique will now be used, as follows.

The former problem, say, Problem 1, is replaced by others, considering the shortest subtour (the one with least arcs) to try to save computation effort. So, 3–4–3 will be chosen. Now, two problems replace the previous one: Problem 2, from Problem 1 but prohibiting 3–4; and Problem 2, from Problem 1 but prohibiting 4–3. The two new problems have the cost matrices in Table 3 and the solutions in Table 4

Table 3 — Cost matrices for Problem 2 (l.-h.) and Problem 3 (r.-h.)

	1	2	3	4	5
1	—	132	217	164	58
2	132	—	290	201	79
3	217	290	—	—	303
4	164	201	113	—	196
5	58	79	303	196	—

	1	2	3	4	5
1	—	132	217	164	58
2	132	—	290	201	79
3	217	290	—	113	303
4	164	201	—	—	196
5	58	79	303	196	—

Table 4 — Solutions to the AP relaxations

	1	2	3	4	5
1	—	0	0	1	0
2	0	—	0	0	1
3	1	0	—	0	0
4	0	0	1	—	0
5	0	1	0	0	—

	1	2	3	4	5
1	—	0	1	0	0
2	0	—	0	0	1
3	0	0	—	1	0
4	1	0	0	—	0
5	0	1	0	0	—

Again, these are not solutions to the TSP, because there are subtours: for Problem 2, 1–4–3–1, 2–5–2; and for Problem 3, 1–3–4–1, 2–5–2. (Symmetry causes some redundancy.) In file *TSP_Wins530.xls*, the complete procedure is shown until the optimum is found.

The solution procedure is usually shown as a tree, constructed progressively as the problem is solved. The one for this problem is (from *TSP_Raffensperger.ppt*) given in Figure 1, with solution 1–5–2–4–3–1 and $z^* = 668$ km.

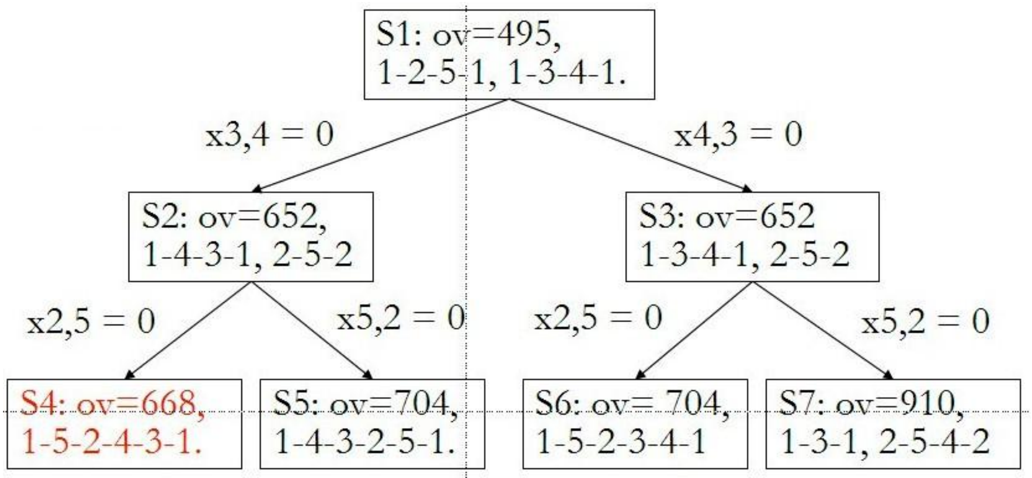

Figure 1 — Tree for the Winston TSP problem.

■

About the B&B, notice that:

(a) (disadvantage) Every replacement of a problem gives rise to two or more "children" and can never improve the value of the objective function (495, 652, 668, increasing, while a minimum was sought); but

(b) (advantage) It is, hopefully, not necessary to investigate all the combinations, this being the real merit of the B&B.

The systematic procedure of B&B permits to avoid the investigation of all the combinations. For a mere $n = 18$ cities —the number of Portugal's main cities ("district capitals"[3])—, these combinations would be $17\,!$, which is $3{,}5 \times 10^{14}$ (11 years of computing at one combination per microsecond).

The motive it is considered that the TSP still "has no solution" is that the size of the tree is not predictable. So, the amount of memory or disk storage occupied can increase beyond the availability, making this not only a matter of time, but often more so one of space.

EXAMPLE: GRID 2

In a manufacture of grids having 4 points in a square arrangement of 2×2, as in Figure 2, all the points have to be treated (*e.g.*, welded, connected, painted) in a certain order. Determine the most economic order.

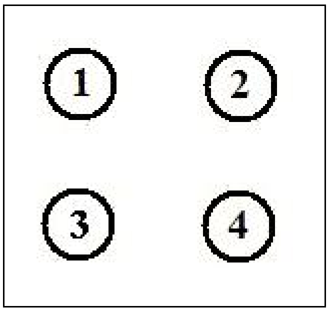

Figure 2 — Grid with 2 by 2 points in a square arrangement.

RESOLUTION

It is obvious that the answer is 1–2–4–3, with $z^* = 4$, if, as is implicit, the distances between the adjacent points (horizontally or vertically) is one. If, however, the problem is to be solved by a convenient algorithm[4], the cost matrix has to be provided and is given in Table 5. In order to use the web site, supply –1 as infinity (diagonal entries).

Table 5 — Cost matrix for the "Grid 2"

	1	2	3	4
1	—	1	1	2
2	1	—	2	1
3	1	2	—	1
4	2	1	1	—

[3] The *district* ("distrito"), of which there are 18 (average 5 500 km^2) in continental Portugal, is roughly equivalent in area to one French *département* (totalling 96) or two Italian *province* (110).

[4] The one on the web, based on Carpaneto *et al.*.

Example: Grid 3

Now suppose that (previous example) it is the manufacture of grids having 9 points in a square arrangement of 3×3, as in Figure 3. Determine the most economic order.

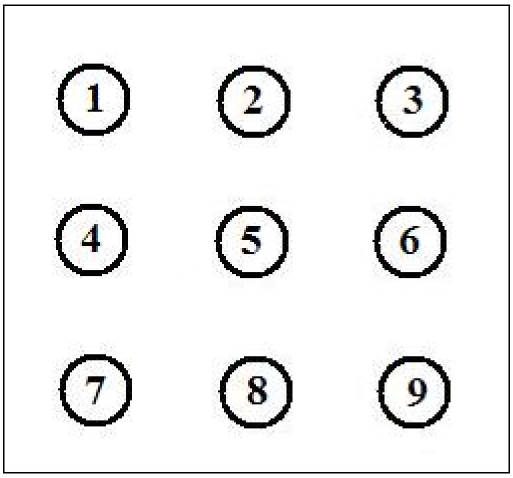

Figure 3 — Grid with 3 by 3 points in a square arrangement.

Resolution

Now the solution is by no means obvious. See *TSP_grids.xls*. The problem is solved assuming both a *taxicab*[5] *geometry*[6] and a *Euclidean geometry*. The taxicab geometry assumes a grid layout, such as the arrangement of streets in certain zones of the cities, hence the mention to the taxicab, so (as in the previous example) only "vertical" and "horizontal" movements are possible. The cost matrix is given in Table 6.

Table 6 — Taxicab cost matrix for the "Grid 3"

	1	2	3	4	5	6	7	8	9
1	—	1	2	1	2	3	2	3	4
2	1	—	1	2	1	2	3	2	3
3	2	1	—	3	2	1	4	3	2
4	1	2	3	—	1	2	1	2	3
5	2	1	2	1	—	1	2	1	2
6	3	2	1	2	1	—	3	2	1
7	2	3	4	1	2	3	—	1	2
8	3	2	3	2	1	2	1	—	1
9	4	3	2	3	2	1	2	1	—

The solution is 1-2-5-4-7-8-9-6-3-1, with $z^* = 10$. Looking at these results, one may wonder if allowing other movements may result in a better value. This leads to the common, Euclidean geometry. Now the cost matrix is given in Table 7.

The solution to the Euclidean geometry Grid 3 problem, as shown in Table 7, is 1-4-7-8-9-6-3-5-2-1, with $z^* = 9{,}41$. As expected, allowing a less constrained geometry has led to a better result.

---

[5] From *taximeter cabriolet*.

[6] See, *e.g.*, http://en.wikipedia.org/wiki/Taxicab_geometry .

Table 7 — Solution for the Euclidean cost matrix for the "Grid 3"

	1	2	3	4	5	6	7	8	9
1	—	1	2	1	1,41421	2,23607	2	2,23607	2,82843
2	1	—	1	1,41421	1	1,41421	2,23607	2	2,23607
3	2	1	—	2,23607	1,41421	1	2,82843	2,23607	2
4	1	1,41421	2,23607	—	1	2	1	1,41421	2,23607
5	1,41421	1	1,41421	1	—	1	1,41421	1	1,41421
6	2,23607	1,41421	1	2	1	—	2,23607	1,41421	1
7	2	2,23607	2,82843	1	1,41421	2,23607	—	1	2
8	2,23607	2	2,23607	1,41421	1	1,41421	1	—	1
9	2,82843	2,23607	2	2,23607	1,41421	1	2	1	—

3. Conclusions

The TSP is a remarkable problem both for the contrast between the simplicity of its formulation and the complexity of its resolution, and the variety of its applications. The exact resolution was presented using the branch-and-bound technique applied to Assignment Problem relaxations. A more advanced algorithm by Carpaneto *et al.* is made available for the solution of this type of problems.

Examples of typical situations were presented, namely, a problem with taxicab and Euclidean distances as costs.

Acknowledgements

This work was done at "Centro de Processos Químicos do IST" (Chemical Process Research Centre of IST), Department of Chemical Engineering, Technical University of Lisbon. Computations were done on the central system of CIIST (Computing Centre of IST).

Bibliography

– [1995CAR] CARPANETO, G., M. DELL'AMICO and P. TOTH, 1995, "Exact solution of large-scale asymmetric traveling salesman problems", *ACM Transactions on Mathematical Software*, 21:394–409.

– [2012CAS] CASQUILHO, M., http://web.ist.utl.pt/~mcasquilho/compute/or/Fx-TSP.php, accessed May 2012..

– [1985LAW] LAWLER, E. L., J. K. LENSTRA, A. H. G. RINNOOY KAN, D. B. Shmoys, "The travelling salesman problem", editors, Chichester (UK). ISBN 0-471-90413-9.

– [2003WIN] WINSTON, Wayne L., and M. VENKATARAMANAN, 2003, "Introduction to Mathematical Programming", Thomson, New York, NY (USA). ISBN 0-534-35964-7

❖