On the Invariance of Ant System

Mauro Birattari[1], Paola Pellegrini[1,2], and Marco Dorigo[1]

[1] IRIDIA, CoDE, Université Libre de Bruxelles, Brussels, Belgium
[2] Department of Applied Mathematics, Università Ca' Foscari, Venice, Italy
mbiro@ulb.ac.be, paolap@pellegrini.it, mdorigo@ulb.ac.be

Abstract. It is often believed that the performance of ant system, and in general of ant colony optimization algorithms, depends somehow on the scale of the problem instance at hand. The issue has been recently raised explicitly [1] and the *hyper-cube framework* has been proposed to eliminate this supposed dependency.

In this paper, we show that although the internal state of ant system—that is, the *pheromone* matrix—depends on the scale of the problem instance under analysis, this does not affect the external behavior of the algorithm. In other words, for an appropriate initialization of the pheromone, the sequence of solutions obtained by ant system does not depend on the scale of the instance.

As a second contribution, the paper introduces a straightforward variant of ant system in which also the pheromone matrix is independent of the scale of the problem instance under analysis.

1 Introduction

The *hyper-cube framework* [1] has been recently introduced with the aim of implementing ant colony optimization algorithms (ACO) [2] that are invariant with respect to a linear rescaling of problem instances. The need for the introduction of the hyper-cube framework has been explicitly motivated by the observation that

> *in standard ACO algorithms the pheromone values and therefore the performance of the algorithms, strongly depend on the scale of the problem.* [1]

In this paper, we formally show that this statement is only partially correct: Indeed, in standard ant colony optimization algorithms the pheromone trail and the heuristic values depend on the scale of the problem. Nonetheless, for an appropriate initialization of the pheromone, the sequence of solutions they find is independent of the scaling.

For definiteness, the paper focuses on ant system [3,4,5] for the traveling salesman problem. The theorems we enunciate in the paper are proved first for this specific algorithm and for this specific problem. The conditions under which these results extend to other problems are discussed in the following.

Although this paper shows that the main motivation for the introduction of the hyper-cube framework does not hold, the work of Blum and Dorigo [1] has the

M. Dorigo et al. (Eds.): ANTS 2006, LNCS 4150, pp. 215–223, 2006.

main merit of having explicitly attracted the attention of the research community on some important issues. Indeed, the fact that pheromone and heuristic values depend on the scale of the problem complicates the analysis of the algorithm and might cause numerical problems in the implementations. The hyper-cube framework is definitely a solution to this problem. Nonetheless, the hyper-cube version of ant system is effectively a new algorithm which shares with the original ant system the underlying ideas but that produces a different sequence of solutions. In other words, the hyper-cube ant system and the original ant system are not *functionally equivalent*. In this paper we propose *si*AS which is a trivial modification of ant system. Similar to the hyper-cube ant system, *si*AS has the property that the pheromone and the heuristic values do not depend on the scaling of the problem. Nevertheless, contrary to the hyper-cube ant system, *si*AS is *functionally equivalent* to the original ant system. This last property is particularly desirable: all theoretical and empirical studies previously performed on ant system immediately extend to *si*AS.

In this paper, we focus our attention on ant system. Nonetheless the same invariance property can be proved for other ACO algorithms. We refer the reader to [6] for an analysis of the invariance of $\mathcal{MAX}$–$\mathcal{MIN}$ ant system [7,8] and of ant colony system [9]. Moreover, in [6] the algorithms *si*$\mathcal{MM}$AS and *si*ACS are defined, which are *functionally equivalent* to $\mathcal{MAX}$–$\mathcal{MIN}$ ant system and ant colony system, respectively, and in which the pheromone and the heuristic values do not depend on the scaling of the problem.

The rest of the paper is organized as follows. Section 2 introduces some preliminary concepts. Section 3 defines ant system and formally proves its invariance. Section 4 introduces the *si*AS algorithm. Finally, Sect. 5 concludes the paper.

2 Preliminary Definitions

This section introduces a number of fundamental concepts that will be needed in the following.

Definition 1 (Linear transformation of a problem instance). *If I is an instance of a generic combinatorial optimization problem, $\bar{I} = fI$, $f > 0$, is a* linear transformation *of I if $\bar{I}$ is obtained by multiplying all costs in I by the coefficient f. In particular, it results that the cost $\bar{C}$ of a solution $\bar{T}$ of instance $\bar{I}$ is f times the cost C of the corresponding solution T of instance I.*

Definition 2 (Linear transformation of a traveling salesman instance). *With $\bar{I} = fI$, $f > 0$, we indicate that the instance $\bar{I}$ is a* linear transformation *of the instance I: The two instances have the same number of cities and the cost $\bar{c}_{ij}$ of traveling from city i to city j in $\bar{I}$ is f times the corresponding cost c_{ij} in instance I. Formally:*

$$\bar{c}_{ij} = fc_{ij}, \text{ for all } \langle i, j\rangle. \tag{1}$$

Remark 1. The cost $\bar{C}$ of a solution $\bar{T}$ of instance $\bar{I}$ is f times the cost C of the corresponding solution T of instance I. Formally:

$$(\bar{I} = fI) \wedge (\bar{T} = T) \implies \bar{C} = fC. \tag{2}$$

Remark 2. In the following, if x is a generic quantity that refers to an instance I, then $\bar{x}$ is the corresponding quantity for what concerns instance $\bar{I}$, when $\bar{I}$ is a linear transformation of I.

Ant colony optimization algorithms are *stochastic*: Solutions are constructed incrementally on the basis of stochastic decisions that are biased by the pheromone and by some heuristic information. The following hypothesis will be used in the paper.

Hypothesis 1 (Pseudo-random number generator). When solving two instances I and $\bar{I}$, the stochastic decisions taken while constructing solutions are made on the basis of random experiments based on pseudo-random numbers produced by the same pseudo-random number generator. We assume that this generator is initialized in the same way (for example, with the same seed) when solving the two instances so that the two sequences of pseudo-random numbers that are generated are the same in the two cases.

Definition 3 (Invariance). *An algorithm A is **invariant** to linear transformations if the sequence of solutions S_I generated when solving an instance I and the sequence of solutions $S_{\bar{I}}$ generated when solving an instance $\bar{I}$ are the same, whenever $\bar{I}$ is a linear transformation of I.*

If A is a stochastic algorithm, it is said to be invariant if it is so under Hypothesis 1.

Definition 4 (Strong and weak invariance). *An algorithm A is said to be **strongly-invariant** if, beside generating the same solutions on any two linearly related instances I and $\bar{I}$, it also enjoys the property that the heuristic information and the pheromone at each iteration are the same when solving I and $\bar{I}$. Conversely, the algorithm A is **weakly-invariant** if it obtains the same solutions on linearly related instances but the heuristic information and the pheromone assume different values.*

If A is stochastic, it is said to be strongly-invariant (or weakly-invariant) if it is so under Hypothesis 1.

3 Ant System

Ant system is the original ant colony optimization algorithm proposed by Dorigo et al. [3,4,5]. The pseudo-code of the algorithm is given in Fig. 1. In our analysis, we refer to the application of ant system to the well-known traveling salesman problem, which consists in finding the Hamiltonian circuit of least cost on an edge-weighted graph.

Ant system:
Initialize pheromone trail
while (termination condition not met) **do**
Construct solutions via the random proportional rule
Update pheromone
end

Fig. 1. Pseudo-code of ant system

Definition 5 (Random proportional rule). *At the generic iteration h, suppose that ant k is in node i. Let $\mathcal{N}_i^k$ be the set of feasible nodes. The node $j \in \mathcal{N}_i^k$, to which ant k moves, is selected with probability:*

$$p_{ij,h}^k = \frac{[\tau_{ij,h}]^\alpha [\eta_{ij}]^\beta}{\sum_{l \in \mathcal{N}_i^k} [\tau_{il,h}]^\alpha [\eta_{il}]^\beta},$$

where α and β are parameters, $\tau_{ij,h}$ is the pheromone value associated with arc $\langle i, j \rangle$ at iteration h, and η_{ij} represents heuristic information *on the desirability of visiting node j after node i.*

Definition 6 (Heuristic information). *When solving the traveling salesman problem, the heuristic information η_{ij} is the inverse of the cost of traveling from city i to city j:*

$$\eta_{ij} = \frac{1}{c_{ij}}, \textit{ for all } \langle i, j \rangle.$$

Definition 7 (Pheromone update rule). *At the generic iteration h, suppose that m ants have generated the solutions $T_h^1, T_h^2, \dots, T_h^m$ of cost $C_h^1, C_h^2, \dots, C_h^m$, respectively. The pheromone on each arc $\langle i, j \rangle$ is updated according to the following rule:*

$$\tau_{ij,h+1} = (1 - \rho)\tau_{ij,h} + \sum_{k=1}^{m} \Delta_{ij,h}^k,$$

where ρ is a parameter called evaporation rate *and*

$$\Delta_{ij,h}^k = \begin{cases} 1/C_h^k, & \textit{if } \langle i, j \rangle \in T_h^k; \\ 0, & \textit{otherwise.} \end{cases} \tag{3}$$

Definition 8 (Ant system). *Ant system is an ant colony optimization algorithm in which solutions are constructed according to the random proportional rule given in Definition 5, and the pheromone is updated according to the rule given in Definition 7. The evaporation rate ρ, the number of ants m, and the exponents α and β are parameters of the algorithm.*

When ant system is used for solving the traveling salesman problem, it is customary to initialize the pheromone as follows.

Definition 9 (Nearest-neighbor pheromone initialization). *At the first iteration $h = 1$, the pheromone on all arcs is initialized to the value:*

$$\tau_{ij,1} = \frac{m}{C^{nn}}, \textit{ for all } \langle i, j \rangle,$$

where m is the number of ants considered at each iteration, and C^{nn} is the cost of the solution T^{nn} obtained by the nearest-neighbor heuristic.

The following theorem holds true.

Lemma 1. *The random proportional rule is invariant to concurrent linear transformations of the pheromone and of the heuristic information. Formally:*

$$(\bar{\tau}_{ij,h} = g_1 \tau_{ij,h}) \wedge (\bar{\eta}_{ij} = g_2 \eta_{ij}), \textit{ for all } \langle i, j \rangle \implies \bar{p}^k_{ij,h} = p^k_{ij,h}, \textit{ for all } \langle i, j \rangle.$$

where $\bar{p}^k_{ij,h}$ is obtained on the basis of $\bar{\tau}_{ij,h}$ and $\bar{\eta}_{ij}$, according to Definition 5.

Proof. According to Definition 5:

$$\begin{aligned}
\bar{p}^k_{ij,h} &= \frac{[\bar{\tau}_{ij,h}]^\alpha [\bar{\eta}_{ij}]^\beta}{\sum_{l \in \mathcal{N}_i^k} [\bar{\tau}_{il,h}]^\alpha [\bar{\eta}_{il}]^\beta} = \frac{[g_1 \tau_{ij,h}]^\alpha [g_2 \eta_{ij}]^\beta}{\sum_{l \in \mathcal{N}_i^k} [g_1 \tau_{il,h}]^\alpha [g_2 \eta_{il}]^\beta} \\
&= \frac{[g_1]^\alpha [g_2]^\beta [\tau_{ij,h}]^\alpha [\eta_{ij}]^\beta}{\sum_{l \in \mathcal{N}_i^k} [g_1]^\alpha [g_2]^\beta [\tau_{il,h}]^\alpha [\eta_{il}]^\beta} = \frac{[g_1]^\alpha [g_2]^\beta [\tau_{ij,h}]^\alpha [\eta_{ij}]^\beta}{[g_1]^\alpha [g_2]^\beta \sum_{l \in \mathcal{N}_i^k} [\tau_{il,h}]^\alpha [\eta_{il}]^\beta} \\
&= \frac{[\tau_{ij,h}]^\alpha [\eta_{ij}]^\beta}{\sum_{l \in \mathcal{N}_i^k} [\tau_{il,h}]^\alpha [\eta_{il}]^\beta} = p^k_{ij,h}.
\end{aligned}$$

□

Theorem 1. *The ant system algorithm for the traveling salesman problem is weakly-invariant, provided that the pheromone is initialized as prescribed by Definition 9.*

Proof. Let us consider two generic instances I and $\bar{I}$ such that

$$\bar{I} = fI, \text{ with } f > 0.$$

The theorem is proved by induction: We show that if at the generic iteration h some set of conditions $\mathcal{C}$ holds, then the solutions generated for the two instances I and $\bar{I}$ are the same and the set of conditions $\mathcal{C}$ also holds for the following iteration $h + 1$. The proof is concluded by showing that $\mathcal{C}$ holds for the very first iteration. With few minor modifications, this technique is adopted in the following for proving all theorems enunciated in the paper.

According to Definition 6, and taking into account (1), it results:

$$\bar{\eta}_{ij} = \frac{1}{f} \eta_{ij}, \text{ for all } \langle i, j \rangle.$$

According to Lemma 1, if at the generic iteration h, $\bar{\tau}_{ij,h} = \frac{1}{f}\tau_{ij,h}$, for all $\langle i,j \rangle$, then $\bar{p}^k_{ij,h} = p^k_{ij,h}$, for all $\langle i,j \rangle$. Under Hypothesis 1,

$$\bar{T}^k_h = T^k_h, \text{ for all } k = 1, \ldots, m,$$

and therefore, according to (2),

$$\bar{C}^k_h = f C^k_h, \text{ for all } k = 1, \ldots, m.$$

According to (3):

$$\begin{aligned}
\bar{\Delta}^k_{ij,h} &= \begin{cases} 1/\bar{C}^k_h, & \text{if } \langle i,j \rangle \in \bar{T}^k_h; \\ 0, & \text{otherwise;} \end{cases} = \begin{cases} 1/fC^k_h, & \text{if } \langle i,j \rangle \in \bar{T}^k_h = T^k_h; \\ 0/f, & \text{otherwise;} \end{cases} \\
&= \frac{1}{f} \begin{cases} 1/C^k_h, & \text{if } \langle i,j \rangle \in T^k_h; \\ 0, & \text{otherwise;} \end{cases} = \frac{1}{f}\Delta^k_{ij,h},
\end{aligned}$$

and therefore, for any arc $\langle i,j \rangle$:

$$\begin{aligned}
\bar{\tau}_{ij,h+1} &= (1-\rho)\bar{\tau}_{ij,h} + \sum_{k=1}^{m} \bar{\Delta}^k_{ij,h} = (1-\rho)\frac{1}{f}\tau_{ij,h} + \sum_{k=1}^{m} \frac{1}{f}\Delta^k_{ij,h} \\
&= (1-\rho)\frac{1}{f}\tau_{ij,h} + \frac{1}{f}\sum_{k=1}^{m} \Delta^k_{ij,h} = \frac{1}{f}\left((1-\rho)\tau_{ij,h} + \sum_{k=1}^{m} \Delta^k_{ij,h} \right) = \frac{1}{f}\tau_{ij,h+1}.
\end{aligned}$$

In order to provide a basis for the above defined induction and therefore to conclude the proof, it is sufficient to observe that at the first iteration $h = 1$, the pheromone is initialized as:

$$\bar{\tau}_{ij,1} = \frac{m}{\bar{C}^{nn}} = \frac{m}{fC^{nn}} = \frac{1}{f}\tau_{ij,1}, \text{ for all } \langle i,j \rangle.$$

$\square$

Remark 3. Theorem 1 holds true for any way of initializing the pheromone, provided that for any two instances $\bar{I}$ and I such that $\bar{I} = fI$, $\bar{\tau}_{ij,1} = \frac{1}{f}\tau_{ij,1}$, for all $\langle i,j \rangle$.

Remark 4. Theorem 1 extends to the application of ant system to problems other than the traveling salesman problem, provided that the initialization of the pheromone is performed as prescribed in Remark 3 and for any two instances $\bar{I}$ and I such that $\bar{I} = fI$, with $f > 0$, there exists a coefficient $g > 0$ such that $[\bar{\eta}_{ij}]^\beta = [g\eta_{ij}]^\beta$, for all $\langle i,j \rangle$. In particular, it is worth pointing out here that one notable case in which this last condition is satisfied is when $\beta = 0$, that is, when no heuristic information is used.

4 Strongly-Invariant Ant System

A strongly invariant version of ant system (*si*AS) can be easily defined. For definiteness, we present here a version of *si*AS for the traveling salesman problem.

Definition 10 (Strongly-invariant heuristic information). *When solving the traveling salesman problem, the heuristic information* η_{ij} *is*

$$\eta_{ij} = \frac{C^{nn}}{nc_{ij}}, \text{ for all } \langle i, j \rangle. \tag{4}$$

where c_{ij} *is the cost of traveling from city* i *to city* j*,* n *is the number of cities, and* C^{nn} *is the cost of the solution* T^{nn} *obtained by the nearest-neighbor heuristic.*

Definition 11 (Strongly-invariant pheromone update rule). *The pheromone is updated using the same rule given in Definition 7, with the only difference that* $\Delta^k_{ij,h}$ *is given by:*

$$\Delta^k_{ij,h} = \begin{cases} C^{nn}/mC^k_h, & \text{if } \langle i, j \rangle \in T^k_h; \\ 0, & \text{otherwise;} \end{cases}$$

where C^{nn} *is the cost of the solution* T^{nn} *obtained by the nearest-neighbor heuristic and* m *is the number of ants generated at each iteration.*

Definition 12 (Strongly-invariant pheromone initialization). *At the first iteration* $h = 1$*, the pheromone on all arcs is initialized to the value:*

$$\tau_{ij,1} = 1, \text{ for all } \langle i, j \rangle.$$

Definition 13 (Strongly-invariant ant system). *The strongly-invariant ant system (siAS) is a variation of ant system. It shares with ant system the random proportional rule for the construction of solutions, but in siAS the heuristic values are set as in Definition 10, the pheromone is initialized according to Definition 12 and the update is performed according to Definition 11.*

Remark 5. In the definition of *si*AS given above, the nearest-neighbor heuristic has been adopted for generating a reference solution, the cost of which is then used for normalizing the cost of the solutions found by *si*AS. Any other algorithm could be used instead, provided that the solution it returns does not depend on the scale of the problem.

Remark 6. It is worth noting here that the presence of the term n in the denominator of the left hand side of (4) is not needed for obtaining an invariant heuristic information. It has been included for achieving another property. Indeed, η_{ij} as defined in (4) assumes values that do not depend on the size of the instance under analysis—that is, on the number n of cities. If this term were not present, since the numerator C^{nn} grows with n, η_{ij} would have been relatively larger in large instances and smaller in small ones.

Remark 7. Similarly, it should be noticed that by initializing the pheromone to $\tau_{ij,1} = 1/m$, for all $\langle i,j\rangle$, and by defining $\Delta^k_{ij,h}$ as:

$$\Delta^k_{ij,h} = \begin{cases} C^{nn}/C^k_h, & \text{if } \langle i,j\rangle \in T^k_h; \\ 0, & \text{otherwise}; \end{cases}$$

one would have obtained nonetheless an invariant algorithm. The advantage of the formulation given in Definitions 11 and 12 is that the magnitude of the pheromone deposited on the arcs does not depend on the number m of ants considered.

The strongly-invariant ant system is *functionally equivalent* to the original ant system, that is, the two algorithms produce the same sequence of solutions for any given instance, provided that the pheromone is properly initialized, their respective pseudo-random number generators are the same, and these generators are initialized with the same seed. Formal proofs of the functional equivalence of *si*AS and ant system and of the strong invariance of *si*AS, are given in [6].

5 Conclusions

We have formally proved that, contrary to what previously believed [1], ant system is invariant to the rescaling of problem instances. The same holds [6] for the two main other members of the ant colony optimization family of algorithms, namely, $\mathcal{MAX}$–$\mathcal{MIN}$ ant system and ant colony system.

Moreover, we have introduced *si*AS, which is a straightforward strongly-invariant version of ant system. In this respect, *si*AS is similar to the hyper-cube ant system [1] which is the first strongly-invariant version of ant system ever published in the literature. The main advantage of *si*AS over the hyper-cube ant system is that, while the latter is effectively a new algorithm, *si*AS is functionally equivalent to the original ant system. As a consequence, one can immediately extend to *si*AS all understanding previously acquired about ant system and all empirical results previously obtained. Following the strategy adopted in the definition of *si*AS, a strongly-invariant version of any ACO algorithm can be defined. In particular, *si*$\mathcal{MM}$AS and *si*ACS are introduced in [6]. These two algorithms are the strongly-invariant versions of $\mathcal{MAX}$–$\mathcal{MIN}$ ant system and ant colony system, respectively. Like *si*AS, also *si*$\mathcal{MM}$AS and *si*ACS are functionally equivalent to their original counterparts.

Acknowledgments. This work was supported by the ANTS project, an *Action de Recherche Concertée* funded by the Scientific Research Directorate of the French Community of Belgium. Marco Dorigo acknowledges support from the Belgian National Fund for Scientific Research (FNRS), of which he is a Research Director.

References

1. Blum, C., Dorigo, M.: The hyper-cube framework for ant colony optimization. IEEE Transactions on Systems, Man, and Cybernetics—Part B **34**(2) (2004) 1161–1172
2. Dorigo, M., Stützle, T.: Ant Colony Optimization. MIT Press, Cambridge, MA, USA (2004)
3. Dorigo, M., Maniezzo, V., Colorni, A.: The Ant System: An autocatalytic optimizing process. Technical Report 91-016 Revised, Dipartimento di Elettronica, Politecnico di Milano, Milano, Italy (1991)
4. Dorigo, M.: Ottimizzazione, apprendimento automatico, ed algoritmi basati su metafora naturale. PhD thesis, Politecnico di Milano, Milano, Italy (1992) In Italian.
5. Dorigo, M., Maniezzo, V., Colorni, A.: Ant System: Optimization by a colony of cooperating agents. IEEE Transactions on Systems, Man, and Cybernetics—Part B **26**(1) (1996) 29–41
6. Birattari, M., Pellegrini, P., Dorigo, M.: On the invariance of ant colony optimization. Technical Report TR/IRIDIA/2006-004, IRIDIA, Université Libre de Bruxelles, Brussels, Belgium (2006) Submitted for journal publication.
7. Stützle, T., Hoos, H.H.: The $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System and local search for the traveling salesman problem. In Bäck, T., Michalewicz, Z., Yao, X., eds.: Proceedings of the 1997 IEEE International Conference on Evolutionary Computation (ICEC'97), Piscataway, NJ, USA, IEEE Press (1997) 309–314
8. Stützle, T., Hoos, H.H.: $\mathcal{MAX}$–$\mathcal{MIN}$ ant system. Future Generation Computer Systems **16**(8) (2000) 889–914
9. Dorigo, M., Gambardella, L.M.: Ant Colony System: A cooperative learning approach to the traveling salesman problem. IEEE Transactions on Evolutionary Computation **1**(1) (1997) 53–66