Convergence of Ant Colony Optimization on First-Order Deceptive Systems

Yixin Chen
Washington University in St. Louis
Department of Computer Science & Engineering
St. Louis, MO 63130, USA
chen@cse.wustl.edu

Haiying Sun
Yangzhou University
Department of Computer Science
Yangzhou, 225009, China

Abstract

Deceptive problems have been considered difficult for ant colony optimization (ACO) and it was believed that ACO will fail to converge to global optima of deceptive problems. This paper presents a convergence analysis of ACO on deceptive systems.

This paper proves, for the first time, that ACO can achieve reachability convergence but not asymptotic convergence for a class of first order deceptive systems (FODS) without assuming a minimum pheromone at each iteration. Experimental results confirm the analysis.

1 Introduction

Ant colony optimization (ACO) is a popular method for hard discrete optimization problems [6, 5, 2, 9].

Although there are a huge amount of experimentations and variants of ACO, its theoretical foundation is still in its early development [5]. Recently, there has been an increased effort to deepen the understanding of the convergence behavior of ACO. Guijahr [7, 8] proves the convergence of a particular implementation of ACO called the graph-based ant system (GBAS). However, GBAS is quite different from common ACO implementations and its practical performance is unknown. Dorigo et al. shows the convergence of another class of ACO [10, 6], in which there is a lower bound τ_{min} to all pheromone values. Such a method is denoted as ACO$_{\tau_{min}}$. Typically, there are two types of convergence of a stochastic optimization algorithm:

- **Asymptotic convergence.** An algorithm has asymptotic convergence if $\lim_{t\to\infty} P_s(t) = 1$, where $P_s(t)$ is the probability that the algorithm generates an optimal solution in the t^{th} iteration.

- **Reachability convergence.** An algorithm has reachability convergence if $\lim_{t\to\infty} P_r(t) = 1$, where $P_r(t)$ is the probability that the algorithm generates an optimal solution at least once in the 1^{st} to t^{th} iterations.

Dorigo et al. [10, 6] show that ACO$_{\tau_{min}}$ achieves reachability convergence under certain assumptions. Dorigo et al. [10, 6] also show the asymptotic convergence of ACO$_{\tau_{min}(t)}$, in which the pheromone lower bound $\tau_{min}(t)$ changes over time, under the assumption that $\tau_{min}(t) = d/ln(t+1)$, where d is a constant.

The main limitation of studying ACO$_{\tau_{min}}$ and ACO$_{\tau_{min}(t)}$ is that they do not allow for an exponentially fast decrement of the pheromone trails resulted by using a constant evaporation factor, which is used by most ACO implementations. In this paper, we consider an ACO algorithm that uses the exponentially fast decrement of the pheromone and prove its reachability convergence.

Our results are established for a particular class of optimization problem, called the *n-bit trap problem*. The n-bit trap problem is deemed a difficult problem because it is a **first-order deceptive system (FODS)**. FODS are characterized by local optimal fix-points with large basins of attractions. It is known that both genetic algorithms (GA) and ACO may fall into the local optimal traps and fail to find the global optimal solution [1, 3, 4]. We present an ACO algorithm that follows closely to the most common ACO implementations and prove that our algorithm achieves reachability convergence for the n-bit trap problem.

We also show that the algorithm does not achieve asymptotic convergence. We prove that, the standard ACO algorithm with an exponential decrease of pheromone, cannot converge to optimal solutions of the n-bit trap problems. This result is meaningful as it provides a first explanation of the lack of asymptotic convergence of ACO on deceptive problems observed by many others [1, 3, 4].

Algorithm 1: Ant colony optimization (ACO)

```
initialize pheromone values τ_i^j for each i = 1, ··· , n
and j = 1, ··· , |D_i|;
foreach iteration t = 1, 2, ··· do
    foreach ant k = 1, 2, ··· , m do
        s ⟵ ConstructSolution(τ);
        if s is a feasible solution then
            if f(s) < f(s_b) or s_b =null then
                s_b ⟵ s;
            Ψ_t ⟵ Ψ_t ∪ {s};
    UpdatePheromone(τ, Ψ_t, s_b);
```

2 Ant Colony Optimization and Deceptive Problems

In this section, we introduce the framework of the ACO algorithm we will use and review the concept of deceptive problems. ACO are designed for solving constrained optimization (CO) problems defined as follows [5].

Definition 1 *A constrained optimization problem is defined by a model $\mathcal{P} = (\mathcal{S}, \Omega, f)$, where: 1) $\mathcal{S}$ is a search space defined over a finite set of discrete variables; 2) Ω is a set of constraints on the variables; and 3) $f : \mathcal{S} \to \mathbf{R}$ is the objective function to be maximized.*

In the search space $\mathcal{S}$, there are n decision variables $X_i, i = 1, \cdots, n$, where X_i can take values from the set $D_i = \{c_i^1, \cdots, c_i^{|D_i|}\}$. A variable assignment is written as $X_i = c_i^j$. A complete assignment to all X_i gives a solution instantiation. $\mathcal{S}$ is the set of all such complete assignments. We also denote the set of all solution components as $R = \{c_i^j | i = 1, \cdots, n, j = 1, \cdots, |D_i|\}$.

A solution $s \in \mathcal{S}$ is called a **feasible solution** if it satisfies all the constraints in Ω. A feasible solution s^* is a **global optimum** if $f(s^*) \geq f(s)$ for all $s \in \mathcal{S}$. The set of all global optima is denoted by $\mathcal{S}^* \subseteq \mathcal{S}$. The general ACO algorithm considered in this paper is shown in Algorithm 1. In the ACO algorithm, we assign a value τ_i^j, called the **pheromone**, to each of the solution component c_i^j. The vector of all pheromone is denoted by τ.

The central component of ACO is the ConstructSolution(τ) procedure that each ant uses to construct a solution. This construction procedure assembles solutions as sequences of elements from the finite set of solution components $\mathcal{S}$. Let s^p be the partial solution constructed by an ant. Initially $s^p = \{\}$ is empty. The ant at each construction step adds a feasible solution component from the set $R(s^p)$, where $R(s^p) \subseteq R \setminus \{s^p\}$ is the set of feasible solution components that satisfy the constraints in Ω given the partial solution s^p. When selecting the component for variable X_i, the probability that the ant chooses c_i^j is, $\forall j = 1, \cdots, |D_i|$,

$$P(c_i^j | s^p) = \frac{[\tau_i^j]^\alpha [\eta(c_i^j)]^\beta}{\sum_{c_i^k \in R(s^p)} [\tau_i^k]^\alpha [\eta(c_i^k)]^\beta}. \tag{1}$$

Here, $\eta(c_i^j)$ is a heuristic function for the component c_i^j. $\alpha > 0$ and $\beta > 0$ are two parameters controlling the relative importance of the pheromone and heuristic information.

In iteration t, after all ants have constructed solutions, ACO calls UpdatePheromone(τ, Ψ_t, s_b) to update the pheromone vector τ. Here, Ψ_t is the set of solutions constructed by the ants in iteration t and s_b is the incumbent best solution. The **pheromone update rule** is, $\forall i = 1, \cdots, n, j = 1, \cdots, |D_i|$,

$$\tau_i^j = \rho \tau_i^j + \frac{1}{|S_i^j|} \sum_{s \in S_i^j} F(s), \tag{2}$$

where $\rho \in (0, 1)$ is an **evaporation factor** and S_i^j is the set of solutions in Ψ_t that have c_i^j as a component, and $F : \mathcal{S} \to \mathbf{R}^+$ is a **quality function** such that, for any $s, s' \in \mathcal{S}$, if $f(s) > f(s')$, then $F(s) \geq F(s')$.

The **expected iteration quality** is denoted as $W_F(\tau|t)$ where t is the iteration number. $W_F(\tau|t)$ is defined as:

$$W_F(\tau|t) = \sum_{s \in \mathcal{S}} F(s) P(s|\tau), \tag{3}$$

where $P(s|\tau)$ is the probability that the solution s is generated by an ant given the pheromone vector τ.

Definition 2 *Given a constrained optimization problem P, an ACO algorithm is a **local optimizer** for P if for any initial pheromone values, the expected iteration quality satisfies:*

$$W_F(\tau|t+1) \geq W_F(\tau|t), \quad \forall t \geq 0. \tag{4}$$

We can now introduce the notion of deception for local optimizers [3].

Definition 3 *Given a constrained optimization problem P, it is called a **first-order deceptive system (FODS)** for an ACO algorithm, if the ACO algorithm is a local optimizer, and there exists an initial setting of pheromone values such that the algorithm does not in expectation converge to a global optimum.*

3 Convergence Analysis of ACO on FODS

Currently, whether a problem is a FODS is mostly proved by empirical studies, and lacks theoretical analysis. For example, experiments have been run to prove that the n-bit trap and job-shop scheduling problems are deceptive

s	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
bit 4	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
bit 3	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
bit 2	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
bit 1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
$f(s)$	5	1	1	2	1	2	2	3	1	2	2	3	2	3	3	4

Table 1. Fitness function value of the 4-bit trap problem.

systems by showing that the ACO does not converge to the global optimal solutions [3]. However, there is no proof as to why the ACO cannot achieve asymptotic convergence for such problems. Also, it is yet to be determined if the ACO can achieve reachability convergence on deceptive systems. Further, understanding the time complexity for solving deceptive problems is important because that it sheds insights into the worst-case performance of ACO, as the deceptive problems are harder problems for ACO. In this paper, we study these issues on the n-bit trap problem, a well-known example of FODS that has also been studied for evolutionary algorithms.

3.1 The n-bit trap problem

The n-bit trap problem is to find among the 2^n binary numbers, from 0 to $2^n - 1$, the one with the highest fitness. The fitness of a binary number s is defined as:

$$f(s) = \begin{cases} h(s) & s \neq 0 \\ n+1 & s = 0 \end{cases}, \tag{5}$$

where $h(s)$ is the Hamming distance between s and 0. Obviously, the global optimum is $s^* = 0$.

To solve the n-bit trap problem, we design the following ACO algorithm which fits into the general ACO framework in Algorithm 1. We use m ants. In each iteration, an ant sequentially fixes the j^{th} bit of the binary number in order of $j = 1, 2, \cdots, n$. For the j^{th} bit, the ant has two choices c_j^0 and c_j^1, corresponding to setting the j^{th} bit to 0 and 1, respectively. The pheromone of c_j^0 and c_j^1 are τ_j^0 and τ_j^1, respectively.

Define the set $G_j^k = \{(b_1 b_2 \cdots b_n) \in \mathcal{S} | b_j = k\}$, where $k \in \{0, 1\}$ and b_j is the value of the j^{th} bit. Hence, G_j^k is the set of binary numbers whose j^{th} bit is k. We define the function $F(G_j^k)$ as the average fitness of all the numbers in G_j^k:

$$F(G_j^k) = \frac{1}{|G_j^k|} \sum_{s \in G_j^k} f(s). \tag{6}$$

For example, when $n = 4$, we have $F(G_j^0) = 2.125$ and $F(G_j^1) = 2.5$ for all $j = 1, 2, 3, 4$.

We initialize pheromone τ_j^k as $\tau_j^k(0) = F(G_j^k)$, for all $j = 1, \cdots, n$ and $k \in \{0, 1\}$. For the j^{th} bit, the probability an ant selects value k, $k = 0$ or 1, is:

$$P(c_j^k, t) = \frac{\tau_j^k(t)}{\tau_j^0(t) + \tau_j^1(t)}, \tag{7}$$

which is derived from (1) by setting $\alpha = 1$ and $\beta = 0$. In each iteration, the pheromone is updated as:

$$\tau_j^k(t+1) = \rho \tau_j^k(t) + \sum_{s \in S_j^k} f(s), \tag{8}$$

where S_j^k is the set of solutions generated at the t^{th} iteration that have c_j^k as the j^{th} bit.

3.2 Asymptotic convergence of ACO on the n-bit trap problem

Lemma 1 *For the n-bit trap problem and $F(G_j^k)$ function defined in (6), we have:*

$$F(G_j^1) = \frac{n+1}{2} \ \ and \ \ F(G_j^0) = \frac{n-1}{2} + \frac{n+1}{2^{n-1}}. \tag{9}$$

Proof. We have from the definition in (6) that:

$$F(G_j^1) = \frac{1}{2^{n-1}} \sum_{k=1}^{n} C_{n-1}^{k-1} k = \frac{n+1}{2} \tag{10}$$

$$\begin{aligned} F(G_j^0) &= \frac{1}{2^{n-1}}((n+1) + \sum_{k=1}^{n-1} C_{n-1}^{k} k) \\ &= \frac{n-1}{2} + \frac{n+1}{2^{n-1}} \end{aligned} \tag{11}$$

■

Lemma 2 *Let $\beta > 0$ be a constant, we have*

$$\lim_{T \to \infty} \prod_{t=\lceil \beta \rceil + 1}^{T} (1 - \frac{\beta}{t}) = 0 \tag{12}$$

Proof. We consider four cases.

1) In case $\beta = 1$, we have

$$\begin{aligned} \lim_{T \to \infty} \prod_{t=\lceil \beta \rceil + 1}^{T} (1 - \frac{\beta}{t}) &= \lim_{T \to \infty} \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdots \frac{T-1}{T} \\ &= \lim_{T \to \infty} \frac{1}{T} = 0. \end{aligned} \tag{13}$$

2) In case $\beta > 1$, we have $\prod_{t=\lceil \beta \rceil + 1}^{T} (1 - \frac{\beta}{t}) < \prod_{t=\lceil \beta \rceil + 1}^{T} (1 - \frac{1}{t})$. Let $\prod_{t=2}^{\lceil \beta \rceil} (1 - \frac{1}{t}) = r$, we have:

$$r \prod_{t=\lceil \beta \rceil + 1}^{T} (1 - \frac{\beta}{t}) = \prod_{t=2}^{T} (1 - \frac{1}{t}). \tag{14}$$

Again, because $\lim_{T\to\infty}\prod_{t=2}^{T}(1-\frac{1}{t}) = 0$, we have $\lim_{T\to\infty} r\prod_{t=\lceil\beta\rceil+1}^{T}(1-\frac{\beta}{t}) = 0$ which means $\lim_{T\to\infty}\prod_{t=\lceil\beta\rceil+1}^{T}(1-\frac{\beta}{t}) = 0$.

3) In case $\beta < 1$, and $\beta = \frac{1}{q}$ where q is an integer larger than 1. Let $T = p\cdot q$ and $c = \prod_{t=2}^{q-1}(1-\frac{1}{t})$, then we have:

$$\begin{aligned}\prod_{t=2}^{T}(1-\frac{1}{t}) &= c\prod_{t=q}^{T}(1-\frac{1}{t})\\ &= c\prod_{j=1}^{p-1}\prod_{i=0}^{q-1}(1-\frac{1}{jq+i}) > c\prod_{j=1}^{p-1}(1-\frac{1}{jq})^q.\end{aligned}$$

Therefore,

$$\begin{aligned}&\lim_{T\to\infty}\prod_{t=2}^{T}(1-\tfrac{1}{t}) > c\lim_{p\to\infty}\prod_{j=1}^{p}(1-\tfrac{1}{jq})^q\\ &\quad= c\Big[\lim_{p\to\infty}\prod_{j=1}^{p}(1-\tfrac{1}{jq})\Big]^q = 0 \qquad (15)\end{aligned}$$

Since we know $\lim_{T\to\infty}\prod_{t=2}^{T}(1-\frac{1}{t}) = 0$ and and $c\Big[\lim_{p\to\infty}\prod_{j=1}^{p}(1-\frac{1}{jq})\Big]^q \geq 0$, combined with (15) we get $c\Big[\lim_{p\to\infty}\prod_{j=1}^{p}(1-\frac{1}{jq})\Big]^q = 0$, which implies $\lim_{p\to\infty}\prod_{j=1}^{p}(1-\frac{1}{jq}) = 0$, which is equivalent to:

$$\lim_{T\to\infty}\prod_{t=1}^{T}(1-\frac{1}{qt}) = \lim_{T\to\infty}\prod_{t=1}^{T}(1-\frac{\beta}{t}) = 0.$$

4) In case $\beta < 1$, and $\beta = \frac{p}{r}$ where p and r are both integers larger than 1. We can always find $\beta' = \frac{1}{q}$ where q is an integer larger than 1 such that $\beta' < \beta$. Since $\prod_{t=1}^{T}(1-\frac{\beta'}{t}) > \prod_{t=1}^{T}(1-\frac{\beta}{t})$ and $\lim_{T\to\infty}\prod_{t=1}^{T}(1-\frac{\beta'}{t}) = 0$ from case 3), we have $\lim_{T\to\infty}\prod_{t=1}^{T}(1-\frac{\beta}{t}) = 0$. ■

We consider the ACO algorithm described in Algorithm 1 using (7) and (8) as the selection and update functions, respectively. We name this algorithm ACO_{n-bit}. For the ACO_{n-bit} algorithm, define:

$$\alpha_j(t) = \frac{\tau_j^0(t)}{\tau_j^1(t)}, \quad \forall j = 1,2,\cdots,n, t = 0,1,2,\cdots. \quad (16)$$

Namely, $\alpha_j(t)$ is the ratio of the pheromone on value 0 to the pheromone on value 1 for bit j in iteration t.

Theorem 1 *The ACO_{n-bit} algorithm cannot achieve asymptotic convergence on the n-bit trap problem when $n > 3$.*

Proof. For the n-bit trap problem, since the global optimal solution is $s^* = \{0,0,\cdots,0)$, we show that $\forall j = 1,2,\cdots,n$,

$$\lim_{t\to\infty} E[\alpha_j(t)] = \lim_{t\to\infty} E\left[\frac{\tau_j^0(t)}{\tau_j^1(t)}\right] = 0. \quad (17)$$

Suppose we have m ants. We first prove by induction that

$$E[\alpha_j(t)] < 1, \quad \forall j = 1,2,\cdots,n, t = 0,1,2,\cdots. \quad (18)$$

When $t = 0$, $\tau_j^0(0)$ is initialized to $F(G_j^0)$ and $\tau_j^1(0)$ is initialized to $F(G_j^1)$. From Lemma 1, while $n > 3$, $F(G_j^1) > F(G_j^0)$. For a fixed n, we set $F(G_j^0) = \alpha_0 F(G_j^1)$, where α_0 is a constant and $0 < \alpha_0 < 1$. Hence, we have our base case:

$$E[\alpha_j(0)] = \frac{\tau_j^0(0)}{\tau_j^1(0)} = \frac{F(G_j^0)}{F(G_j^1)} = \alpha_0 < 1.$$

To prove (18) by induction, we assume that $E[\alpha_j(t)] < 1$. According to (8), we have that in the $(t+1)^{th}$ iteration, the expected value of pheromone satisfies

$$\begin{aligned}E[\tau_j^0(t+1)] &\leq \rho E[\tau_j^0(t)] + E\left[\frac{m\alpha_j(t)}{1+\alpha_j(t)}F(G_j^0)\right],\\ E[\tau_j^1(t+1)] &\geq \rho E[\tau_j^1(t)] + E\left[\frac{m}{1+\alpha_j(t)}F(G_j^1)\right].\end{aligned}$$

Given (16), we have

$$\begin{aligned}&E[\alpha_j(t+1)] = \tfrac{E[\tau_j^0(t+1)]}{E[\tau_j^1(t+1)]}\\ &\leq \frac{\rho E[\tau_j^0(t)] + \frac{mE[\alpha_j(t)]}{1+E[\alpha_j(t)]}F(G_j^0)}{\rho E[\tau_j^1(t)] + \frac{m}{1+E[\alpha_j(t)]}F(G_j^1)}\\ &= \frac{\rho(1+E[\alpha_j(t)])E[\alpha_j(t)]E[\tau_j^1(t)] + E[\alpha_j(t)]mF(G_j^0)}{\rho(1+E[\alpha_j(t)])E[\tau_j^1(t)] + mF(G_j^1)}\end{aligned}$$

Hence, we have

$$\begin{aligned}&E[\alpha_j(t+1)]\\ =\;& E[\alpha_j(t)]\left(1 - \frac{(1-\alpha_0)mF(G_j^1)}{\rho(1+E[\alpha_j(t)])E[\tau_j^1(t)] + mF(G_j^1)}\right)\\ <\;& E[\alpha_j(t)].\end{aligned}$$

Therefore, we get, for any $j = 1,\cdots,n$,

$$\begin{aligned}1 &> \alpha_0 = E[\alpha_j(0)] > E[\alpha_j(1)] > \cdots\\ &> E[\alpha_j(t)] > E[\alpha_j(t+1)] > \cdots > 0. \quad (19)\end{aligned}$$

Thus, we have proved (18) by induction.

Further, since

$$\begin{aligned}E[\tau_j^1(t)] &= \rho E[\tau_j^1(t-1)] + \frac{1}{1+E[\alpha_j(t)]}mF(G_j^1)\\ &< E[\tau_j^1(t-1)] + mF(G_j^1), \quad (20)\end{aligned}$$

recursively applying (20) leads to

$$E[\tau_j^1(t)] < E[\tau_j^1(0)] + tmF(G_j^1) < (t+1)mF(G_j^1). \quad (21)$$

Substituting (21) into (19) yields

$$E[\alpha_j(t+1)] < E[\alpha_j(t)]\left(1 - \frac{(1-\alpha_0)}{\rho(1+E[\alpha_j(t)])(t+1)+1}\right)$$

From (19), we know $1 + E[\alpha_j(t)] < 2$. Therefore,

$$\begin{aligned} E[\alpha_j(t+1)] &< E[\alpha_j(t)]\left(1 - \tfrac{(1-\alpha_0)}{2\rho(t+1)+1}\right) \\ &< E[\alpha_j(t)]\left(1 - \tfrac{(1-\alpha_0)}{(2\rho+3)t}\right) \end{aligned}$$

Let $\beta = \frac{1-\alpha_0}{2\rho+3}$, we have

$$E[\alpha_j(t+1)] < E[\alpha_j(t)]\left(1 - \frac{\beta}{t}\right) < \prod_{i=1}^{t}\left(1 - \frac{\beta}{i}\right)\alpha_0$$

From Lemma 2, we get $\lim_{t\to\infty} E[\alpha_j(t)] = 0$. Therefore, the ACO_{n-bit} algorithm will not converge to the optimal solution $s^* = \{0, 0, \cdots, 0\}$. ■

Theorem 1 gives an explanation why the ACO algorithms have difficulties in solving deceptive problems such as the n-bit trap problem.

3.3 Reachability convergence of ACO on the n-bit trap problem

Although ACO_{n-bit} cannot achieve asymptotic convergence, we will show that it can achieve reachability convergence. We need some preliminary results first.

Lemma 3 *The pheromone in the t^{th} iteration of ACO_{n-bit} satisfies $\tau_j^0(t) \geq \frac{1-\rho^{t+1}}{1-\rho}\varphi F(G_j^0)$, where φ is a constant, $0 < \varphi < 1$, for $j = 1, 2, \cdots, n$ and $t = 1, 2, \cdots$.*

Proof. From the pheromone update rule in (8), we have

$$\tau_j^k(t) = \rho\tau_j^k(t-1) + \sum_{s\in S_j^k} f(s), \tag{22}$$

Let $\theta_j^0 = \min\left\{f(s)|s = (b_1b_2\cdots b_n), s \in \mathcal{S}, b_j = 0\right\}$ and $\theta_j^0 = \varphi F(G_j^0)$. Obviously, $0 < \varphi < 1$. Since $\tau_j^0(t) \geq \rho\tau_j^0(t-1) + \theta_j^0$, we have $\tau_j^0(t) \geq \rho\tau_j^0(t-1) + \varphi F(G_j^0)$. Therefore, we get

$$\begin{aligned} \tau_j^0(t) &\geq \rho^t\tau_j^0(0) + (\rho^{t-1} + \rho^{t-2} + \cdots + 1)\varphi F(G_j^0) \\ &= (\rho^t + \rho^{t-1} + \rho^{t-2} + \cdots + 1)\varphi F(G_j^0) \\ &= \frac{1-\rho^{t+1}}{1-\rho}\varphi F(G_j^0). \end{aligned}$$

Lemma 4 *For any positive integer n and a real number $\rho \in (0, 1)$, there exists a positive integer t_0 such that $\frac{1-\rho^{t+1}}{1-\rho} > t^{-\frac{1}{n}}$ for any $t > t_0$.*

Proof. We see that fact that

$$\lim_{t\to\infty} \frac{t^{-\frac{1}{n}}}{\frac{1-\rho^{t+1}}{1-\rho}} = \lim_{t\to\infty} \frac{1-\rho}{(1-\rho^{t+1})\sqrt[n]{t}} = 0.$$

Thus, there exists an integer $t_0 > 0$ such that $\frac{1-\rho^{t+1}}{1-\rho} > t^{-\frac{1}{n}}$ for $t > t_0$. ■

From Lemma 3 and Lemma 4, we see that when t is large enough, $\tau_j^0(t) \geq t^{-\frac{1}{n}}\varphi F(G_j^0)$. We set $\tau_{min}^0(j,t) = t^{-\frac{1}{n}}\varphi F(G_j^0)$ and $\tau_{min}^0(t) = \min_{1\leq j\leq n}\tau_{min}^0(j,t)$. If we set $\theta_{min} = \min_{1\leq j\leq n}\varphi F(G_j^0)$, then from Lemma 3 and Lemma 4,

$$\tau_{min}^0(t) \geq t^{-\frac{1}{n}}\theta_{min}, \quad \forall t > t_0. \tag{23}$$

Lemma 5 *The pheromone in the t^{th} iteration of ACO_{n-bit} satisfies $\tau_j^k(t) \leq \frac{1}{1-\rho}F(G_j^k)$, for $k = 0, 1$, $j = 1, 2, \cdots, n$ and $t = 1, 2, \cdots$.*

Proof. From the pheromone update rule in (8), we have

$$\tau_j^k(t) = \rho\tau_j^k(t-1) + \sum_{s\in S_j^k} f(s), \tag{24}$$

Let $\lambda_j^k = \max\left\{f(s)|s = (b_1b_2\cdots b_n), s \in \mathcal{S}, b_j = k\right\}$ and $\lambda_j^k = \phi F(G_j^k)$ where $\phi > 1$ is a constant for a given n. Since $\tau_j^k(t) \leq \rho\tau_j^k(t-1) + m\lambda_j^k$, we have $\tau_j^k(t) \leq \rho\tau_j^k(t-1) + \phi mF(G_j^k)$. Therefore, we get

$$\begin{aligned} \tau_j^k(t) &\leq \rho^t\tau_j^k(0) + (\rho^{t-1} + \rho^{t-2} + \cdots + 1)\phi mF(G_j^k) \\ &\leq \frac{1}{1-\rho}\phi mF(G_j^k). \end{aligned}$$ ■

We denote $\tau_{max}^k(j) = \frac{1}{1-\rho}\phi mF(G_j^k)$ and $\tau_{max} = \max_{1\leq j\leq n, k=0,1}\tau_{max}^k(j)$.

Theorem 2 *The ACO_{n-bit} algorithm achieves reachability convergence on the n-bit trap problem.*

Proof. Although there are multiple ants, we only need to prove that one ant can achieve reachability convergence in order to establish the result. From Lemma 4 and Lemma 5, we know that for $t > t_0$, where t_0 is defined in Lemma 4,

$$P(c_j^0, t) = \frac{\tau_j^0(t)}{\tau_j^0(t) + \tau_j^1(t)} \geq \frac{\tau_{min}^0(t)}{2\tau_{max}}. \tag{25}$$

Hence, let $P(t)$ be the probability that the ant can generate the optimal solution $s^* = (0, 0, \cdots, 0)$ in iteration t, we have

$$P(t) \geq \left[\frac{\tau_{min}^0(t)}{2\tau_{max}}\right]^n \tag{26}$$

Let $P_{succ}(T)$ and $P_{fail}(T)$, respectively, be the probability that the ant does and does not find s^* in the first T iterations. From (23) and Lemma 5, we know

$$\begin{aligned} P_{fail}(T) &= \prod_{t=1}^{T}(1 - P(t)) \leq \prod_{t=t_0}^{T}\left[1 - \left(\frac{\tau_{min}^0(t)}{2\tau_{max}}\right)^n\right] \\ &\leq \prod_{t=t_0}^{T}\left[1 - \left(\frac{\theta_{min}}{2\tau_{max}}\right)^n\left(t^{-\frac{1}{n}}\right)^n\right]. \end{aligned}$$

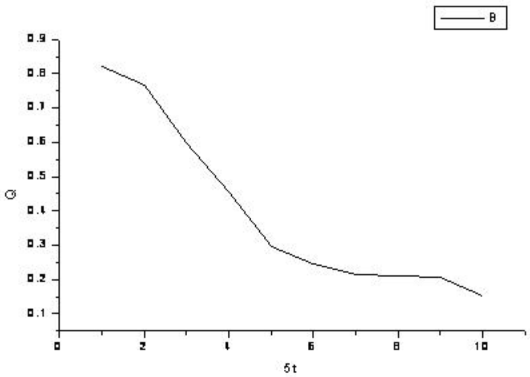

Figure 1. Average solution pheromone.

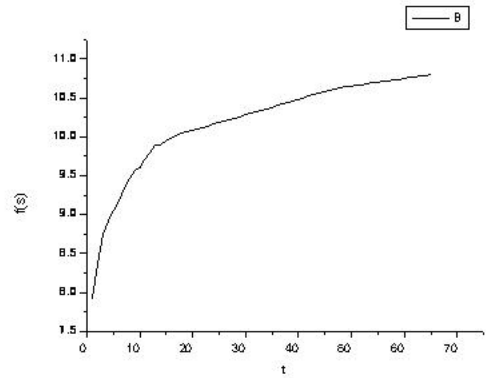

Figure 2. Average incumbent solution.

Denote $(\frac{\theta_{min}}{2\tau_{max}})^n = \gamma$, then $\lim_{T\to\infty} P_{fail}(T) \leq \lim_{T\to\infty} \prod_{t=t_0}^{T} [1 - \frac{\gamma}{t}]$. From Lemma 2, we have that $\lim_{T\to\infty} P_{fail}(T) = 0$ and thus $\lim_{T\to\infty} P_{succ}(T) = 1 - \lim_{T\to\infty} P_{fail}(T) = 1$. ■

4 Experimental Results

We have implemented an ACO on the n-bit trap problem. We set $n = 11$, evaporation rate $\rho = 0.95$, and use 20 ants. First, to test the asymptotic convergence, we make 100 runs and record Q, the solution pheromone at each iteration t. Figure 1 shows the average pheromone at each iteration of the 100 runs. We see that Q decreases as t increases, indicating that ACO does not have asymptotic convergence on the n-bit trap problem. Next, to show reachability convergence, we plot in Figure 2 the average fitness of the incumbent solution f versus the iteration number t. We see that the average f approaches the optimal value 11 as t increases, showing the reachability convergence of ACO.

5 Conclusions

In this paper, we have theoretically analyzed ACO algorithms on deceptive problems, a class of problems that are considered difficult for ACO. We have presented a first attempt towards a convergence analysis of ACO on deceptive systems. Using the n-trap problem as an example of first-order deceptive problems, we have proved that that ACO can achieve reachability convergence but not asymptotic convergence for some first order deceptive systems without assuming a minimum pheromone at each iteration. We have also presented experimental results that confirm the analysis.

References

[1] C. Blum. Theoretical and practical aspects of ant colony optimization. *Dissertations in Artificial Intelligence*, 2463, 2004.

[2] C. Blum. Ant colony optimization: Introduction and recent trends. *Physics of Life Reviews*, 2(4):353–373, December 2005.

[3] C. Blum and M. Dorigo. Deception in ant colony optimization. In *Proceedings of ANTS2004, Lecture Notes in Computer Science*, pages 119–130, 2004.

[4] C. Blum and M. Dorigo. Search bias in ant colony optimization: on the role of competition-balanced systems. *IEEE Transactions on Evolutionary Computation*, 9:159–174, 2005.

[5] M. Dorigo and C. Blum. Ant colony optimization theory: A survey. *Theoretical Computer Science*, 344:243–278, 2005.

[6] M. Dorigo and T. Stützle. *Ant Colony Optimization*. MIT Press, 2004.

[7] W. Gutjahr. A Graph-based Ant System and its convergence. *Applied Soft Computing*, 16:873–888, 2000.

[8] W. Gutjahr. ACO algorithms with guaranteed convergence to the optimal solution. *Info. Processing Lett.*, 82:145–153, 2002.

[9] S. Shtovba. Ant algorithms: Theory and applications. *Programming and Computer Software*, 31:167–178, 2005.

[10] T. Stützle and M. Dorigo. A Short Convergence Proof for a Class of Ant Colony Optimization Algorithms. *IEEE Transactions on Evolutionary Computation*, 6:358–365, 2002.