ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 4, No. 4, 2009, *pp*. 259-264

Ant Colony Optimization for Optimal Control Problems

Akbar H. Borzabadi [1, +] and Hamed H. Mehne [2]

[1] Dept. of Appl. Math., Damghan Univ. of Basic Sciences, 36715-364, Iran.

[2] Aerospace Research Institute, Tehran, 15875-3885, Iran. E-mail: hashemi@math.um.ac.ir

(*Received December 26, 2008, accepted April 17, 2009*)

Abstract. Ant algorithm optimization is used to solve a general class of optimal control problems with single input control. The discretized form of the optimal control problem is converted to a quasi quadratic assignment problem in the time-control space. By applying the ant optimization algorithm on this problem, a piece-wise constant approximation is obtained for the optimal control. Implementation of the method and numerical results are also discussed.

Keywords: Meta-heuristic algorithms, ant colony optimization, optimal control problems.

1. Introduction

In early decades optimal control theory as one of the most applicable and technological issues has been taken into consideration. The analytical solutions for problems of optimal control are not always available. Thus to find approximate solution is the most logical way to solve them. To this end, various approaches such as discretization [14], measure theory [10], polynomial parametrization [9, 8], etc., have been proposed. Some heuristic algorithms such as genetic algorithms [7] have been also applied to solve optimal control problems (OCP's). Our aim is to apply Ant Colony Optimization (ACO) method to construct approximate optimal control function for a general class of OCPs. To implement this method we first discretize the time-control space. This control discretization enables us to examine different choices of controls to find the optimal one in assignment of constant controls to sub-intervals. This assignment nature of the problem encourages us using a meta heuristic as ACO method to solve the problem. The advantages may be encounter as the method is self-starting i.e. it doesn't need any approximate solution to be started, and the type of dynamical system and performance index doesn't have serious effect on the method because it uses direct evaluations of controls.

Metaheuristics incorporate concepts from very different fields such as genetics, biology, artificial intelligence, mathematics and physics and neuro-science among others. Examples of methahuristics include simulated annealing, tabu search, iterated local search, variable neighborhood search algorithms, greedy randomized adaptive search procedures and evolutionary algorithms.

ACO is currently one of the best available meta-heuristic for some problems and is amongthe most competitive approaches for discrete optimization problems [1, 4, 6]. The essential framework of the ACO is search over several constructive computational threads, based on a memory structure incorporating the information about the effectiveness of previously obtained fragments of solutions. This structure is maintained dynamically by deposit, evaporation and detection of conceptual pheromone.

Several algorithms have been proposed in the literature following the ACO metaheuristic, (see [3]). The first ACO algorithm, called Ant System (AS) [6], was initially proposed by Dorigo et al. and then this algorithm was applied to the well-known traveling salesman problem as a benchmark problem [5]. AS has been the prototype of many following ACO algorithms with which many other NP-hard combinatorial optimization problems can be solved successfully. Ant algorithms have also been applied to the Facilities Layout Problem which can be shown to be a Quadratic Assignment Problem (QAP).

[+] Corresponding author. Tel.: +98-232-5235316; fax: +98-232-5235316.
E-mail address: borzabadi@dubs.ac.ir

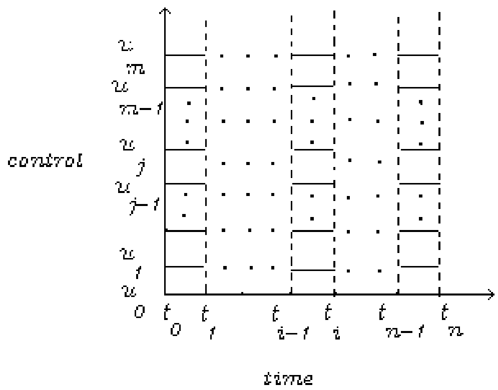

Fig. 1. A typical control function in time-control discretization

2. Discretized Optimal Control Problem

There are many examples in science and engineering involving optimal control problems. Mathematically, an OCP's deals with optimizing a performance index having the pair state control satisfying a dynamical system. There may be also bounding conditions on the state and some restrictions on control. Different types of the performance index and dynamical systems accompanied with various bounding conditions lead to a wide variety of definitions for OCPs. Exact and numerical methods for solving OCPs are highly dependent to the form of the problem. As we are going to develop a numerical method for solving general OCPs, let consider the minimizing of a performance index like

$$I(x(t),u(t)) = \int_0^{t_f} f_\circ(t,x(t),u(t))dt \tag{1}$$

where $t_f > 0$ is given and $f_\circ$ is an integrable function with no restriction about linearity and differentiability. The state and control satisfy a dynamical system as

$$\dot{x}(t) = g(t,x(t),u(t)), \quad t \in (0,t_f) \tag{2}$$

with $x(0) = x_0$ and $x(t_f) = x_f$ as initial and final given conditions. The single valued control function gives its values from a known interval $[u_\alpha, u_\beta]$.

To find the optimal solution we must examine the performance index in the set of all possibilities of control-state pairs. This set is called the set of admissible pairs consisting of pairs like (x,u) satisfying in (2) and other mentioned conditions. If we choose a control function u and solve (2) with initial conditions, then resulting state may not reach to x_0 at $t = t_f$ and a miss distance between $x(t_f)$ and x_0 is introduced. Now if the norm of miss distance is added to the performance index as a penalty, then minimizing $I + M\|x(t_f) - x_0\|$ forces the control to produce an admissible state. This enables us to reduce the admissible set of control-states to the admissible set of controls only. So we could search for the optimal solution in the set of all controls. This process of constructing optimal solutions from control function is a popular method in optimal control theory which appears in literature under control parametrization [9, 15] and control discretization [10].

Here we develop a control discretization based method where the time interval is divided to n sub-interval $[t_0,t_1],[t_1,t_2],\cdots,[t_{n-1},t_n]$. On the other hand the set of control values is divided to constant values $u_1,u_2,\cdots,u_m$. In this way the time-control space is discretized if the control function assumes to be constant at each time sub-interval. A typical dicsretization is given in Fig.1 with $n = 7$ and $m = 6$. The bold pattern in this picture shows a control function.

3. Converting OCP to AS

Discretization proposes to consider control function as a sequence of u_j segments corresponding to time sub-intervals. Now a trivial way to find the nearly optimal solution is to calculate all possible patterns and compare the corresponding trade offs. This trivial method of total enumeration needs m^n evaluation. Avoiding of such a huge number of computations, we introduce a method of evaluating special patterns guiding us to the optimal one. The main drawback of the total enumeration is that the method evaluates all of the control patterns independently i.e. the evaluated performance index of the current pattern doesn't have a role in construction the next pattern. With AS approach we construct patterns based on the performance index of pervious iterations leading to a method with computations less than the total enumeration.

For converting the OCP to AS we use a similar framework of solving QAP by ACO. In fact we decide to assign for every interval $[t_{i-1}, t_i], i = 1, 2, \cdots, n,$, a constant $u_k \in \{u_0, u_1, \cdots, u_m\}$.

Special form of our problem suggests us to use another version of ACO, called Max-Min Ant System (MMAS). This method, the first used to solve TSP ([11, 12]) and then used for solving QAP in [12]. In fact this method is one of the best performing extensions of AS. It extends the basic AS in the following aspects which is quoted from [3]:

1. After each tour, updating of trial will be done by an ant, i.e. that ant which obtains the best solution in currently tour or the best solution from the first tour until current tour. After all ants have constructed a solution, first every pheromone trial is evaporated:

$$\tau_{rs} \leftarrow (1-\rho)\tau_{rs},$$

and next pheromone is deposited according to:

$$\tau_{rs} \leftarrow \tau_{rs} + f(C(S_{best})), \ \forall a_{rs} \in S_{best},$$

where $f(C(S))$ is the amount of pheromone released depends on the quality $C(S)$, $C(S)$ is a cost associated to each solution S, S_{best} is the best solution and $\rho \in (0,1]$ is evaporation rate. The best ant that is allowed to add pheromone may be the iteration-best or the global best solution. Experimental results have shown that the best performance is obtained by gradually increasing the frequency of choosing the global-best solution for the pheromone trial update. In addition in MMAS typically the ants solutions are improved using local optimizers before the pheromone update.

2. Also in MMAS, the value of pheromone is restricted in a closed interval $[\tau_{\min}, \tau_{\max}]$. The chance of algorithm stagnation is thus decrease by giving each connection some, although very small, probability of being chosen. In practice, heuristics exist for setting $\tau_{\min}$ and $\tau_{\max}$. First it can be shown that, because of the pheromone evaporation, the maximal possible pheromone trial level is limited to $\tau^*_{\max} = \frac{1}{\rho.C(S^*)}$, where S^* is the optimal solution.

Based on this result, the global best solution can be used to estimate $\tau_{\max}$ by replacing S^* with $S_{global-best}$t in the equation for $\tau^*_{\max}$. For $\tau_{\min}$ it is often enough to choose it as some constant factor lower than $\tau_{\max}$. As a means for further increasing the exploration of solution, MMAS also uses the occasional re-initialization of pheromone trails.

3. Instead of initializing the pheromones to a small amount, in MMAS the pheromone trails are initialized to an estimate of the maximum allowed pheromone trail value. This leads to an additional diversification component in the algorithm, because at the beginning the relative difference of the pheromone trails will not be very marked, which is different when initializing the pheromone trails to some very small value.

The procedure of MMAS can be found in [11] with more details. Two special properties of this method are as follows:

i) items are chosen randomly,

ii) pheromone trails refers to the desirability of assigning item i to location j in iteration s is as:

$$p^k_{ij}(s) = \frac{\tau_{ij}(s)}{\sum_l \tau_{il}(s)}.$$

For using a method which is based on MMAS, we need to define a criteria for measuring the objective and updating the trials of pheromone. For this purpose, we suppose that after the iteration s we obtain the control $u^{<s>}$ as:

$$u^{<s>}(t) = \sum_{\upsilon=1}^{n} u_{\upsilon}^{<s>} \xi_{[t_{\upsilon-1}, t_{\upsilon}]}(t),$$

where $u_{\upsilon}^{<s>}$ is the selection of the υth value of control in interval $[t_{\upsilon-1}, t_{\upsilon}]$. As it is mentioned in Section 2 we consider an approximation of trajectory $x(t)$ corresponding to $u^{<s>}(t)$ from (2) and initial value, and we call it $x^{<s>}(t)$. We also denote the estimation of the criteria for updating tour (pattern) in $s-$th iteration by J_s and define it as follows:

$$J_s = \int_0^{t_f} f_{\circ}(t, x^{<s>}(t), u^{<s>}(t))dt + M \left\| x^{<s>}(t_f) - x_f \right\| \tag{3}$$

where M is a large and positive real number which we consider it as a penalty value for obtaining desired final value x_f. In fact J_s is the same $C(S)$ which is defined in the above. To be taken note that the definition of criteria estimation can be done in different manners. On basis of the discussion in this section, we present the following procedure for obtaining an approximate solution of the optimal control problem in format of MMAS as following pseudo-code:

0.(Initialize) Control-state parametrization, method parameter settings

1. For $s=1$ to $s :=$ the maximum number of iterations, do

2. For $k=1$ 1 to $m :=$ the number of ants, do

3. Repeat until ant k has completed a tour

4. Determine the best probability p_{ij}^{k}

5. Calculate the objective J_s of the tour generated by ant k

6.Call the best objective between all tours of ants as J^* and update general pheromone.

Numerical Examples

In this section we present some numerical examples to show the implementation and accuracy confirmation of the proposed method.

Example 1. In the first example we consider an OCP of minimizing

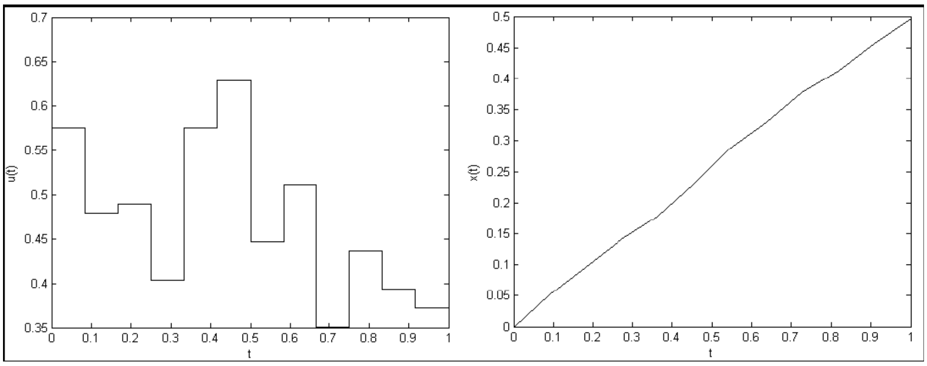

Fig.2. The resulting piece-wise control of Example 1.

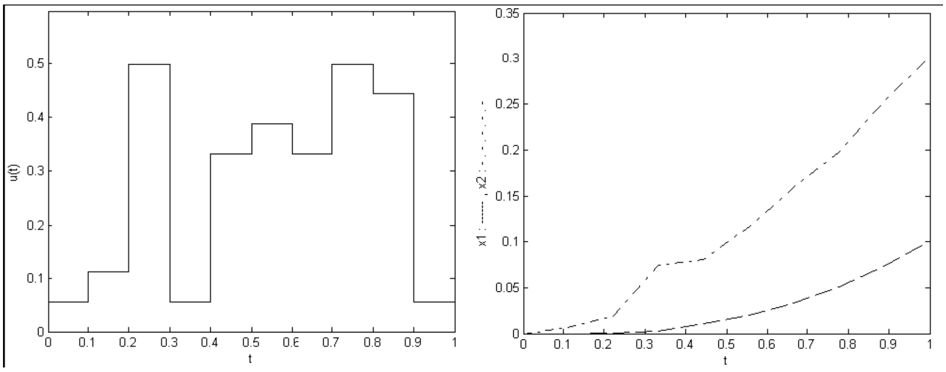

Fig.3. The resulting state of Example 1.

$$I(u(\cdot)) = \int_0^1 u^2(t)dt,$$

subject to

$$\dot{x} = \frac{1}{2}x^2 \sin x + u$$

with $x(0) = 0$, $x(1) = 0.5$ as initial and final conditions. Here the control function values are in $[0.3, 0.7]$. For control-state division we choose $n = 10$ and $m = 10$. By applying the procedure of Section 4 with 100 ants we obtain $I^* = 0.2293$ in 50 iterations. The approximate optimal control in piece-wise linear form is shown in Fig.2. If we substitute this function in the system equations, then an initial valued problem is left to solve by a numerical method.

We solve this initial value problem by using Rung-Kutta method of forth order ([2]) to find the corresponding state function as depicted in Fig.3. The value of trajectory corresponding to the final time t_f is $x^*(1) = 0.4968$ which shows the accuracy of the method in final condition.

Example 2. In the second example we consider a nonlinear OCP involving minimization of

$$I(x_1(\cdot), x_2(\cdot)) = \int_0^1 ({x_1}^2(t) + {x_2}^2(t))dt,$$

where the pair of control-state satisfy in the following non-linear dynamical system:

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = 10x_1^3 + u,$$

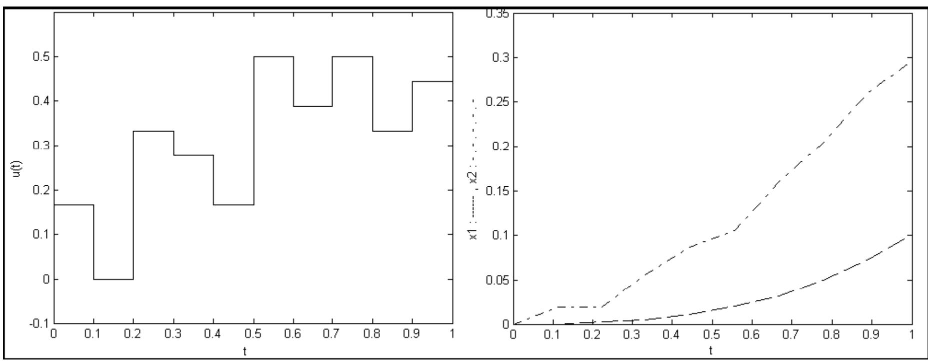

Fig.4. The resulting states of Example 2.

It is desired that the system state moves from $(0,0)$ at $t = 0$ to $(0.1, 0.3)$ at $t = 1$. The control value interval is given by $[0, 0.5]$ which is divided to $m = 10$ portion. The time interval is also divided to

$n = 10$ sub-intervals. By using 200 ants in this example, the method converges to the solution in only 80 iterations. The _nal value of approximate optimal trajectories are obtained with low miss distances as $x_1^*(1) = 0.1009$ and $x_2^*(1) = 0.2969$. The resulting approximate trajectories which have been found by solving the di_erential equation with the resulting control function and initial conditions are depicted in Fig.4.

4. Conclusions

In this paper we tried to apply the benefits of one of the best evolutionary algorithm, ant colony optimization, to obtain approximate solution of optimal control problems. To this means, we proposed an special discretization of control state and then change the procedure of MMAS to obtain the best solution. Numerical results show the accuracy of the method in final conditions. Of course it seems that the number of iterations, ants and the form of discretizing effect on the complexity of method, but the nonlinearity of the objective and system have no serious effect on the procedure. In the case of large discrete optimal control problem, the method may be implemented on parallel computers to save the computational time.

5. References

[1] E. Bonabeau, M. Dorigo, and G. Theraulaz. Inspiration for Optimization from Social Insect Behaviour. *Nature.* 2000, **406**: 39-42.

[2] L. Collatz. *The Numerical Treatment of Differential Equations*. Berlin: Springer, 1960.

[3] O. Cordon, F. Herrera, and T. Stuzle. A review on the Ant Colony Optimization metaheuristic: basis, models and new trends. *Mathware and Soft Computing.* 2002, **9**: 135.

[4] M. Dorigo, Optimization. *Learning and Natural Algorithms.* Ph.D. thesis, Politecnico di Milano, 1992.

[5] M. Dorigo, and L. M. Gambardella. Ant Colonies for the Travelling Salesman Problem. *Biosystems.* 1997, 43: 73-81.

[6] M. Dorigo, V. Maniezzo, and A. Colorni. The Ant System: Optimization by a colony of cooperating agents. *IEEE Transactions on Systems., Man, and Cybernetics - Part B.* 1996, **26**: 113.

[7] O. S. Fard, and A. H. Borzabadi. Optimal control problem, quasi-assignment problem and genetic algorithm. *Enformatika, Transaction on Engin., Compu. and Tech.*. 2007, **19**: 422-424.

[8] H. M. Jaddu. *Numerical Methods for solving optimal control problems using chebyshev polynomials.* PhD thesis, School of Information Science, Japan Advanced Institute of Science and Technology, 1998.

[9] H. H. Mehne, A. H. Borzabadi. A numerical method for solving optimal control problems using state parametrization. *Numerical Algorithms.* 2006, **42**(2): 165-169.

[10] J. E. Rubio. *Control and Optimization the Linear Treatment of Non-linear Problems.* Manchester, U. K., Manchester University Press, 1986.

[11] T. Stuzle, and H. H. Hoos. The MAX - MIN Ant System and local search for the traveling salesman problem. *In T. Back, Z. Michalewicz, and X. Yao, editors, Proceedings of the 1997, IEEE International Conference on Evolutionary Computation (ICEC97).*1997, pp. 309-314.

[12] T. Stuzle, and H. H. Hoos. MAX-MIN Ant System. *Future Generation Computer Systems Journal.* 2000, **16**(8):. 889-914.

[13] T. Stuzle. MAX-MIN Ant System for the quadratic assignment problem. *Technical Report AIDA974.* FG Intellektik, FB Informatik, TU Darmstadt, July 1997.

[14] K. L Teo, C. J. Goh, and K. H. Wong. A unified computational approach to optimal control problems. *Longman Scientific and Technical*, 1991.

[15] K. L Teo, L. S. Jennings, H. W. J. Lee, V. Rehbook. Control parametrization enhancing transform for constrained optimal control problem. *J. Aust. Math. Soc.*. 1999, **B40**: 314-335.