ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 1, No. 3, 2006, *pp*. 161-167

An Improved Annealing Algorithm Based on Multi-Agent

Lichao Wang [1], Jianliang Li [2 +] and Yun Zou [1]

[1] School of Automation, Nanjing University of Science & Technology, Nanjing210094, P. R.China

[2] School of Science, Nanjing University of Science & Technology，Nanjing 210094, P. R. China

(*Received November 26, 2005, Accepted February 18, 2006*)

Abstract. Considering the characteristics of Simulated Annealing Algorithm (SAA), based on the agents' perception and retroaction to their surroundings, in this paper a new algorithm called Multi-agent Annealing Algorithm (MAA) is presented. And the global convergence of MAA is gotten. Finally, we make several numerical experiments to compare MAA with SAA and Genetic Algorithm (GA) by three classical test functions in optimization, which suggest that MAA is superior to GA and SAA.

Keywords: Multi-agent, Annealing Algorithm, Global Convergence, Numerical Analysis

1. Introduction

Many engineering problems can be concluded to a global optimization as follows

$$\min_{\mathbf{x}\in S} f(x) \qquad S \subseteq \mathrm{R}^n \tag{1}$$

where $f(x)$ is the objective function and $S = \{(x_1, x_2, ..., x_n) \mid \underline{x}_i \le x_i \le \overline{x}_i, i = 1, 2, ..., n\}$ is its solution space. Methods for global optimization can be divided into two classes: deterministic and stochastic. The former such as analytic method has a high convergence, but it cannot be used to non-differentiable functions and is sensitive to initial guess points. It is usually trapped in a local solution. Therefore, the latter including Genetic Algorithm (GA) and Simulated Annealing Algorithm (SAA) is widely used in engineering problems recently. GA has characteristics of strong robustness and parallel processing, but its weak capacity of climbing hill limits its application [1-2]. SAA is a new method for global optimizations, which can avoid being trapped in a local solution by carefully allowing the configuration of input variables to temporarily make the output worse, enabling the solution to jump out of a local solution and fall into a more productive path toward the global solution [3-6]. But single-point searching, the main drawback of SAA, brings on a contradiction between accuracy and running time, especially when the objective function is complicated or multidimensional, the running time of SAA is unacceptable. Based on their perception and retroaction to their surroundings in artificial intelligence, agents are utilized in SAA to overcome the algorithm's drawbacks.

2. Simulated Annealing Algorithm and Agents

2.1. Annealing Algorithm and Agents

SAA is a stochastic method based on Monte Carlo iterative algorithm. It derives from a simulation of solid matter's annealing process. This algorithm utilizes a decreasing cooling schedule. During the annealing process, particles in the solid matter get more regular and the energy keeps decreasing. The essence of SAA is to search the energy solution of the solid matter, that is, to repeat searching stochastic in solution space by Metropolis sampling technique as the temperature descends until the optimum solution is obtained.

An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through effectors. For each possible percept sequence, an ideal rational agent should do whatever action is expected to maximize its performance measure on the basis of the evidence provided by the percept sequence and whatever built-in knowledge.[7-9] According to the definition, we can give different descriptions for different problems.

[+] Corresponding author. *Tel.*: +86-25-85408876, *E-mail address*: ljl6006@hotmail.com.

Published by World Academic Press, World Academic Union

For the optimization problem (1), an agent represents a solution point that is a real vector $\left(x_1,...,x_n\right)^T$. Many agents are placed in a gridding whose size is $L_{size} \times L_{size}$. Each agent is fixed in an intersection of the gridding. Each agent can only perceive locally and interact with the agents in its surroundings. Only the conjoint agents can interact.

Definition 1. Let $L_{i,j}\ \left(i, j = 1, 2, ..., L_{size}\right)$ denote the agent at $\left(i, j\right)$ in the intellective gridding, whose neighborhood is denoted by $Nbs_{i,j} = \left\{L_{i_1,j}, L_{i,j_1}, L_{i_2,j}, L_{i,j_2}\right\}$ where

$$i_1 = \begin{cases} i-1 & i \neq 1 \\ L_{size} & i = 1 \end{cases},\ j_1 = \begin{cases} j-1 & j \neq 1 \\ L_{size} & j = 1 \end{cases},\ i_2 = \begin{cases} i+1 & i \neq L_{size} \\ 1 & i = L_{size} \end{cases},\ j_2 = \begin{cases} j+1 & j \neq L_{size} \\ 1 & j = L_{size} \end{cases}$$

Every agent has certain energy which should embody the objective functional value. The purpose of the agents' evolution is to decrease its energy, which leads to competition between the conjoint agents. At last, the agents with higher energy will die off, whose position will be replaced by new agents, because the agents are intelligent and they can learn from their surroundings to generate new agents.

2.2. The agents' abilities of competition and learning

When the temperature is T_m, we suppose that the intellective gridding L equals to the intellective gridding L^{T_m}, $L_{i,j}^{T_m} = \left(l_1^m, l_2^m, ..., l_n^m\right)^T$ denotes the agent at $\left(i, j\right)$ in intellective gridding L, $Max_{i,j}^{T_m} = \left(m_1^m, m_2^m, ..., m_n^m\right)^T$ denotes the agent with the lowest energy in the neighborhood of $L_{i,j}^{T_m}$, $Best^{T_m} = \left(b_1^m, b_2^m, ..., b_n^m\right)^T$ denotes the best agent in the intellective gridding, $FL_{i,j}$ denotes the energy of $L_{i,j}^{T_m}$ and $FM_{i,j}$ denotes the energy of $Max_{i,j}^{T_m}$. Then there is the competitive rule or surviving rule as follows

$$\begin{cases} L_{i,j}^{T_m}\ \ keep\ alive & exp\left(\dfrac{FM_{i,j} - FL_{i.j}}{\beta \times T_m}\right) \geq rand(.)\ and \quad FM_{i,j} \neq FL_{i,j} \\ L_{i,j}^{T_m}\ \ die\ off & otherwise \end{cases} \tag{2}$$

where $rand(.)$ is a random digit in $\left[0,\ \ 1\right]$, β is a parameter related to the size of independent variables. The dead agent will be replaced by a new agent $New_{i,j} = \left(e_1, e_2, ..., e_n\right)^T$ generated by the simulated annealing search, which is the embodiment of agents' learning ability. First, a new real vector is generated, that is $Z = \left(z_1, z_2, \cdots, z_n\right)^T$, where

$$z_i = \frac{W_i}{\sqrt{\sum_{j=1}^{n} W_j^2}}\left[\frac{1}{u_i^3} - 1\right],\ i = 1, 2, \cdots, n$$

where $W_1, W_2, \cdots W_n$ is a set of independent stochastic variables of uniform distribution in $\left[-1,\ \ 1\right]$, $u_1, u_2, \cdots u_n$ is a set of independent stochastic variables of uniform distribution in $\left[0,\ \ 1\right]$,and $\left\{W_i\ \right\}_1^n$ and $\left\{u_i\ \right\}_1^n$ are independent to each other. Suppose the dead agent is $Old_{i,j} = \left(d_1, d_2, ..., d_n\right)^T$, then a new agent $New_{i,j}$ is generated as follows

$$New_{i,j} = Old_{i,j} + 2 \times Rand(.) \times \left(Max_{i,j} - Old_{i,j}\right) + LENGTH \times Z$$

where $Rand(.)$ is a diagonal matrix whose diagonal elements are a set of independent stochastic variables of uniform distribution in $\left[0,\ \ 1\right]$. If the new agent satisfies surviving rule (2), it will be placed at $\left(i, j\right)$ in the intellective gridding. Otherwise, the searching for the new agent continues. A new agent will not be generated at random until the searching for a new agent at $\left(i, j\right)$ has repeated for certain times.

3. Multi-agent Annealing Algorithm

3.1. The agents' abilities of competition and learning

Algorithm

Step 1 Initialize of the intellective gridding L^{T_0} and search for $Best^{T_0}$.

Step 2 When the temperature is T_m, let the agents in L^{T_m} compete in their surroundings and learn from their environment.

Step 3 Search for $Best^{T_m}$ in L^{T_m}.

Step 4 If the stopping criterion is satisfied, $Best^{T_m}$ will be outputted as the approximate optimum solution and the energy corresponding as the approximate optimum value, then the iteration is over; Otherwise, $m \leftarrow m+1$, go to **Step 2**.

In MAA, we assume the renewal function of temperature to be

$$T_m = T_0 \big/ (m+1)^2 .$$

For problem (1), if the precision of the *i*th variable is ε_i, the solution space S can be thought to be an discrete space whose size is

$$|S| = \prod_{i=1}^{n} \left\lceil \frac{(\overline{x}_i - \underline{x}_i)}{\varepsilon_i} \right\rceil$$

Let $E = \{Eng(X) \mid X \in S\}$ where *Eng(X)* denotes the energy of the agent *X*, obviously $|E| \le |S|$, then $E = \{E^1, E^2, ..., E^{|E|}\}$ where $E^1 < E^2 < ... < E^{|E|}$. According to the agents' energy, S can be divided into many subsets $\{S^i\}$, where

$$S^i = \{X \mid X \in S, and, Eng(X) = E^i\},\ i = 1, 2, ..., |E|$$

$$\sum_{i=1}^{|E|} |S^i| = |S|;\ S^i \ne \varnothing,\ \forall i \in \{1, 2, ..., |E|\} \tag{3}$$

$$S^i \cap S^j = \varnothing,\ i \ne j,\ \bigcup_{i=1}^{|E|} S^i = S$$

Obviously, E^1 is the optimum value and S^1 includes all the agents whose energy is E^1.

Let WG denote the set which includes all the intellective griddings, then $|WG| = |S|^{L_{size} \times L_{size}}$. To evaluate the intellective griddings, we define the energy of intellective gridding L to be

$$Eng(L) = \text{Min}\{\ Eng(L_{i,j}) \mid \quad i, j = 1, 2, \cdots, L_{size}\ \}$$

Then, for $\forall L \in WG$, the following holds: $E^1 \le Eng(L) \le E^{|E|}$. So WG can be divided into many subsets $\{WG^i\}$, where

$$WG^i = \{L \mid L \in WG, and, Eng(L) = E^i\},\quad i = 1, 2, ..., |E|;$$

$$\sum_{i=1}^{|E|} |WG^i| = |WG|;\qquad WG^i \ne \varnothing,\quad \forall i \in \{1, 2, ..., |E|\};$$

$$WG^i \cap WG^j = \varnothing,\quad i \ne j;\ \bigcup_{i=1}^{|E|} WG^i = WG.$$

Accordingly, WG^1 includes all the intellective griddings whose energy is E^1

Let $L^{ij}\ (i = 1, 2, ..., |E|,\quad j = 1, 2, ..., |WG^i|)$ denote the *j*th intellective gridding in WG^i. In the algorithm process, when the temperature is T_m, we let $L^{ij} \xrightarrow{T_m} L^{kl}$ denote the state transition form L^{ij} to L^{kl}, $p_{ij,kl}^m$ denote the trasition probability from L^{ij} to L^{kl}, $p_{ij,k}^m$ denote the transition probability from L^{ij} to any intellective gridding in WG^k, and $p_{i,k}^m$ denote the transition probability from any intellective gridding in WG^i to any intellective gridding in WG^k. Then

$$p_{ij,k}^m = \sum_{l=1}^{\left|WG^k\right|} p_{ij,kl}^m , \quad \sum_{k=1}^{|E|} p_{ij,k}^m = 1 , \quad p_{i,k}^m \ge p_{ij,k}^m .$$

3.2. Convergence of MAA

Theorem 1. In MAA, for $\forall i,k \in \{1,2,\dots,|E|\}$, $\forall m \in N$, the following holds

$$p_{i,k}^m = \begin{cases} >0 & k \le i \\ =0 & k > i \end{cases} .$$

Proof: Firstly, for $\forall L^{ij} \in WG^i \left(i = 1,2,\dots,|E|, j = 1,2,\dots,\left|WG^i\right|\right)$, $\exists X^* = \left(x_1, x_2, \dots, x_n\right)^T \in L^{ij}$, and $Eng\left(X^*\right) = E^i$ .According to the rules by which the new agents generate, the iteration of MAA only brings greater intellective griddings. Hence for $\forall i,k \in \{1,2,\dots,|E|\}$, $\forall m \in N$, if $i < k$, $p_{i,k}^m = 0$.

Secondly, when the temperature is T_m, we assume $X^1 = \left(x_1^1, x_2^1, \dots, x_n^1\right)^T$ be the best agent in the neighborhood of dead agent $X^0 = \left(x_1^0, x_2^0, \dots, x_n^0\right)^T$, and $X^2 = \left(x_1^2, x_2^2, \dots, x_n^2\right)^T$ be the new agent who replaces the $X^0 = \left(x_1^0, x_2^0, \dots, x_n^0\right)^T$ in the intellective gridding L^{ij}. By formula (3), we can conclude that $\exists X' = \left(x_1', x_2', \dots, x_n'\right)^T \in S^k \left(k \le i\right)$, as long as $\sum_{X' \in S^k} P\left(X^2 = X'\right) > 0$, we can get that $p_{i,k}^m > 0$.

$$\left(X^2 = X'\right) \supseteq \bigcap_{i=1}^{n} \left(x_i' = x_i^0 + 2 \times rand\left(.\right) \times \left(x_i^1 - x_i^0\right) + LENGTH \times z_i\right)$$

$$= \bigcap_{i=1}^{n} \left(x_i' = x_i^0 + 2 \times rand\left(.\right) \times \left(x_i^1 - x_i^0\right) + LENGTH \times \frac{W_i}{\sqrt{\sum_{j=1}^{n} W_j^2}} \left[\frac{1}{u_i^3} - 1\right] \right)$$

$$\supseteq \bigcap_{i=1}^{n} \left(\begin{array}{c} M_i = x_i' - x_i^0 - 2 \times rand\left(.\right) \times \left(x_i^1 - x_i^0\right) \\ W_i = k_i \times M_i \\ u_i = \sqrt[3]{\dfrac{k_i \times LENGTH}{k_i \times LENGTH + \sqrt{\sum_{i=1}^{n} \left(k_i \times M_i\right)^2}}} \end{array} \right) \triangleq A$$

where k_i is used to ensure that W_i is an stochastic variable in $[0,1]$.

Because the solution space S is a discrete space, $P\left(A\right) = \upsilon > 0$ .Then $p_{i,k}^m > 0$ if $i \ge k$.

From the conclusions above, we can prove the global convergence of MAA as follows. To describe easily, we let T_m denote the temperature when the intellective gridding has altered for m times, P^m denote a stochastic matrix. Each $WG^i \left(i = 1,2,\dots,|E|\right)$ can be seen as a state in a finite Markov chain. According to Theorem 1, P^m, the transition matrix of that Markov chain, can be described as follows

$$P^m = \begin{pmatrix} p_{1,1}^m & 0 & \dots & 0 \\ p_{2,1}^m & p_{2,2}^m & \dots & 0 \\ \dots & \dots & \dots & \dots \\ p_{|E|,1}^m & p_{|E|,2}^m & \dots & p_{|E|,|E|}^m \end{pmatrix} = \begin{pmatrix} C^m & 0 \\ R^m & T^m \end{pmatrix} \quad (4)$$

where $C^m = \left(p_{1.1}^m\right) = (1)$, $R^m = \left(p_{2,1}^m, p_{3,1}^m, \dots, p_{|E|,1}^m\right)^T > 0$,

$$T^m = \begin{pmatrix} p_{2,2}^m & 0 & \dots & 0 \\ p_{3,2}^m & p_{3,3}^m & \dots & 0 \\ \dots & \dots & \dots & \dots \\ p_{|E|,1}^m & p_{|E|,2}^m & \dots & p_{|E|,|E|}^m \end{pmatrix} \geq 0$$

Theorem 2. In the condition of theorem 1, MAA is global convergent.

Proof: By the formula (4), we can get that

$$\prod_{m=0}^{\infty} P^m = \lim_{M\to\infty} \prod_{m=0}^{M} P^m = \lim_{M\to\infty} \begin{pmatrix} 1 & 0 \\ R^0 + \sum_{m=1}^{M} \left(\prod_{i=1}^{m-1} T^i \right) R^m & \prod_{m=0}^{M} T^m \end{pmatrix}$$

For $\forall i \in N$, $T^i\delta = \delta - R^i$ where $\delta = (1,1,\dots,1)^T$, then $R^i = \left(E - T^i\right)\delta$ where E is an identity matrix. So

$$R^0 + \sum_{m=1}^{M} \left(\prod_{i=1}^{m-1} T^i \right) R^m = R^0 + \sum_{m=1}^{M} \left(\prod_{i=1}^{m-1} T^i \right) R^m = R^0 + \sum_{m=1}^{M} \left(\prod_{i=1}^{m-1} T^i \right) \left(E - T^m\right)\delta$$

$$= R^0 + \sum_{m=1}^{M} \left(\prod_{i=1}^{m-1} T^i - \prod_{i=1}^{m} T^i \right)\delta = R^0 + T^0\delta - \prod_{i=1}^{M} T^i\delta$$

$$= \delta - \prod_{i=1}^{M} T^i\delta$$

According to the characteristics of stochastic matrix, we can get that $\left\|T^i\right\|_\infty < 1 - \upsilon$, then

$$\begin{matrix} \prod_{m=0}^{M} T^m \to 0 \\ R^0 + \sum_{m=1}^{M} \left(\prod_{i=1}^{m-1} T^i \right) R^m \to \delta \end{matrix}, \quad M \to \infty$$

therefore

$$\prod_{m=0}^{\infty} P^m = \begin{pmatrix} 1 & 0 \\ \delta & 0 \end{pmatrix}$$

Accordingly,

$$\lim_{m\to\infty} P\left\{ Eng\left(L^{T^m}\right) = E^1 \right\} = 1$$

That is, SAA converges to the global solution.

4. Numerical experiment

To verify the efficiency of MAA, we make several numerical experiments to compare MAA with SAA and GA by three classical test functions in optimization. The test functions are listed as follows:

1) Generalized Griewank function

$$f_1(X) = \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \qquad x_i \in [-600, 600]$$

2) Ackley function

$$f_2(X) = -20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right) - \exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\right) + 20 + e \quad x_i \in [-30,30]$$

3) Rastrigin function

$$f_3(X) = \sum_{i=1}^{n}\left[x_i^2 - 10\cos(2\pi x_i) + 10\right] \quad x_i \in [-5.12,5.12]$$

The three functions have the same solution (zero vector) and the same optimum value (zero).

Table 1. The results for functions of 5 variables

Function	Algorithm	Approximate solution	App. value	CPU time
Generalized Griewank	SA	$(0.0050032,-0.0370384,0.018544, 0.049868,0.0581997)^T$	0.969304	1046
	GA	$(0.180314,-0.0860113,-0.0261525, 0.220705,0.441092)^T$	26.5153	1250
	MAA	$(0.000229365,0.0082141,-7.70802e-005, -0.000359218,0.0113358)^T$	0.0952871	343
Ackley	SA	$(0.038154,0.00120962,0.0145285, 0.997239,-0.0209483)^T$	1.73091	1187
	GA	$(0.180314,-0.0860113,-0.0261525, 0.220705,0.127479)^T$	1.41461	1265
	MAA	$(0.000372097,-0.00255419,0.00151736, 0.00342458,0.00182566)^T$	0.0090241	468
Rastrigin	SA	$(0.00560495,0.00529649,7.825e-005, 0.987643,0.0118406)^T$	1.04515	953
	GA	$(-0.0127323,-0.13267,0.963046, -0.0230955,-0.131381)^T$	7.86428	1109
	MAA	$(0.000400278,-0.00145453,-0.00284663, 0.00281354,-0.000764573)^T$	0.00374522	234

In the experiments, the operating condition is CPU 1.80GHZ, and VC++. To use the same program, the intervals of definition of the three functions are all turned into $[-1,1]$ by linear transformation. For SAA[10], we let the step size *Length*=1E-7, initial temperature T_0=2000000 and the stopping criterion be $t \le$ 1E-5. For GA[11], we let populations size *M*=61, evolutional generation size k=400, crossover rate *pcross*=0.9, and mutation rate *pmutation*=0.02. For MAA, we let the number of times we search for a new agent each time *c*0=4, L_{SIZE}=7, the step size *Length*=1E-5, and the stopping criterion be $t \le$ 1E-5. In the experiments, the unit of time is 10^{-3}s. The results in tables are the average of independent experiments for10 times.

Table 2. The results for functions of 20/100 variables

Function	Algorithm	Approximate optimum value	CPU time
Generalized Griewank	SA	2.90129/895.59	2578/6609
	GA	253.715/2069.02	1953/7640
	MAA	1.23907/39.6328	1140/5953
Ackley	SA	0.724477/2.97288	2812/6531
	GA	3.08849/3.48865	2046/6640
	MAA	0.379777/0.781705	1093/3093
Rastrigin	SA	11.0912/448.108	3203/10781
	GA	142.185/879.943	1578/6031
	MAA	6.691/97.7543	1296/3875

From table 1, we can conclude that MAA is superior to GA and SAA on approximate optimum value and running time. The table 2 suggests that the single-point searching in SAA leads to the poor rapidity of convergence, and that GA is trapped in a local solution early and hard to get out, which is the embodiment of its weak capacity of climbing hill.

5. Conclusion

A multi-agent Annealing Algorithm for global optimizations is proposed in this paper, and whose global convergence is testified. The numerical results reveal that MAA is superior to GA and SAA.

6. References

[1] J. H. Holland, Genetic algorithms. *Scientific American*, 1992, **7**: 66-72.

[2] M. Srinivas, L. M. Pataik, genetic algorithms. *A survey, IEEE*, 1994, **27**(6): 17-26.

[3] H. H. Szu, Non-convex optimization, *Proc. SPIE*, 1986, **698**: 59-65.

[4] L. Ingber, Simulated annealing: practice versus theory. *Math. and Comp. Modeling*, 1993, **18**: 29-57.

[5] R. Yang, J. Gu. An efficient simulated annealing algorithm for global optimization. *Systems Engineering-Theory & Practice*, 1997, 5: 29-35.

[6] S. Kirkpatrick, et al. Optimization by simulated annealing. *Science*, 1983, **220**: 671-680.

[7] S. Russell, et al. Artivicial intelligence: a modern approach. *New York*, Prentice-Hall,1995.

[8] J. Han, Q. Cai, Emergent intelligence in AER model. *Pattern Recognition and Artificial Intelligence*, 2002, **15**(2): 134-142.

[9] W. Zhong, et al. Higher-dimensions iptimization using multi-agent evolutionary algorithm. *Chinese Journal of Cmputers*, 2004, **27**(10): 1341-1353.

[10] C. Wang, G. Li, X. Zhao. VaR-GARCH model using simulated annealing algorithm. *J. of Sys. Eng*., 2003, **18**(1): 1-7.

[11] Genetic Algorithms and Their Application[C]. *Proceedings of International Conference on Genetic Algorithm*, l993.