INFORMATICA, 1996, Vol. 7, No. 2, 167–174

OPTIMIZATION OF A "MIXTURE" OF PURE HEURISTICS AND MONTE CARLO ALGORITHMS

Jonas MOCKUS

Vytautas Magnus University
Vileikos St. 8, 2042 Kaunas, Lithuania

Elvyra SENKIENĖ

Institute of Mathematics and Informatics
Akademijos St. 4, 2600 Vilnius, Lithuania

Abstract. We consider a stochastic algorithm of optimization in the presented paper. We deal here with the average results of a "mixture" of the deterministics heuristics algorithm and uniform random search. We define the optimal "mixture".

Key words: average results, optimization, heuristics, mixture, uniform random search.

1. Introduction. Denote by $(h, h, \ldots, h)$ the case of applying pure heuristics k times. Denote average results as $E_k(h, h, \ldots, h)$. Assume, for simplicity that

y_k are values of the objective function uniformly distributed in a unit interval [0,1],

$v(k, 0)$ are the results of heuristics expressed as

$$v(k,0) = E_k(h,h,\ldots,h) = E_k y^0 \left(\delta + \frac{1-\delta}{(1+k)^2}\right) = \delta + \frac{1-\delta}{(k+1)^2}. \quad (1)$$

We assume here that the starting point $y^0 = 1$. The index k is the number of stages and the parameter $\delta \geqslant 0$ denotes the asymptotic residual value as $k \to \infty$. Thus expression (1) roughly represents the results of some local minimum δ.

After simple calculations we get the expression of the average results $E_k(r, r, \ldots, r)$ of the k-th repetition of Monte Carlo search

$$v(0,k) = E_k(r,r,\ldots,r) = \frac{1}{k+1}. \quad (2)$$

Here we prove equality (2).

$$F(y) = \mathbf{P}\{\min_{1<i\leqslant k} y_i < y\} = 1 - (1-y)^k,$$

(see Wilks, 1967; Gnedenko, 1965; Senkienė, 1987). Then

$$\frac{dF}{dy} = k(1-y)^{k-1},$$

and

$$v(0,k) = E_k(r,r,\dots,r) = k\int_o^1 y(1-y)^{k-1}dy = \frac{1}{k+1}.$$

If we first repeat pure heuristics for l times and afterwards Monte Carlo for $m = k - l$ times, then

$$v(l,m) = E_k(h,\dots,h,r,\dots,r) = \frac{1}{m+1}\Big[1 - \big(1 - v(l,0)\big)^{m+1}\Big]. \quad (3)$$

Now we prove equality (3).

Denote by $E(y) = v(l,m) = E_k(h,\dots,h,r,\dots,r)$, $\varepsilon = v(l,0)$, and $\xi = \min_{1\leqslant i\leqslant m} y_i$. Then

$$F(y) = \mathbf{P}\{\min(\varepsilon,\xi) < y\} = \begin{cases} 1-(1-y)^m, & y \leqslant \varepsilon, \\ 1, & y > \varepsilon. \end{cases}$$

$$\frac{dF}{dy} = \begin{cases} m(1-y)^{m-1}, & y \leqslant \varepsilon, \\ \exists, & y > \varepsilon, \end{cases}$$

$$E(y) = m\int_0^\varepsilon y(1-y)^{m-1}dy + \varepsilon\big[F(\varepsilon+\triangle) - F(\varepsilon)\big] = A + B,$$

$$A = m\int_0^\varepsilon y(1-y)^{m-1}dy = m\int_{1-\varepsilon}^1 (1-v)v^{m-1}dv$$

$$= 1-(1-\varepsilon)^m - \frac{m}{m+1}\big(1-(1-\varepsilon)^{m+1}\big),$$

$$B = \varepsilon\Big(1-\big(1-(1-\varepsilon)^m\big)\Big) = \varepsilon(1-\varepsilon)^m,$$

$$E(y) = A + B = \frac{1}{m+1}\left(1-(1-\varepsilon)^{m+1}\right),$$

$$v(l,m) = E_k(h,\dots,r,\dots,r) = \frac{1}{m+1}\left(1-(1-\varepsilon)^{m+1}\right).$$

If $m = 1$, then

$$v(l,1) = v(l,0) - \frac{v(l,0)^2}{2}.$$

If we start from repeating Monte Carlo for l times and afterwards heuristics for $m = k - l$ times, then

$$v(m,l) = E_k(r,\dots,r,h,\dots,h) = \frac{1}{l+1}\left(\delta + \frac{1-\delta}{(m+1)^2}\right). \tag{4}$$

Now we are ready to consider a "mixture" of pure heuristics and Monte Carlo. Denote by q_i a probability of choosing Monte Carlo at the stage i, then the probility of pure heuristics at this stage will be $p_i = 1 - q_i$. Assume that probalilities q_i are independent.

Denote by $v_q(m,l)$ the averages of $v(m,l)$ given $q = (q_1,\dots,q_k)$. For example, using Monte Carlo at all stages

$$v_q(0,k) = \prod_{i=1}^{k} q_i v(0,k), \tag{5}$$

using the pure heuristics at all stages

$$v_q(k,0) = \prod_{i=1}^{k}(1-q_i)v(k,0), \tag{6}$$

repeating the pure heuristics for l times and afterwards Monte Carlo for $m = k - l$ times

$$v_q(l,m) = \prod_{i=1}^{l}(1-q_i)\prod_{i=1}^{m} q_i v(l,m), \tag{7}$$

repeating Monte Carlo for l times and afterwards heuristics for $m = k - l$ times

$$v_q(m,l) = \prod_{i=1}^{l} q_i \prod_{i=1}^{m}(1-q_i)v(m,l). \tag{8}$$

2. Two-stage case. A pair (r,h), for example, denotes that we use Monte Carlo at the first stage and heuristics at the second one. Assume, for simplicity, that $q_1 = q_2 = q$. Then from (5) – (8) the average of $v_q(1,1)$

$$\begin{aligned} v(1,1) &= Ev_q(1,1) \\ &= q^2 v(r,r) + q(1-q)\,[v(r,h) + v(h,r)] + (1-q)^2 v(h,h) \\ &= [v(r,r) + v(h,h) - (v(r,h) + v(h,r))]\,q^2 \\ &\quad + [2v(h,h) - (v(r,h) + v(h,r))]\,q + v(h,h), \qquad (9) \\ \frac{\partial v(1,1)}{\partial q} &= 2\,[v(r,r) + v(h,h) - (v(r,h) + v(h,r))]\,q \\ &\quad + [2v(h,h) - (v(r,h) + v(h,r))]\,, \\ \frac{\partial^2 v(1,1)}{\partial q^2} &= 2\,[v(r,r) + v(h,h) - (v(r,h) + v(h,r))]\,. \end{aligned}$$

The optimal value $q = q^*$

$$q^* = \frac{2v(h,h) - (v(r,h) + v(h,r))}{2\,[v(r,r) + v(h,h) - (v(r,h) + v(h,r))]}.$$

From $0 \leqslant q^* \leqslant 1$, we get

1) $2v(h,h) - (v(r,h) + v(h,r)) \geqslant 0$,

2) $2v(h,h) - (v(r,h) + v(h,r))$
$\quad < 2v(r,r) + 2v(h,h) - 2\,(v(r,h) + v(h,r))$,
$2v(r,r) - (v(r,h) + v(h,r)) \geqslant 0$,

3) $v(r,r) + v(h,h) - (v(r,h) + v(h,r)) > 0$.

Then the optimal value $q = q^*$ is defined by the expression

$$q^* = \begin{cases} q^0 & \text{if } \; 2v(h,h) - (v(h,r) + v(r,h)) \geqslant 0, \\ & \text{and } \; 2v(r,r) - (v(h,r) + v(r,h)) \geqslant 0, \\ & \text{and } \; v(r,r) + v(h,h) - (v(h,r) + v(r,h)) > 0; \\ 0 & \text{if } \; 2v(h,h) - (v(h,r) + v(r,h)) < 0; \\ 1 & \text{if } \; 2v(r,r) - (v(h,r) + v(r,h)) < 0. \end{cases} \qquad (10)$$

From (5) – (8)

$$v(r,r) = \frac{1}{3},$$
$$v(r,h) = \frac{1}{2}\left(\delta + \frac{1-\delta}{4}\right),$$

$$v(h,r) = \left(\delta + \frac{1-\delta}{4}\right) - \frac{1}{2}\left(\delta + \frac{1-\delta}{4}\right)^2,$$
$$v(h,h) = \left(\delta + \frac{1-\delta}{9}\right),$$
$$v(1,1) = q^2 \cdot \frac{1}{3} + q(1-q)\left[\frac{3}{2}\left(\delta + \frac{1-\delta}{4}\right) - \frac{1}{2}\left(\delta + \frac{1-\delta}{4}\right)^2\right]$$
$$+ (1-q^2)\left(\delta + \frac{1-\delta}{9}\right),$$
$$2v(h,h) - (v(r,h) + v(h,r)) = \frac{1}{288}\left(81\delta^2 + 242\delta - 35\right),$$
$$v(r,r) + v(h,h) - (v(r,h) + v(h,r)) = \frac{1}{288}\left(81\delta^2 - 14\delta + 29\right),$$
$$2v(r,r) - (v(r,h) + v(h,r)) = \frac{1}{288}\left(81\delta^2 - 270\delta + 93\right),$$

$$q^0 = \frac{1}{2}\frac{81\delta^2 + 242\delta - 35}{81\delta^2 - 14\delta + 29}; \tag{11}$$

$$2v(h,h) - (v(h,r) + v(r,h)) < 0, \quad \text{if} \quad 81\delta^2 + 242\delta - 35 < 0, \tag{12}$$
$$2v(r,r) - (v(h,r) + v(r,h)) < 0, \quad \text{if} \quad 81\delta^2 - 270\delta + 93 > 0. \tag{13}$$

When solving (11) – (13) we get

$$q^* = \begin{cases} q^0 & \text{if} \quad 0.138 \leqslant \delta \leqslant 0.390, \\ 0 & \text{if} \quad \delta < 0.138, \\ 1 & \text{if} \quad \delta > 0.390. \end{cases} \tag{14}$$

3. Three-stage case. A set (r,r,h), for example, denotes that we use Monte Carlo at the first two stages and heuristics at the third one. $q_1 = q_2 = q_3 = q$. Then from (5) – (8)

$$v = q^3 v(r,r,r) + q(1-q)^2\Big[v(r,h,h) + v(h,r,h) + v(h,h,r)\Big]$$
$$+ q^2(1-q)\left[v(r,r,h) + v(h,r,r) + v(r,h,r)\right] + (1-q)^3 v(h,h,h)$$
$$= \Big[v(r,r,r) + v(r,h,h) + +v(h,r,h) + v(h,h,r)$$
$$- v(r,r,h) - v(h,r,r) - v(r,h,r) - v(h,h,h)\Big]q^3$$

$$
\begin{aligned}
&+\Big[-2v(r,h,h)-2v(h,r,h)-2v(h,h,r)\\
&+v(r,r,h)+v(h,r,r)+v(r,h,r)+3v(h,h,h)\Big]q^2\\
&+\Big[v(r,h,h)+v(h,r,h)+v(h,h,r)-3v(h,h,h)\Big]q+v(h,h,h);
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial v}{\partial q} &= 3\Big[v(r,r,r)+v(r,h,h)+v(h,r,h)+v(h,h,r)-v(r,r,h)\\
&-v(h,r,r)-v(r,h,r)-v(h,h,h)\Big]q^2\\
&-2\Big[-2v(r,h,h)-2v(h,r,h)-2v(h,h,r)\\
&+v(r,r,h)+v(h,r,r)+v(r,h,r)+3v(h,h,h)\Big]q\\
&+[v(r,h,h)+v(h,r,h)+v(h,h,r)-3v(h,h,h)]\\
&= Bq^2+2Aq+C=0. \qquad (15)
\end{aligned}
$$

Here

$$
\begin{aligned}
v(r,r,r) &= \frac{1}{4},\\
v(r,h,h) &= \frac{1}{2}\left(\delta+\frac{1-\delta}{9}\right),\\
v(h,r,h) &= \left(\delta+\frac{1-\delta}{9}\right)-\frac{1}{2}\left(\delta+\frac{1-\delta}{9}\right)^2,\\
v(h,h,r) &= \left(\delta+\frac{1-\delta}{9}\right)-\frac{1}{2}\left(\delta+\frac{1-\delta}{9}\right)^2,\\
v(r,r,h) &= \frac{1}{3}\left(\delta+\frac{1-\delta}{4}\right),\\
v(h,r,h) &= \left(\delta+\frac{1-\delta}{4}\right)-\left(\delta+\frac{1-\delta}{4}\right)^2+\frac{1}{3}\left(\delta+\frac{1-\delta}{4}\right)^3,\\
v(r,h,r) &= \frac{1}{2}\left(\delta+\frac{1-\delta}{4}\right)-\frac{1}{4}\left(\delta+\frac{1-\delta}{4}\right)^2,\\
v(h,h,h) &= \left(\delta+\frac{1-\delta}{16}\right),\\
A &= 0.14\delta^3-3.91\delta^2-1.09\delta-0.22,\\
B &= -0.42\delta^3+6.15\delta^2+1.05\delta+0.88,\\
C &= 1.58\delta^2-0.19\delta+0.11.
\end{aligned}
$$

Solving (15), we get

$$q^0 = \frac{-A \pm \sqrt{A^2 - BC}}{B}. \tag{16}$$

The optimal value of $q = q^*$ is

$$q^* = \begin{cases} q^0 & \text{if} \quad 0 \leqslant \delta \leqslant 0.4, \\ 0 & \text{if} \quad \delta < 0, \\ 1 & \text{if} \quad \delta > 0.4. \end{cases} \tag{17}$$

4. Conclusion. The examples show that using a "mixture" of heuristics and Monte Carlo one may get better results as compared with using them separately. It seems natural, since using the "mixture" we optimize the real variable $q \in [0, 1]$ instead of the Boolean one corresponding to the choice between the pure heuristics $q = 0$ and Monte Carlo $q = 1$.

REFERENCES

Wilks, S. (1967). *Mathematical Statistics.* Nauka, Moscow (in Russian).

Gnedenko, B.V. (1965). *Course of Theory Probability.* Nauka, Moscow (in Russian).

Senkienė, E. (1987). Some properties of the distribution of the minimal term in the sequence of independent normal random variables. *Works of the Lithuanian Academy of Sciences,* Ser.B, **4**(161), 117–122 (in Russian).

Received February 1996

J. Mockus graduated Kaunas Technological University, Lithuania, in 1952. He got his Doctor habilitus degree in the Institute of Computers and Automation, Latvia, in 1967. He is a head of Optimal Decision Theory Department, Institute of Mathematics and Informatics, Vilnius, Lithuania and professor of Kaunas Technological University.

E. Senkienė graduated Vilnius University, Lithuania, in 1966. She received Ph. D. degree of Mathematical Sciences from Vilnius University, Vilnius, Lithuania, in 1974. She is a senior researcher at the Department of Optimal Decision Theory, Institute of Mathematics and Informatics, Vilnius, Lithuania. Her research interests include stochastic processes and optimization problems.

DETERMINUOTO HEURISTINIO IR MONTE CARLO ALGORITMŲ MIŠINIO OPTIMIZAVIMAS

Jonas MOCKUS ir Elvyra SENKIENĖ

Straipsnyje nagrinėjamas stochastinis optimizacijos algoritmas. Čia nagrinėjami determinuoto heuristinio ir tolydinės atsitiktinės paieškos algoritmų mišinio optimizavimo rezultatai. Apibrėžiamas optimalus mišinys.