Multi Armed Bandits for Many Task Optimization

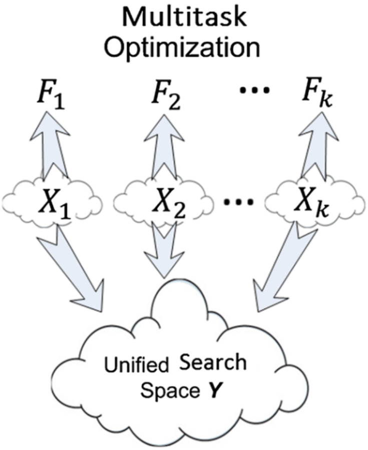
Le Tien Thanh - Ta Bao Thang
Supervisor: Professor Huynh Thi Thanh Binh

January 9, 2021

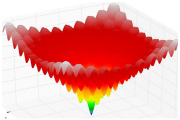
Agenda

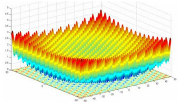
Agenda

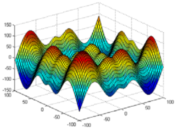
Many task optimization - MaTSOO benchmark



Single Objective Optimization (SOO)







Agenda

MFEA

MFEA

A. Gupta, Y.-S. Ong, and L. Feng, "Multifactorial evolution: Toward evolutionary multitasking," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 3, pp. 343–357, 2015

Problem

- Do not group ideal assist tasks

MaTGA

MaTGA

Y. Chen, J. Zhong, L. Feng, *et al.*, "An adaptive archive-based evolutionary framework for many-task optimization," *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2019

Problem

- Store past solutions
- Measure KLD between past solutions
- Complex, many parameters

Agenda

Task selection as a Multi Armed Bandit (MAB) problem

Action

- Given a K task MTO problem
- Given a parent p_1, skill factor τ_{source}
- **Action** is $\tau_{target} \in \{1, \ldots, K | \tau_{source} \neq \tau_{source}\}$
- Then, select p_2 from $P_{\tau_{target}}$ to reproduce.

Reward

- Given $y_{\tau_{source}}$, list of fitness of $P_{\tau_{source}}$
- **Reward**

$$r = \begin{cases} 1 \text{ if } y_c < min(y_{\tau_{source}}) \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

UCB function to solve MAB

Estimate expected reward

Each source task τ_{source} estimate the expected reward of selecting a target task

$$Q[\tau_{target}] = Q[\tau_{target}] + \alpha(r - Q[\tau_{target}]) \tag{2}$$

where $Q \in \mathbb{R}^{K-1}$ is the estimated expected reward, r is the given reward, α is the discount factor.

UCB function

Each source task τ_{source} select target task by

$$\tau_{target} = \underset{k}{\operatorname{argmin}}\ Q[k] + c\sqrt{\frac{log(t)}{N[k]}} \tag{3}$$

where c is exploration-exploitation trade off coefficient, t is the total number of selection, $N[k]$ is number of times task k is selected.

Linear decay of rmp

High *rmp* first, low *rmp* after

$$rmp = \frac{rmp_{max} - rmp_{min}}{T} \times (T - t) \tag{4}$$

Basic structure of MAB-MFEA

Parameters

- α - discounted factor, c - exploration exploitation trade-off
- rmp_{max}, rmp_{min}

MAB-MFEA evolution step

Algorithm 1 Multi-step SR

Get current generation's rmp
for $i \in \{1, 2, \dots\}$ **do**
  Select parents p_1, p_2 randomly.
  if $\tau_1 = \tau_2$ **then**
    SBX, Polynomial p_1, p_2 to c_1, c_2, evaluate them
  else if $rand() < rmp$ **then**
    Select τ_{target} for τ_1 (Equation.3)
    Select p_2 from τ_{target}
    SBX, Polynomial p_1, p_2 to c_1, c_2, evaluate them
    Update estimate (Equation.2)
  else
    Select p_2 from τ_1
    SBX, Polynomial p_1, p_2 to c_1, c_2, evaluate them
  end if
end for
Elitist selection

Agenda

Benchmark

Task	Function	Ideal Assisted Task
T_1	Sphere1	None
T_2	Sphere2	None
T_3	Sphere3	None
T_4	Weierstrass25D	None
T_5	Rosenbrock	T_1
T_6	Ackley	T_2
T_7	Weierstrass50D	T_3, T_4
T_8	Schwefel	None
T_9	Griewank	T_4
T_10	Rastrigin	None

Parameter setting

MFEA

- Population size: 30
- rmp: 0.3
- sbxdi: 2
- pmdi: 5

MAB-MFEA

- rmp_{min}: 0.1
- rmp_{max}: 0.9
- α: 0.01 - discounted coefficient
- c: 0.5 - UCB explore/exploit trade-off coefficient

MaTGA

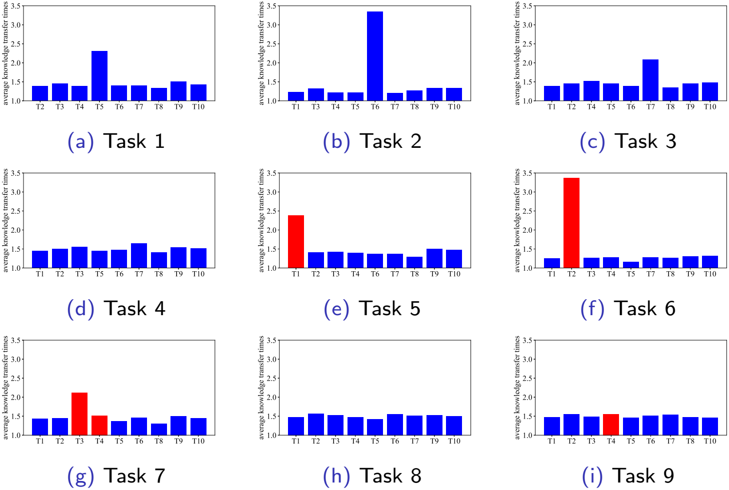
- NP:
- λ: 0.8 - reward shrink rate
- ρ: 0.8 - attenuation coefficient
- α: 0.1 - knowledge transfer rate
- UR: 0.2 - archive update rate
- AcS: 300 - archive size

Table

	MFEA	**MaTGA**	**MAB-MFEA**
T_1	1.32E+00	**2.45E-04**	2.85E-03
T_2	1.27E+00	4.75E-04	**1.15E-12**
T_3	1.17E+00	**0**	**6.30E-09**
T_4	3.37E+00	**1.00E-06**	5.37E-04
T_5	8.15E+02	**2.16E-02**	1.01E+01
T_6	1.99E+01	3.61E-03	**1.05E-07**
T_7	1.05E+01	**5.57E-04**	1.52E-03
T_8	2.34E+03	**3.40E-02**	9.21E+02
T_9	4.11E-02	4.52E-03	**4.50E-03**
$T_1$0	1.55E+01	1.57E+01	**7.77E+00**

Table 1: The performance of 3 algorithms for 30 independent runs

Analysis



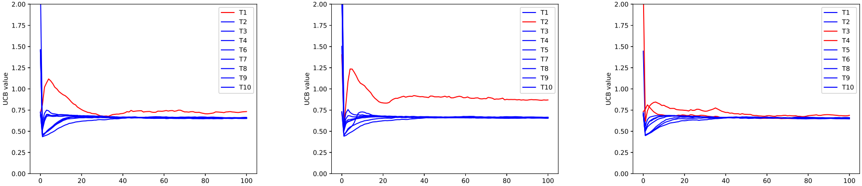
(a) Task 1 (b) Task 2 (c) Task 3

(d) Task 4 (e) Task 5 (f) Task 6

(g) Task 7 (h) Task 8 (i) Task 9

Figure 1: The average times of choosing knowledge transfer target, the assisted task is highlighted in red

Analysis



(a) Task 5 (b) Task 6 (c) Task 7

Figure 2: The UCB function value 3 for choosing knowledge transfer target, the assisted task is highlighted in red

Conclusion

Contribution

- Trade of between explore and exploit task relationship.
- Better than MFEA, less complex but performs equally well to MaTGA.
- The best on 5/10 tasks.
- Analyse on 10 task benchmark shows MAB-MFEA correctly identify ideal assist task.

Future task to complete the study

- Adjust *rmp_min* and *rmp_max* paramter.
- Finish experiment on CEC Complex 50 tasks benchmark.

Thank you for your attention!