Multi-agent Dynamic Algorithm Configuration

Ke Xue[1*], Jiacheng Xu[1*], Lei Yuan[1], Miqing Li[2],
Chao Qian[1†], Zongzhang Zhang[1], Yang Yu[1]

NeurIPS'22

Email: {xuek, qianc}@lamda.nju.edu.cn
Nanjing University[1]
University of Birmingham[2]

Background

Deep learning

engineering features → learning features

SIFT (Lowe et. al. 1999)
HOG (Dalal et. al. 2005)

LeNet (LeCun et. al. 1998)
AlexNet (Krizhevsky et. al. 2012)

Meta learning

engineering to learn → learning to learn

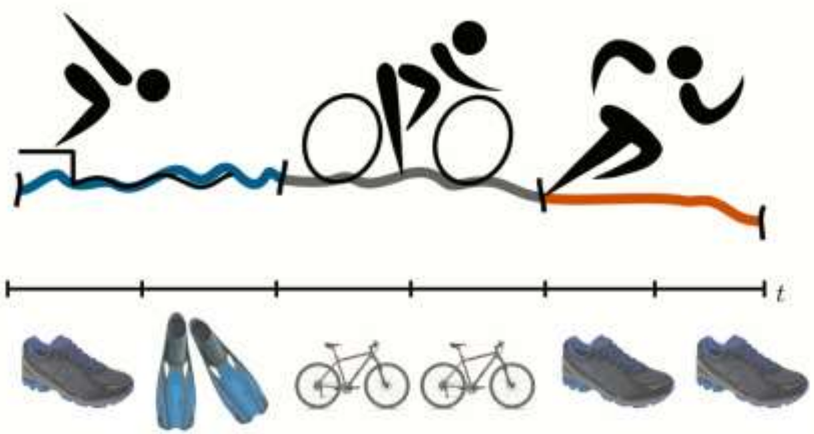
SGD (Robbins et. al. 1951, Bottou 2010)
Autoencoders (Hinton et. al. 2006)

Learning To Learn (Hochreiter et. al. 2001)
Learned Optimizers (Andrychowicz et. al. 2016, Li et. al. 2016, Wichrowska et. al. 2017, Metz et. al. 2018, 2019)

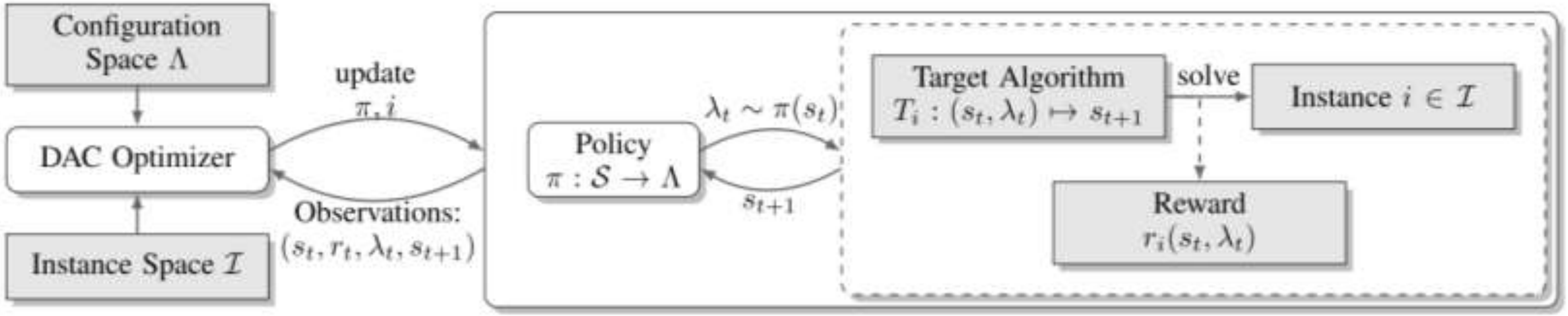
Background

Dynamic algorithm configuration (DAC) is a new trend in Auto-ML.

Algorithm Configuration (AC)



Dynamic Algorithm Configuration (DAC)



$$\pi^* \in \arg\min_{\pi \in \Pi} \int_{i \in \mathcal{I}} p(i) c(\pi, i) \mathrm{d}i$$

[Eimer et al., IJCAI'21]

https://www.automl.org/automated-algorithm-design/dac/

Background

DAC has been found to outperform static methods on many tasks

- learning rate tuning of deep neural networks
- step-size adaptation of evolution strategies
- heuristic selection of AI planning

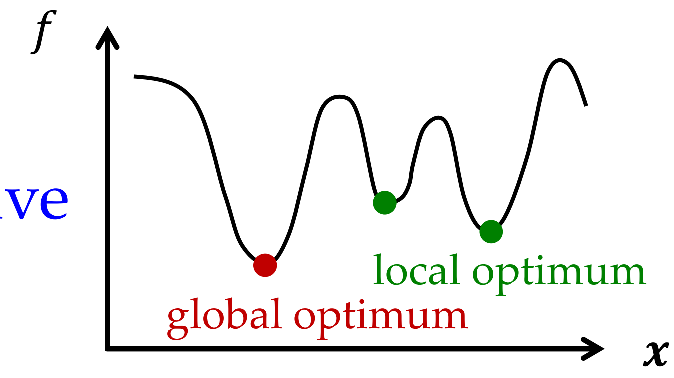
These tasks typically focus on a single type of hyperparameter

However, due to the increasing complexity of real-world problem modeling, there are many algorithms whose performance rests on *multiple types of hyperparameters*.

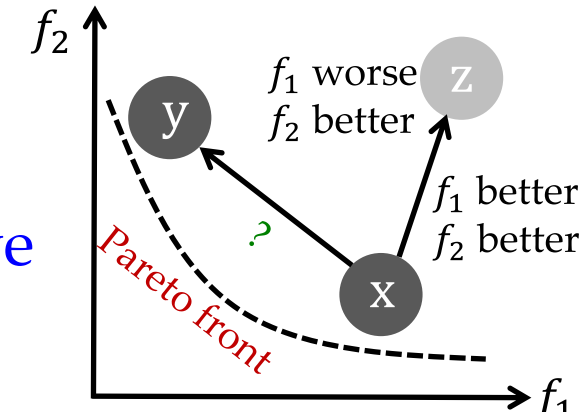
Background

Multi-objective optimization problems (MOPs)

Single-objective

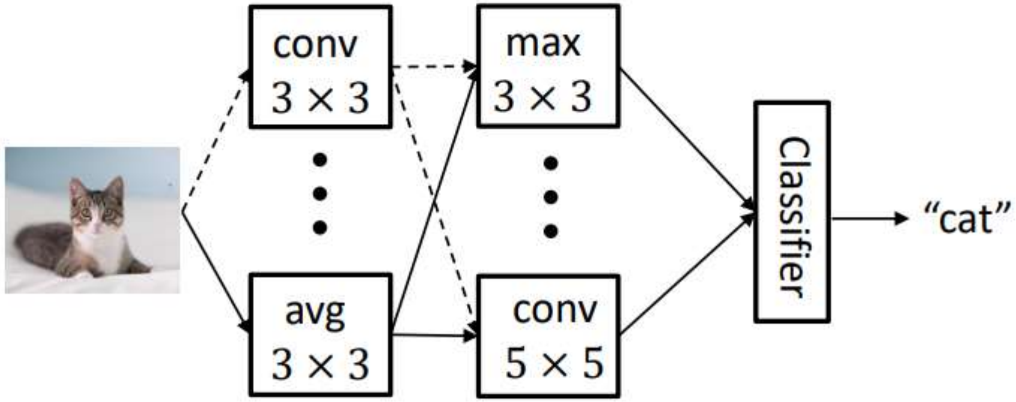


Multi-objective



Neural architecture search

- Max: accuracy
- Min: computation cost



Algorithm 1: MOEA/D

Parameters: Population size N, number T of iterations

Initialize a population $\{\boldsymbol{x}^{(i)}\}_{i=1}^{N}$ of solutions, and a corresponding set $W=\{\boldsymbol{w}^{(i)}\}_{i=1}^{N}$ of weight vectors ;

$t=0$;

while $t<T$ **do**

  for $i=1:N$ **do**

    Randomly select parent solutions from the neighborhood of $\boldsymbol{w}^{(i)}$, denoted as $\Theta^{\boldsymbol{w}^{(i)}}$;

    Use crossover and mutation operators to generate an offspring solution $\boldsymbol{x}'^{(i)}$;

    Evaluate the offspring solution to obtain $\boldsymbol{F}(\boldsymbol{x}'^{(i)})$;

    Update the ideal point $\boldsymbol{z}^*$. That is, for any $j\in\{1,2,\ldots,m\}$, if $f_j(\boldsymbol{x}'^{(i)})<z_j^*$, then $z_j^*=f_j(\boldsymbol{x}'^{(i)})$;

    Update the corresponding solution of each sub-problem within $\Theta^{\boldsymbol{w}^{(i)}}$ by $\boldsymbol{x}'^{(i)}$. That is, for each $\boldsymbol{w}^{(j)}\in\Theta^{\boldsymbol{w}^{(i)}}$, if $g(\boldsymbol{x}'^{(i)}\mid\boldsymbol{w}^{(j)},\boldsymbol{z}^*)<g(\boldsymbol{x}^{(j)}\mid\boldsymbol{w}^{(j)},\boldsymbol{z}^*)$, then $\boldsymbol{x}^{(j)}=\boldsymbol{x}'^{(i)}$

  end

  $t=t+1$

end

Complex and hard to tune

Background

The limitations and further research in a recent paper [Adriaensen et al., JAIR'23]

Automated Dynamic Algorithm Configuration

Steven Adriaensen ADRIAENS@CS.UNI-FREIBURG.DE
André Biedenkapp BIEDENKA@CS.UNI-FREIBURG.DE
Gresa Shala SHALAG@CS.UNI-FREIBURG.DE
Noor Awad AWAD@CS.UNI-FREIBURG.DE
University of Freiburg, Machine Learning Lab

Theresa Eimer EIMER@TNT.UNI-HANNOVER.DE
Marius Lindauer LINDAUER@TNT.UNI-HANNOVER.DE
Leibniz University Hannover, Institute for Information Processing

Frank Hutter FH@CS.UNI-FREIBURG.DE
University of Freiburg, Machine Learning Lab & Bosch Center for Artificial Intelligence

Frank Hutter

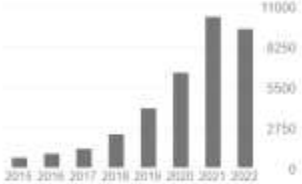
FOLLOWING

Professor of Computer Science, University of Freiburg, Germany
Verified email at cs.uni-freiburg.de - Homepage
AutoML Neural Architecture Search Meta-Learning Deep Learning Machine Learning

TITLE	CITED BY	YEAR
Decoupled weight decay regularization I Loshchilov, F Hutter arXiv preprint arXiv:1711.05101	5552 *	2017
Sgdr: Stochastic gradient descent with warm restarts I Loshchilov, F Hutter arXiv preprint arXiv:1608.03983	3815	2016
Sequential model-based optimization for general algorithm configuration F Hutter, HH Hoos, K Leyton-Brown International conference on learning and intelligent optimization, 507-523	2472	2011

Cited by VIEW ALL

	All	Since 2017
Citations	38773	34076
h-index	75	65
i10-index	154	151



7.2 Limitations and Further Research

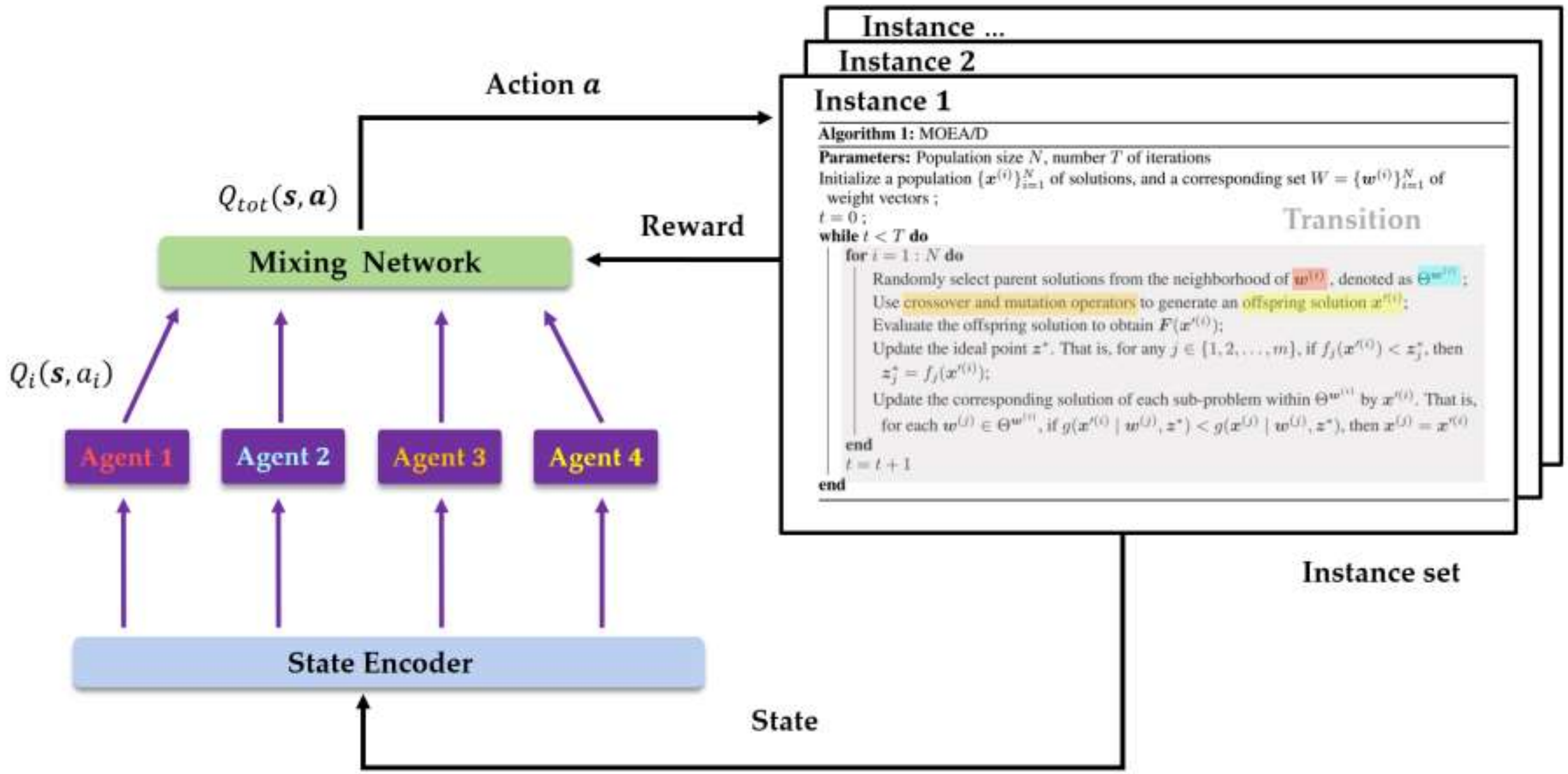
While these case studies and other previous applications provide a "proof of concept" for automated DAC, we point out that much remains to be done to unlock its full potential, and we hope that this work may serve as a stepping stone for further exploring this promising line of research. In what remains, we will discuss some of the limitations of contemporary work and provide specific directions for future research.

Jointly configuring many parameters: While static approaches are capable of jointly configuring hundreds of parameters, the configuration space in contemporary DAC is typically much smaller, often considering only a single parameter. While the configuration space is smaller, the candidate solution space (i.e., the dynamic configuration policy space) grows exponentially with the number of reconfiguration points, in the worst case, and is thus typically drastically larger than static configuration policy spaces. Although modern techniques from reinforcement learning scale much better than ever before, we still know too little about the internal structure of DAC problems to handle this exploding space of possible policies. For example, not much is known regarding interaction effects of parameters in the DAC setting. If there should be only a few interaction effects between parameters as in static AC (Hutter et al., 2014; Wang et al., 2016), learning several independent policies might be a way forward.

How to dynamically adjust a complex algorithm with many hyperparameters?

MA-DAC

The configuration of a complex algorithm → MA-DAC → Cooperative multi-agent problem



one agent works to handle one type of hyperparameter

MA-DAC

State

- *Accessibility*. The state should be accessible in each step.
- *Representability*. The state should reflect the information in the optimization process.
- *Generalizability*. The learned policy is expected to generalize to inner and even outer classes of instances.

Index	Parts of state	Feature	Notes
0	1	$1/m$	m: Number of objectives
1	1	$1/D$	D: Number of variables
2	2	t/T	Computational budget that has been used
3	2	N_{stag}/T	Stagnant count ratio
4	3	HV_t	Hypervolume value
5	3	NDRatio_t	Ratio of non-dominated solutions
6	3	Dist_t	Average distance
7	3	$\text{HV}_t - \text{HV}_{t-1}$	Change of HV between steps t and $t-1$
8	3	$\text{NDRatio}_t - \text{NDRatio}_{t-1}$	Change of NDRatio between steps t and $t-1$
9	3	$\text{Dist}_t - \text{Dist}_{t-1}$	Change of Dist between steps t and $t-1$
10	3	$\text{Mean}(\text{List}(\text{HV}, t, 5))$	Mean of HV in the last 5 steps
11	3	$\text{Mean}(\text{List}(\text{NDRatio}, t, 5))$	Mean of NDRatio in the last 5 steps
12	3	$\text{Mean}(\text{List}(\text{Dist}, t, 5))$	Mean of Dist in the last 5 steps
13	3	$\text{Std}(\text{List}(\text{HV}, t, 5))$	Standard deviation of HV in the last 5 steps
14	3	$\text{Std}(\text{List}(\text{NDRatio}, t, 5))$	Standard deviation of NDRatio in the last 5 steps
15	3	$\text{Std}(\text{List}(\text{Dist}, t, 5))$	Standard deviation of Dist in the last 5 steps
16	3	$\text{Mean}(\text{List}(\text{HV}, t, t))$	Mean of HV in all the steps so far
17	3	$\text{Mean}(\text{List}(\text{NDRatio}, t, t))$	Mean of NDRatio in all the steps so far
18	3	$\text{Mean}(\text{List}(\text{Dist}, t, t))$	Mean of Dist in all the steps so far
19	3	$\text{Std}(\text{List}(\text{HV}, t, t))$	Standard deviation of HV in all the steps so far
20	3	$\text{Std}(\text{List}(\text{NDRatio}, t, t))$	Standard deviation of NDRatio in all the steps so far
21	3	$\text{Std}(\text{List}(\text{Dist}, t, t))$	Standard deviation of Dist in all the steps so far

MA-DAC

Action: the value of the hyperparameter

Reward:

$$r_t^1 = \max\{f(s_t) - f(s_{t+1}), 0\}$$

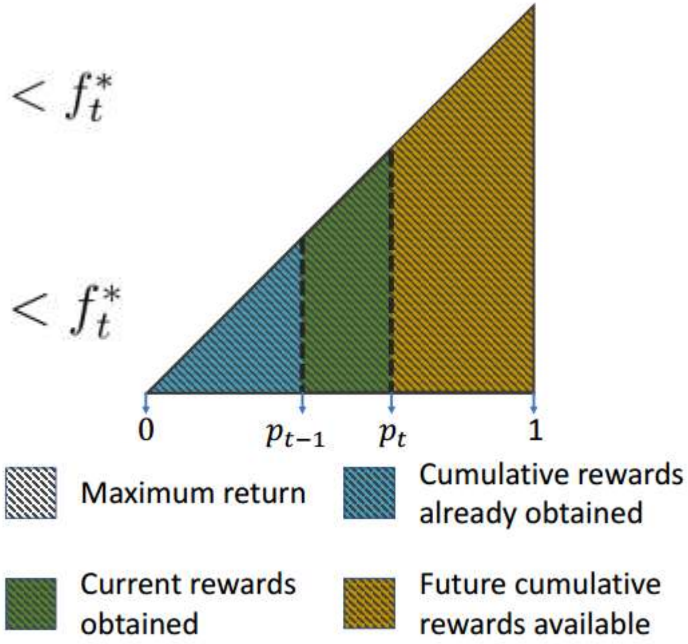
$$r_t^2 = \begin{cases} 10 & \text{if } f(s_{t+1}) < f_t^* \\ 1 & \text{else if } f(s_{t+1}) < f(s_t) \\ 0 & \text{otherwise} \end{cases}$$

$$r_t^3 = \max\left\{\frac{f(s_t) - f(s_{t+1})}{f(s_{t+1}) - f_{\text{opt}}}, 0\right\}$$

$$\Rightarrow$$

$$r_t = \begin{cases} (1/2)\cdot(p_t^2 - p_{t-1}^2) & \text{if } f(s_{t+1}) < f_t^* \\ 0 & \text{otherwise} \end{cases}$$

$$p_{t+1} = \begin{cases} \frac{f(s_0)-f(s_{t+1})}{f(s_0)} & \text{if } f(s_{t+1}) < f_t^* \\ p_t & \text{otherwise} \end{cases}$$



Transition: one step in RL is one generation in MOEA/D

MA-DAC – benchmark

We obtain the Multi-agent RL for Multi-objective optimization (MaMo) benchmark

Table 1: Overview of MARL benchmarks and their properties.

Benchmark	Heterogeneous	# of agents	Stochastic	Application scenarios
Matrix Games [5]	×	2	Low	Game
MPE [20]	×	2-3	Low	Game
MAgent [42]	×	2-1000	Low	Game
SMAC [28]	✓	2-30	Low	Game
Active Voltage Control [38]	×	3-38	Low	Control
MaMo (Ours)	✓	2-4	High	Optimization

We hope it can benefit the MARL community

Experiment

We investigate the following three research questions (RQs):

- RQ1: How does MA-DAC *perform* compared with the baseline and other tuning algorithms?
- RQ2: How is the *generalization ability* of MA-DAC?
- RQ3: How do the *different parts* of MA-DAC affect the performance?

Experiment – RQ1

Problem	M	MOEA/D	DQN	MA-UCB	MA-DAC
DTLZ2	3	4.605E-02 (3.54E-04) –	4.628E-02 (2.96E-04) –	4.671E-02 (3.70E-04) –	**3.807E-02** (5.05E-04)
	5	3.006E-01 (1.55E-03) –	3.016E-01 (1.34E-03) –	3.041E-01 (1.69E-03) –	**2.442E-01** (1.26E-02)
	7	4.455E-01 (1.41E-02) –	4.671E-01 (1.15E-02) –	4.826E-01 (9.59E-03) –	**3.944E-01** (1.17E-02)
WFG4	3	5.761E-02 (5.41E-04) –	6.920E-02 (1.20E-03) –	7.165E-02 (1.83E-03) –	**5.200E-02** (1.19E-03)
	5	3.442E-01 (1.21E-02) –	2.810E-01 (6.86E-03) –	2.859E-01 (6.77E-03) –	**1.868E-01** (2.81E-03)
	7	4.529E-01 (1.79E-02) –	3.725E-01 (1.14E-02) –	3.868E-01 (1.54E-02) –	**3.033E-01** (3.66E-03)
WFG6	3	6.938E-02 (5.50E-03) –	6.834E-02 (1.78E-02) –	6.601E-02 (1.00E-02) –	**4.831E-02** (8.95E-03)
	5	3.518E-01 (2.82E-03) –	3.160E-01 (2.40E-02) –	3.359E-01 (1.47E-02) –	**1.942E-01** (6.90E-03)
	7	4.869E-01 (3.03E-02) –	4.322E-01 (2.95E-02) –	4.389E-01 (3.41E-02) –	**3.112E-01** (4.93E-03)
Train: +/–/≈		0/9/0	0/9/0	0/9/0	
DTLZ4	3	6.231E-02 (8.85E-02) ≈	**5.590E-02** (5.77E-03) –	6.011E-02 (5.08E-03) –	6.700E-02 (6.14E-02)
	5	3.133E-01 (4.45E-02) ≈	3.457E-01 (1.61E-02) –	3.492E-01 (1.69E-02) –	**2.995E-01** (2.10E-02)
	7	4.374E-01 (2.57E-02) –	4.552E-01 (1.47E-02) –	4.756E-01 (2.01E-02) –	**4.182E-01** (1.21E-02)
WFG5	3	6.327E-02 (1.10E-03) –	6.212E-02 (5.54E-04) –	6.118E-02 (7.03E-04) –	**4.730E-02** (7.89E-04)
	5	3.350E-01 (9.77E-03) –	3.077E-01 (6.36E-03) –	3.036E-01 (8.83E-03) –	**1.811E-01** (3.02E-03)
	7	4.101E-01 (2.08E-02) –	4.996E-01 (1.32E-02) –	5.024E-01 (1.38E-02) –	**3.206E-01** (8.04E-03)
WFG7	3	5.811E-02 (6.31E-04) –	5.930E-02 (7.32E-04) –	6.014E-02 (7.11E-04) –	**4.066E-02** (5.31E-04)
	5	3.572E-01 (5.47E-03) –	2.993E-01 (1.43E-02) –	3.207E-01 (1.71E-02) –	**1.858E-01** (2.12E-03)
	7	5.236E-01 (2.19E-02) –	4.576E-01 (2.38E-02) –	4.879E-01 (2.75E-02) –	**3.258E-01** (1.25E-02)
WFG8	3	8.646E-02 (3.44E-03) –	9.280E-02 (1.06E-03) –	9.612E-02 (1.48E-03) –	**7.901E-02** (1.19E-03)
	5	4.258E-01 (8.42E-03) –	3.969E-01 (1.26E-02) –	3.956E-01 (1.32E-02) –	**2.479E-01** (7.20E-03)
	7	5.816E-01 (1.30E-02) –	5.575E-01 (1.39E-02) –	5.642E-01 (1.38E-02) –	**4.127E-01** (5.93E-03)
WFG9	3	5.817E-02 (1.24E-03) –	5.628E-02 (7.29E-04) –	7.953E-02 (2.45E-02) –	**4.159E-02** (6.10E-04)
	5	3.633E-01 (1.20E-02) –	3.258E-01 (1.61E-02) –	3.396E-01 (1.55E-02) –	**1.832E-01** (7.10E-03)
	7	5.538E-01 (2.63E-02) –	5.115E-01 (2.15E-02) –	5.227E-01 (1.79E-02) –	**3.278E-01** (7.21E-03)
Test: +/–/≈		0/13/2	0/15/0	0/15/0	

Significantly better on almost all the 24 problems

Good generalization ability

Experiment – RQ1

Statistic | Learning based | Heuristic

Problem	M	MOEA/D	MOEA/D-OP2	MOEA/D-FRRMAB	MA-DAC	MOEA/D-AWA	MOEA/D-OP2-AWA
DTLZ2	3	4.605E-02 (3.54E-04) –	4.691E-02 (3.97E-04) –	4.668E-02 (2.50E-04) –	**3.807E-02** (5.05E-04)	4.596E-02 (3.54E-04) –	4.670E-02 (3.30E-04) –
	5	3.006E-01 (1.55E-03) –	3.012E-01 (1.51E-03) –	3.031E-01 (1.29E-03) –	**2.442E-01** (1.26E-02)	2.900E-01 (2.73E-03) –	2.764E-01 (3.40E-03) –
	7	4.455E-01 (1.41E-02) –	4.551E-01 (4.43E-03) –	4.724E-01 (7.80E-03) –	**3.944E-01** (1.17E-02)	4.167E-01 (2.37E-02) –	4.436E-01 (8.67E-03) –
WFG4	3	5.761E-02 (5.41E-04) –	7.293E-02 (1.43E-03) –	7.097E-02 (1.63E-03) –	**5.200E-02** (1.19E-03)	5.748E-02 (7.11E-04) –	7.280E-02 (1.33E-03) –
	5	3.442E-01 (1.21E-02) –	2.761E-01 (6.39E-03) –	2.799E-01 (9.44E-03) –	**1.868E-01** (2.81E-03)	3.168E-01 (5.37E-03) –	2.648E-01 (8.15E-03) –
	7	4.529E-01 (1.79E-02) –	3.711E-01 (9.79E-03) –	3.778E-01 (1.01E-02) –	**3.033E-01** (3.66E-03)	4.285E-01 (1.55E-02) –	3.676E-01 (1.06E-02) –
WFG6	3	6.938E-02 (5.50E-03) –	6.714E-02 (1.59E-02) –	6.266E-02 (8.47E-03) –	**4.831E-02** (8.95E-03)	6.846E-02 (4.70E-03) –	6.078E-02 (1.16E-03) –
	5	3.518E-01 (2.82E-03) –	3.285E-01 (2.33E-02) –	3.272E-01 (1.61E-02) –	**1.942E-01** (6.90E-03)	3.190E-01 (3.93E-03) –	3.143E-01 (2.52E-02) –
	7	4.869E-01 (3.03E-02) –	4.797E-01 (3.04E-02) –	4.417E-01 (3.29E-02) –	**3.112E-01** (4.93E-03)	4.727E-01 (3.05E-02) –	4.770E-01 (3.24E-02) –
Train: average rank		3.11	3.00	2.89	1.00	2.56	2.67
Train: +/–/≈		0/9/0	0/9/0	0/9/0		0/9/0	0/9/0
DTLZ4	3	6.231E-02 (8.85E-02) ≈	6.226E-02 (4.05E-03) –	**5.782E-02** (3.48E-03) –	6.700E-02 (6.14E-02)	**4.597E-02** (3.66E-04) ≈	6.219E-02 (3.90E-03) –
	5	3.133E-01 (4.45E-02) ≈	3.413E-01 (1.48E-02) –	3.373E-01 (1.70E-02) –	**2.995E-01** (2.10E-02)	**2.816E-01** (3.24E-03) +	3.283E-01 (1.08E-02) –
	7	4.374E-01 (2.57E-02) –	4.519E-01 (1.15E-02) –	4.681E-01 (1.87E-02) –	**4.182E-01** (1.21E-02)	**3.696E-01** (1.32E-02) +	4.437E-01 (9.46E-03) –
WFG5	3	6.327E-02 (1.10E-03) –	6.177E-02 (8.01E-04) –	6.120E-02 (7.38E-04) –	**4.730E-02** (7.89E-04)	6.376E-02 (9.85E-04) –	6.168E-02 (4.61E-04) –
	5	3.350E-01 (9.77E-03) –	3.052E-01 (7.19E-03) –	3.033E-01 (8.69E-03) –	**1.811E-01** (3.02E-03)	3.173E-01 (5.33E-03) –	3.024E-01 (6.02E-03) –
	7	4.101E-01 (2.08E-02) –	4.988E-01 (1.04E-02) –	5.045E-01 (9.70E-03) –	**3.206E-01** (8.04E-03)	4.095E-01 (1.94E-02) –	4.865E-01 (1.28E-02) –
WFG7	3	5.811E-02 (6.31E-04) –	6.033E-02 (8.84E-04) –	5.976E-02 (7.44E-04) –	**4.066E-02** (5.31E-04)	5.837E-02 (6.25E-04) –	6.017E-02 (6.74E-04) –
	5	3.572E-01 (5.47E-03) –	2.941E-01 (9.66E-03) –	3.042E-01 (1.52E-02) –	**1.858E-01** (2.12E-03)	3.227E-01 (4.19E-03) –	2.885E-01 (1.25E-02) –
	7	5.236E-01 (2.19E-02) –	4.739E-01 (2.51E-02) –	4.762E-01 (2.74E-02) –	**3.258E-01** (1.25E-02)	5.004E-01 (3.80E-02) –	4.560E-01 (2.56E-02) –
WFG8	3	8.646E-02 (3.44E-03) –	9.598E-02 (1.22E-03) –	9.536E-02 (1.14E-03) –	**7.901E-02** (1.19E-03)	8.742E-02 (6.36E-04) –	9.572E-02 (8.39E-04) –
	5	4.258E-01 (8.42E-03) –	3.884E-01 (1.19E-02) –	3.917E-01 (9.00E-03) –	**2.479E-01** (7.20E-03)	4.216E-01 (1.18E-02) –	3.824E-01 (9.74E-03) –
	7	5.816E-01 (1.30E-02) –	5.587E-01 (1.56E-02) –	5.570E-01 (1.60E-02) –	**4.127E-01** (5.93E-03)	5.790E-01 (1.06E-02) –	5.632E-01 (1.27E-02) –
WFG9	3	5.817E-02 (1.24E-03) –	8.122E-02 (2.54E-02) –	6.445E-02 (1.72E-02) –	**4.159E-02** (6.10E-04)	5.809E-02 (1.45E-03) –	6.470E-02 (1.75E-02) –
	5	3.633E-01 (1.20E-02) –	3.300E-01 (1.47E-02) –	3.312E-01 (1.70E-02) –	**1.832E-01** (7.10E-03)	3.517E-01 (2.19E-02) –	3.024E-01 (1.36E-02) –
	7	5.538E-01 (2.63E-02) –	5.001E-01 (2.59E-02) –	5.145E-01 (2.82E-02) –	**3.278E-01** (7.21E-03)	5.108E-01 (2.65E-02) –	4.861E-01 (2.78E-02) –
Test: average rank		3.13	2.87	2.80	1.20	2.53	2.80
Test: +/–/≈		0/13/2	0/15/0	0/15/0		2/12/1	0/15/0

Comparison with specific tuning approaches for MOEA/D

Experiment – RQ2

Problem	M	MA-DAC (M)	MA-DAC (3)	MA-DAC (5)	MA-DAC (7)
DTLZ2	3	3.839E-02 (5.35E-04) +	**3.807E-02** (5.05E-04) +	3.830E-02 (7.24E-04) +	3.837E-02 (5.69E-04) +
	5	2.468E-01 (7.55E-03) +	2.472E-01 (1.56E-02) +	**2.442E-01** (1.26E-02) +	2.569E-01 (1.39E-02) +
	7	3.921E-01 (8.84E-03) +	**3.880E-01** (1.02E-02) +	4.081E-01 (1.52E-02) +	3.944E-01 (1.17E-02) +
WFG4	3	5.220E-02 (9.83E-04) +	**5.200E-02** (1.19E-03) +	5.236E-02 (1.10E-03) +	5.302E-02 (9.78E-04) +
	5	**1.850E-01** (3.14E-03) +	1.867E-01 (3.01E-03) +	1.868E-01 (2.81E-03) +	1.853E-01 (2.67E-03) +
	7	3.091E-01 (5.80E-03) +	3.104E-01 (7.14E-03) +	3.100E-01 (5.89E-03) +	**3.033E-01** (3.66E-03) +
WFG6	3	5.078E-02 (1.20E-02) +	4.831E-02 (8.95E-03) +	**4.599E-02** (9.48E-03) +	5.206E-02 (1.64E-02) +
	5	1.971E-01 (6.40E-03) +	2.003E-01 (6.26E-03) +	**1.942E-01** (6.90E-03) +	1.957E-01 (6.67E-03) +
	7	3.114E-01 (5.08E-03) +	3.242E-01 (9.24E-03) +	3.129E-01 (5.71E-03) +	**3.112E-01** (4.93E-03) +
DTLZ4	3	**6.171E-02** (3.67E-02) +	6.700E-02 (6.14E-02) +	6.618E-02 (4.62E-02) +	8.088E-02 (7.12E-02) +
	5	3.044E-01 (1.66E-02) ≈	**2.974E-01** (1.94E-02) +	2.995E-01 (2.10E-02) ≈	3.036E-01 (1.69E-02) ≈
	7	4.271E-01 (1.45E-02) +	4.313E-01 (1.39E-02) ≈	4.327E-01 (2.15E-02) ≈	**4.182E-01** (1.21E-02) +
WFG5	3	**4.721E-02** (7.15E-04) +	4.730E-02 (7.89E-04) +	4.733E-02 (8.10E-04) +	4.746E-02 (5.90E-04) +
	5	1.811E-01 (2.59E-03) +	1.817E-01 (2.96E-03) +	1.811E-01 (3.02E-03) +	**1.808E-01** (2.83E-03) +
	7	3.256E-01 (5.49E-03) +	3.266E-01 (8.98E-03) +	3.263E-01 (9.73E-03) +	**3.206E-01** (8.04E-03) +
WFG7	3	4.076E-02 (5.33E-04) +	**4.066E-02** (5.31E-04) +	4.077E-02 (5.12E-04) +	4.124E-02 (4.98E-04) +
	5	1.839E-01 (2.38E-03) +	1.881E-01 (3.70E-03) +	1.858E-01 (2.12E-03) +	**1.836E-01** (2.21E-03) +
	7	3.368E-01 (1.54E-02) +	3.461E-01 (1.97E-02) +	3.390E-01 (1.38E-02) +	**3.258E-01** (1.25E-02) +
WFG8	3	**7.828E-02** (1.46E-03) +	7.901E-02 (1.19E-03) +	7.921E-02 (1.36E-03) +	7.944E-02 (1.30E-03) +
	5	2.506E-01 (1.11E-02) +	2.653E-01 (1.51E-02) +	**2.479E-01** (7.20E-03) +	2.532E-01 (9.28E-03) +
	7	4.303E-01 (1.49E-02) +	4.364E-01 (1.38E-02) +	4.242E-01 (9.08E-03) +	**4.127E-01** (5.93E-03) +
WFG9	3	4.324E-02 (7.07E-04) +	**4.159E-02** (6.10E-04) +	4.359E-02 (1.00E-02) +	6.415E-02 (2.64E-02) +
	5	1.858E-01 (7.63E-03) +	**1.814E-01** (4.59E-03) +	1.832E-01 (7.10E-03) +	1.918E-01 (1.13E-02) +
	7	3.328E-01 (1.02E-02) +	3.298E-01 (1.03E-02) +	3.307E-01 (1.37E-02) +	**3.278E-01** (7.21E-03) +
Test: average rank	3	1.4	1.8	2.8	4.0
	5	2.6	2.8	2.2	2.4
	7	2.6	3.4	3.0	1.0
Test: +/−/≈	3	5/0/0	5/0/0	5/0/0	5/0/0
	5	4/0/1	5/0/0	4/0/1	4/0/1
	7	5/0/0	4/0/1	4/0/1	5/0/0

Train on problems with three objectives, but test on all problems

MA-DAC (M) is robust

Good generalization ability

Experiment – RQ3

Problem	M	MA-DAC (M) w/o 1	MA-DAC (M) w/o 2	MA-DAC (M) w/o 3	MA-DAC (M) w/o 4	MA-DAC (M)
DTLZ2	3	4.656E-02 (3.80E-04) −	3.914E-02 (8.43E-04) −	3.935E-02 (6.72E-04) −	3.919E-02 (5.91E-04) −	**3.839E-02** (5.35E-04)
	5	3.086E-01 (7.24E-03) −	2.619E-01 (8.99E-03) −	2.503E-01 (1.30E-02) −	**2.433E-01** (1.59E-02) ≈	2.468E-01 (7.55E-03)
	7	4.970E-01 (1.26E-02) −	4.067E-01 (1.20E-02) −	4.003E-01 (1.19E-02) −	4.228E-01 (1.25E-02) −	**3.921E-01** (8.84E-03)
WFG4	3	7.222E-02 (1.93E-03) −	5.484E-02 (1.01E-03) −	5.410E-02 (8.85E-04) −	5.410E-02 (8.85E-04) −	**5.220E-02** (9.83E-04)
	5	2.868E-01 (1.01E-02) −	1.879E-01 (3.76E-03) ≈	**1.845E-01** (2.17E-03) +	1.846E-01 (2.39E-03) +	1.850E-01 (3.14E-03)
	7	3.758E-01 (1.33E-02) −	3.102E-01 (6.34E-03) −	**3.020E-01** (3.99E-03) ≈	3.032E-01 (4.32E-03) ≈	3.091E-01 (5.80E-03)
WFG6	3	6.864E-02 (8.14E-03) −	5.338E-02 (1.37E-02) −	6.543E-02 (1.69E-02) −	6.067E-02 (2.11E-02) ≈	**5.078E-02** (1.20E-02)
	5	3.480E-01 (1.34E-02) −	2.005E-01 (5.21E-03) −	1.996E-01 (6.51E-03) −	1.979E-01 (6.76E-03) −	**1.971E-01** (6.40E-03)
	7	4.784E-01 (3.37E-02) −	3.147E-01 (5.85E-03) −	3.162E-01 (6.15E-03) −	3.147E-01 (5.73E-03) −	**3.114E-01** (5.08E-03)
Train: +/−/≈		0/9/0	0/8/1	1/7/1	1/5/3	
DTLZ4	3	6.463E-02 (3.85E-02) −	6.242E-02 (4.07E-02) −	**4.496E-02** (2.45E-03) +	4.496E-02 (2.45E-03) +	6.171E-02 (3.67E-02)
	5	3.497E-01 (1.41E-02) −	3.061E-01 (2.12E-02) ≈	3.054E-01 (1.55E-02) ≈	3.130E-01 (1.81E-02) −	**3.044E-01** (1.66E-02)
	7	4.853E-01 (1.82E-02) −	4.275E-01 (1.90E-02) −	**4.208E-01** (1.52E-02) ≈	4.289E-01 (2.07E-02) −	4.271E-01 (1.45E-02)
WFG5	3	6.189E-02 (7.40E-04) −	4.827E-02 (6.31E-04) −	4.766E-02 (5.72E-04) ≈	4.766E-02 (5.72E-04) ≈	**4.721E-02** (7.15E-04)
	5	3.202E-01 (9.77E-03) −	1.821E-01 (2.73E-03) ≈	1.835E-01 (2.90E-03) −	1.822E-01 (2.31E-03) −	**1.811E-01** (2.59E-03)
	7	4.948E-01 (1.47E-02) −	3.290E-01 (8.87E-03) −	3.310E-01 (7.70E-03) −	3.261E-01 (9.26E-03) −	**3.256E-01** (5.49E-03)
WFG7	3	6.004E-02 (9.45E-04) −	4.250E-02 (5.82E-04) −	4.150E-02 (6.60E-04) −	4.150E-02 (6.60E-04) −	**4.076E-02** (5.33E-04)
	5	3.402E-01 (2.49E-02) −	1.873E-01 (3.67E-03) ≈	**1.826E-01** (2.60E-03) +	1.847E-01 (2.80E-03) ≈	1.839E-01 (2.38E-03)
	7	4.877E-01 (5.70E-02) −	3.393E-01 (1.23E-02) −	3.373E-01 (1.16E-02) −	3.377E-01 (1.59E-02) −	**3.368E-01** (1.54E-02)
WFG8	3	9.661E-02 (1.87E-03) −	8.374E-02 (1.70E-03) −	8.029E-02 (1.29E-03) −	8.029E-02 (1.29E-03) −	**7.828E-02** (1.46E-03)
	5	4.119E-01 (1.30E-02) −	2.695E-01 (1.49E-02) −	2.571E-01 (1.09E-02) −	2.632E-01 (9.96E-03) −	**2.506E-01** (1.11E-02)
	7	5.830E-01 (1.59E-02) −	4.345E-01 (9.27E-03) −	**4.260E-01** (8.71E-03) −	4.322E-01 (1.12E-02) −	4.303E-01 (1.49E-02)
WFG9	3	5.894E-02 (9.24E-04) −	**4.321E-02** (7.50E-04) −	4.650E-02 (1.40E-02) −	4.650E-02 (1.40E-02) −	4.324E-02 (7.07E-04)
	5	3.148E-01 (2.06E-02) −	1.875E-01 (5.14E-03) −	1.941E-01 (6.45E-03) −	1.865E-01 (9.02E-03) −	**1.858E-01** (7.63E-03)
	7	5.069E-01 (2.53E-02) −	3.433E-01 (1.35E-02) −	3.402E-01 (1.00E-02) −	3.368E-01 (1.05E-02) −	**3.328E-01** (1.02E-02)
Test: +/−/≈		0/15/0	0/12/3	2/10/3	1/12/2	

Necessary to tune each hyperparameter

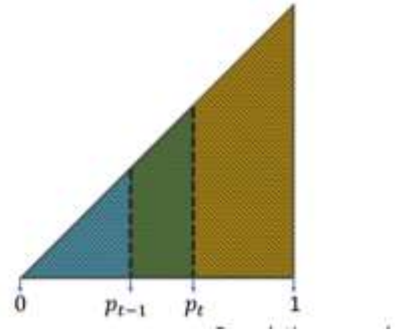
Experiment – RQ3

$$r_t^1 = \max\{f(s_t) - f(s_{t+1}), 0\}$$

$$r_t^2 = \begin{cases} 10 & \text{if } f(s_{t+1}) < f_t^* \\ 1 & \text{else if } f(s_{t+1}) < f(s_t) \\ 0 & \text{otherwise} \end{cases}$$

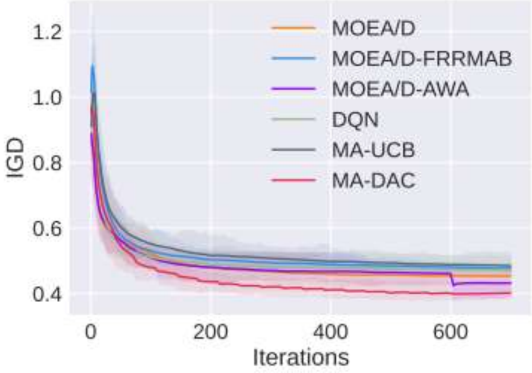
$$r_t^3 = \max\left\{\frac{f(s_t) - f(s_{t+1})}{f(s_{t+1}) - f_{\text{opt}}}, 0\right\}$$

$$r_t = \begin{cases} (1/2) \cdot (p_t^2 - p_{t-1}^2) & \text{if } f(s_{t+1}) < f_t^* \\ 0 & \text{otherwise} \end{cases}$$

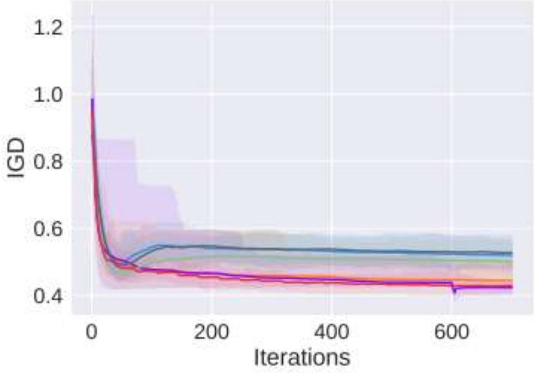


Problem	M	MA-DAC-R1	MA-DAC-R2	MA-DAC-R3	MA-DAC
DTLZ2	3	4.223E-02 (2.50E-03) −	3.853E-02 (5.58E-04) −	3.809E-02 (4.64E-04) ≈	**3.807E-02** (5.05E-04)
	5	2.401E-01 (8.27E-03) −	2.726E-01 (1.51E-02) ≈	**2.364E-01** (1.04E-02) +	2.442E-01 (1.26E-02)
	7	4.142E-01 (1.12E-02) −	4.248E-01 (1.30E-02) −	4.215E-01 (9.03E-03) −	**3.944E-01** (1.17E-02)
WFG4	3	5.989E-02 (5.60E-03) ≈	5.255E-02 (1.14E-03) −	5.309E-02 (8.02E-04) −	**5.200E-02** (1.19E-03)
	5	1.848E-01 (2.61E-03) +	1.851E-01 (2.43E-03) +	**1.846E-01** (2.20E-03) +	1.868E-01 (2.81E-03)
	7	3.028E-01 (3.19E-03) +	**3.008E-01** (3.51E-03) ≈	3.029E-01 (3.36E-03) ≈	3.033E-01 (3.66E-03)
WFG6	3	7.920E-02 (1.81E-02) +	4.909E-02 (1.50E-02) −	**4.814E-02** (1.22E-02) ≈	4.831E-02 (8.95E-03)
	5	1.977E-01 (6.17E-03) −	2.037E-01 (4.49E-03) −	1.975E-01 (5.78E-03) −	**1.942E-01** (6.90E-03)
	7	**3.110E-01** (4.86E-03) −	3.151E-01 (5.01E-03) ≈	3.148E-01 (4.05E-03) −	3.112E-01 (4.93E-03)
Train: average rank		2.67	3.11	2.11	2.11
Train: +/−/≈		3/5/1	1/5/3	2/4/3	
DTLZ4	3	**5.567E-02** (7.33E-03) −	7.236E-02 (6.19E-02) −	6.144E-02 (5.10E-02) ≈	6.700E-02 (6.14E-02)
	5	3.119E-01 (1.91E-02) −	3.221E-01 (2.12E-02) −	3.119E-01 (1.58E-02) −	**2.995E-01** (2.10E-02)
	7	4.354E-01 (1.29E-02) −	4.385E-01 (1.23E-02) −	4.275E-01 (1.60E-02) −	**4.182E-01** (1.21E-02)
WFG5	3	4.841E-02 (7.78E-04) −	4.763E-02 (7.73E-04) −	4.773E-02 (6.58E-04) −	**4.730E-02** (7.89E-04)
	5	1.823E-01 (2.49E-03) ≈	1.818E-01 (2.90E-03) −	1.812E-01 (3.06E-03) ≈	**1.811E-01** (3.02E-03)
	7	3.212E-01 (6.60E-03) ≈	**3.174E-01** (6.43E-03) ≈	3.196E-01 (5.99E-03) ≈	3.206E-01 (8.04E-03)
WFG7	3	4.555E-02 (1.26E-03) ≈	4.076E-02 (5.41E-04) −	4.168E-02 (6.40E-04) −	**4.066E-02** (5.31E-04)
	5	1.842E-01 (3.28E-03) ≈	1.865E-01 (2.93E-03) +	**1.841E-01** (3.95E-03) +	1.858E-01 (2.12E-03)
	7	3.335E-01 (1.09E-02) +	**3.199E-01** (9.86E-03) −	3.271E-01 (9.65E-03) ≈	3.258E-01 (1.25E-02)
WFG8	3	8.914E-02 (2.96E-03) ≈	7.911E-02 (1.06E-03) −	8.199E-02 (1.96E-03) −	**7.901E-02** (1.19E-03)
	5	2.551E-01 (1.02E-02) −	2.628E-01 (1.22E-02) −	2.541E-01 (9.08E-03) −	**2.479E-01** (7.20E-03)
	7	4.163E-01 (9.54E-03) ≈	**4.115E-01** (9.80E-03) ≈	4.197E-01 (7.52E-03) −	4.127E-01 (5.93E-03)
WFG9	3	5.003E-02 (9.00E-03) −	4.208E-02 (6.56E-04) −	4.428E-02 (9.97E-03) −	**4.159E-02** (6.10E-04)
	5	1.929E-01 (8.84E-03) ≈	**1.819E-01** (5.73E-03) −	1.951E-01 (9.83E-03) −	1.832E-01 (7.10E-03)
	7	3.342E-01 (8.56E-03) −	3.322E-01 (8.89E-03) −	3.327E-01 (8.02E-03) −	**3.278E-01** (7.21E-03)
Test: average rank		3.27	2.47	2.67	1.60
Test: +/−/≈		1/7/7	1/12/2	1/10/4	

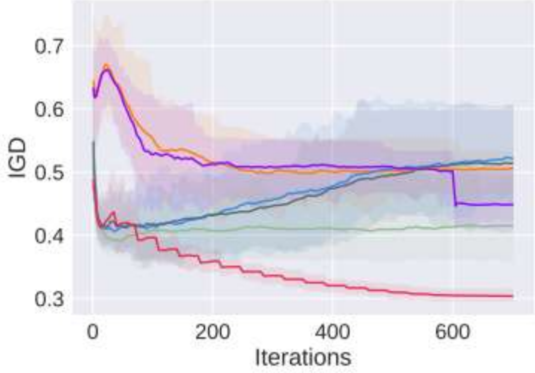
Experiment – IGD curve



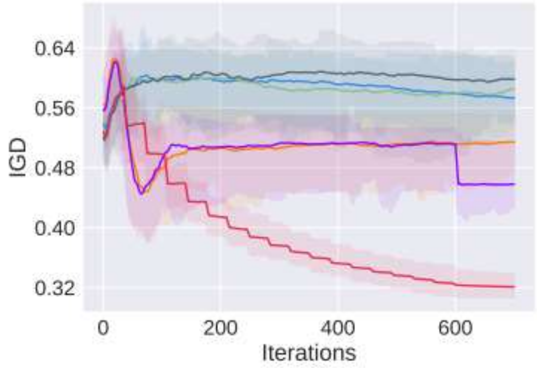
(a) DTLZ2_7



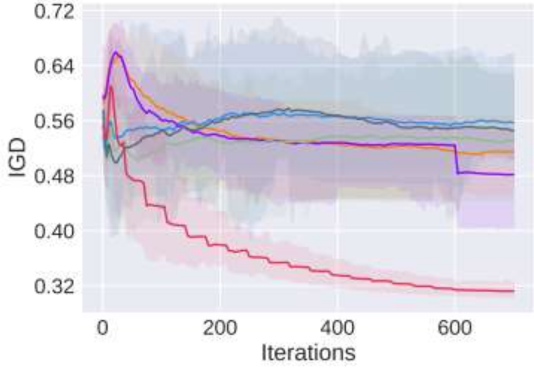
(b) DTLZ4_7



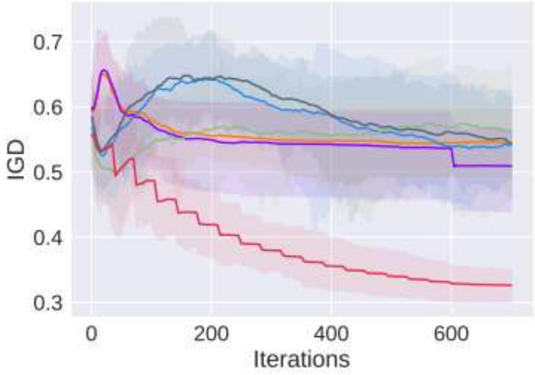
(c) WFG4_7



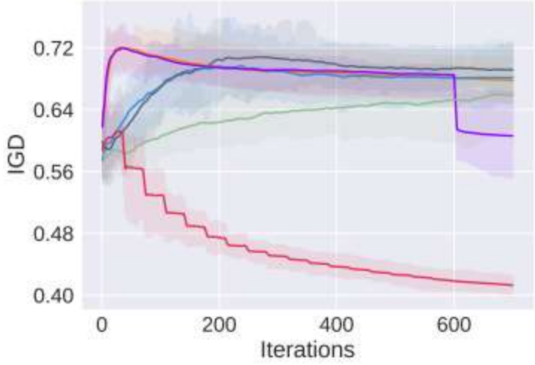
(d) WFG5_7



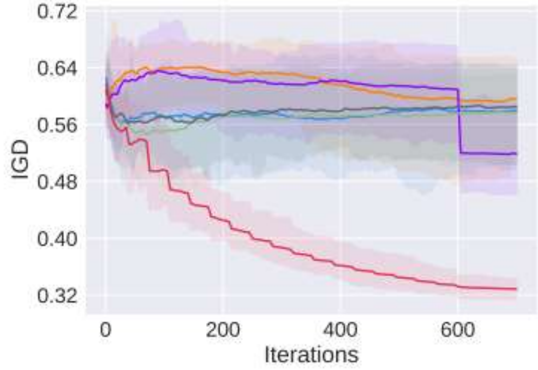
(e) WFG6_7



(f) WFG7_7



(g) WFG8_7



(h) WFG9_7

Experiment – Single DQN

Problem	M	DQN-1	DQN-2	DQN-3	DQN-4
DTLZ2	3	**4.100E-02** (7.64E-04) +	4.283E-02 (5.50E-04) +	4.284E-02 (3.47E-04) +	4.244E-02 (5.46E-04) +
	5	**2.408E-01** (1.03E-02) +	2.469E-01 (3.31E-03) +	2.531E-01 (3.43E-03) +	2.530E-01 (4.35E-03) +
	7	**4.064E-01** (1.14E-02) +	4.164E-01 (8.81E-03) +	4.149E-01 (8.92E-03) +	4.132E-01 (8.89E-03) +
WFG4	3	6.136E-02 (1.80E-03) +	6.383E-02 (1.92E-03) +	**5.885E-02** (1.50E-03) +	6.031E-02 (1.24E-03) +
	5	**1.887E-01** (2.15E-03) +	2.264E-01 (4.37E-03) +	2.351E-01 (5.49E-03) +	2.274E-01 (3.69E-03) +
	7	**3.018E-01** (3.14E-03) +	3.421E-01 (9.24E-03) +	3.427E-01 (9.67E-03) +	3.419E-01 (9.62E-03) +
WFG6	3	5.149E-02 (1.17E-02) +	**5.104E-02** (8.75E-03) +	6.538E-02 (2.37E-02) +	6.120E-02 (2.10E-02) +
	5	**1.991E-01** (9.10E-03) +	2.642E-01 (2.21E-02) +	2.665E-01 (1.73E-02) +	2.712E-01 (2.29E-02) +
	7	**3.204E-01** (7.11E-03) +	3.992E-01 (2.64E-02) +	3.967E-01 (2.29E-02) +	3.824E-01 (1.88E-02) +
Train: +/−/≈		9/0/0	9/0/0	9/0/0	9/0/0
DTLZ4	3	**4.697E-02** (3.78E-03) +	5.249E-02 (1.14E-02) +	6.116E-02 (5.10E-02) +	5.469E-02 (3.65E-02) +
	5	**3.074E-01** (1.24E-02) +	3.136E-01 (1.19E-02) +	3.207E-01 (1.26E-02) +	3.243E-01 (1.75E-02) +
	7	4.263E-01 (1.28E-02) +	**4.179E-01** (1.44E-02) +	4.330E-01 (2.83E-02) +	4.211E-01 (1.21E-02) +
WFG5	3	**4.815E-02** (6.96E-04) +	5.173E-02 (7.46E-04) +	5.188E-02 (5.99E-04) +	5.175E-02 (7.38E-04) +
	5	**1.833E-01** (3.37E-03) +	2.532E-01 (5.66E-03) +	2.537E-01 (5.17E-03) +	2.560E-01 (6.80E-03) +
	7	**3.294E-01** (8.51E-03) +	4.053E-01 (1.80E-02) +	4.090E-01 (1.11E-02) +	4.088E-01 (1.34E-02) +
WFG7	3	**4.624E-02** (6.74E-04) +	4.841E-02 (1.20E-03) +	4.638E-02 (8.69E-04) +	4.690E-02 (8.85E-04) +
	5	**1.835E-01** (3.02E-03) +	2.476E-01 (7.11E-03) +	2.492E-01 (5.16E-03) +	2.445E-01 (9.30E-03) +
	7	**3.274E-01** (1.15E-02) +	3.956E-01 (1.96E-02) +	3.951E-01 (1.88E-02) +	3.843E-01 (1.81E-02) +
WFG8	3	8.658E-02 (1.30E-03) +	8.850E-02 (1.40E-03) +	**8.490E-02** (2.33E-03) +	8.555E-02 (1.59E-03) +
	5	**2.519E-01** (8.57E-03) +	3.216E-01 (1.01E-02) +	3.485E-01 (1.65E-02) +	3.328E-01 (1.29E-02) +
	7	**4.163E-01** (7.50E-03) +	5.153E-01 (1.61E-02) +	5.172E-01 (1.73E-02) +	4.970E-01 (1.12E-02) +
WFG9	3	7.747E-02 (2.46E-02) −	7.320E-02 (2.62E-02) ≈	**4.969E-02** (1.34E-02) +	6.138E-02 (2.42E-02) +
	5	**2.041E-01** (5.21E-03) +	2.542E-01 (1.03E-02) +	2.493E-01 (9.82E-03) +	2.486E-01 (1.18E-02) +
	7	**3.418E-01** (1.18E-02) +	4.088E-01 (1.71E-02) +	4.179E-01 (1.89E-02) +	4.069E-01 (1.73E-02) +
Test: +/−/≈		14/1/0	14/0/1	15/0/0	15/0/0

Single DQN is better than DQN

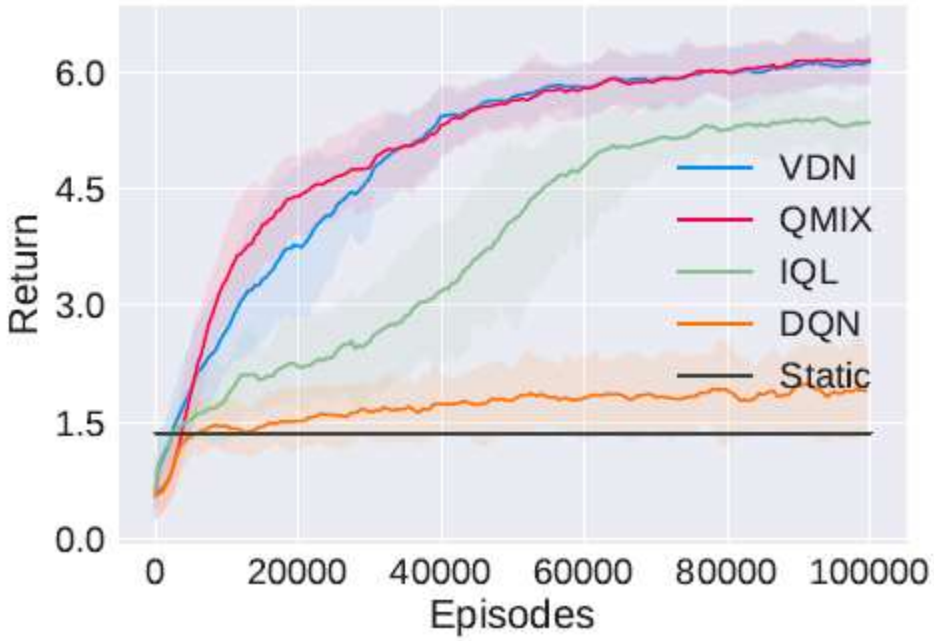
Experiment – MARL algorithms

Problem	M	DQN-1	VDN	IQL	QMIX
DTLZ2	3	4.100E-02 (7.64E-04)	**3.807E-02** (5.05E-04) +	3.933E-02 (4.29E-04) +	3.916E-02 (7.21E-04) +
	5	2.408E-01 (1.03E-02)	2.442E-01 (1.26E-02) ≈	**2.310E-01** (9.29E-03) +	2.419E-01 (1.37E-02) ≈
	7	4.064E-01 (1.14E-02)	**3.944E-01** (1.17E-02) +	4.138E-01 (9.79E-03) −	4.162E-01 (1.59E-02) −
WFG4	3	6.136E-02 (1.80E-03)	**5.200E-02** (1.19E-03) +	5.438E-02 (8.83E-04) +	5.206E-02 (1.16E-03) +
	5	1.887E-01 (2.15E-03)	1.868E-01 (2.81E-03) +	1.879E-01 (2.78E-03) ≈	**1.859E-01** (2.25E-03) +
	7	3.018E-01 (3.14E-03)	3.033E-01 (3.66E-03) ≈	3.046E-01 (3.55E-03) −	**2.998E-01** (4.21E-03) +
WFG6	3	5.149E-02 (1.17E-02)	4.831E-02 (8.95E-03) +	5.592E-02 (1.57E-02) ≈	**4.542E-02** (3.02E-03) +
	5	1.991E-01 (9.10E-03)	**1.942E-01** (6.90E-03) +	1.981E-01 (6.76E-03) ≈	1.975E-01 (6.98E-03) ≈
	7	3.204E-01 (7.11E-03)	**3.112E-01** (4.93E-03) +	3.148E-01 (3.34E-03) +	3.128E-01 (7.39E-03) +
Train: +/−/≈			7/0/2	4/2/3	6/1/2
DTLZ4	3	**4.697E-02** (3.78E-03)	6.700E-02 (6.14E-02) ≈	6.328E-02 (4.48E-02) −	5.094E-02 (2.38E-03) −
	5	3.074E-01 (1.24E-02)	**2.995E-01** (2.10E-02) +	3.021E-01 (1.62E-02) ≈	3.013E-01 (1.80E-02) +
	7	4.263E-01 (1.28E-02)	**4.182E-01** (1.21E-02) +	4.323E-01 (1.43E-02) ≈	4.303E-01 (1.95E-02) ≈
WFG5	3	4.815E-02 (6.96E-04)	**4.730E-02** (7.89E-04) +	4.818E-02 (6.24E-04) ≈	4.736E-02 (7.49E-04) +
	5	1.833E-01 (3.37E-03)	**1.811E-01** (3.02E-03) +	1.812E-01 (2.21E-03) +	1.813E-01 (2.54E-03) +
	7	3.294E-01 (8.51E-03)	3.206E-01 (8.04E-03) +	**3.173E-01** (6.91E-03) +	3.175E-01 (7.19E-03) +
WFG7	3	4.624E-02 (6.74E-04)	**4.066E-02** (5.31E-04) +	4.313E-02 (7.79E-04) +	4.077E-02 (4.94E-04) +
	5	1.835E-01 (3.02E-03)	1.858E-01 (2.12E-03) −	1.832E-01 (2.34E-03) ≈	**1.825E-01** (3.16E-03) ≈
	7	3.274E-01 (1.15E-02)	3.258E-01 (1.25E-02) ≈	**3.219E-01** (1.09E-02) +	3.226E-01 (1.12E-02) +
WFG8	3	8.658E-02 (1.30E-03)	**7.901E-02** (1.19E-03) +	8.652E-02 (2.75E-03) ≈	7.909E-02 (1.60E-03) +
	5	2.519E-01 (8.57E-03)	**2.479E-01** (7.20E-03) +	2.544E-01 (8.10E-03) ≈	2.496E-01 (9.83E-03) ≈
	7	4.163E-01 (7.50E-03)	4.127E-01 (5.93E-03) +	4.175E-01 (7.32E-03) ≈	**4.006E-01** (9.42E-03) +
WFG9	3	7.747E-02 (2.46E-02)	**4.159E-02** (6.10E-04) +	4.423E-02 (7.08E-04) +	4.167E-02 (5.92E-04) +
	5	2.041E-01 (5.21E-03)	**1.832E-01** (7.10E-03) +	1.915E-01 (8.87E-03) +	1.921E-01 (6.43E-03) +
	7	3.418E-01 (1.18E-02)	**3.278E-01** (7.21E-03) +	3.322E-01 (8.41E-03) +	3.298E-01 (8.46E-03) +
Test: +/−/≈			12/1/2	7/1/7	11/1/3

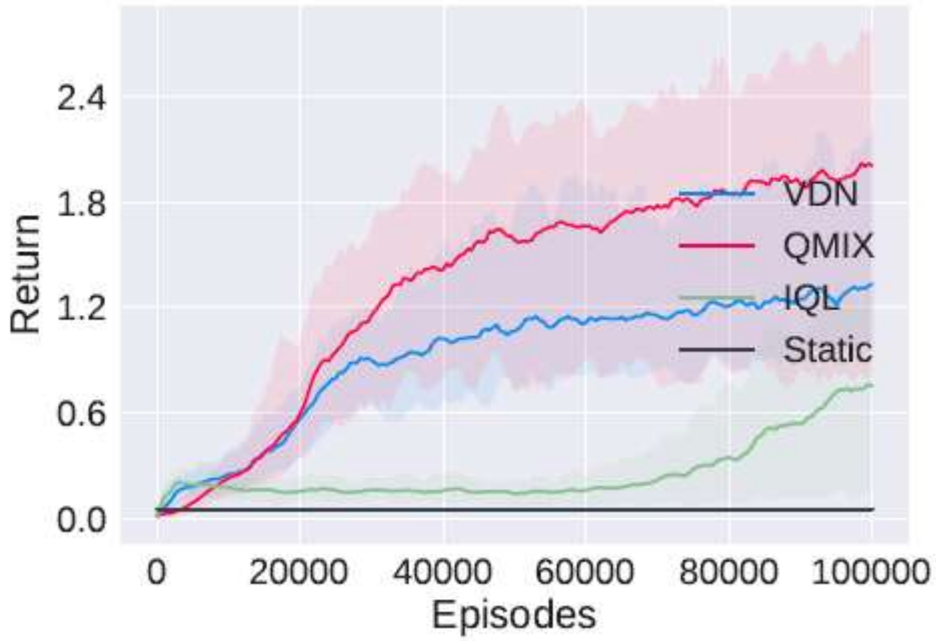
Equipped with different MARL algorithms, all the MA-DAC variants are better than DQN-1

Experiment – Other environments

Experiment on the DACBench [Eimer et al., IJCAI'21]



(a) $5D$-Sigmoid



(b) $10D$-Sigmoid

MA-DAC variants (VDN, QMIX, IQL) are better than DQN and static policy

Contribution

1) To the best of our knowledge, MA-DAC is the **first one** to address dynamic configuration of algorithms with multiple types of hyperparameters.

2) The contextual MMDP formulation of MA-DAC is analyzed, and experimental results show that the presented formulation works well and has good **generalization** ability.

3) The instantiation of configuring MOEA/D in this work can be used as a benchmark problem for MARL.
    1) The **heterogeneity** of MOEA/D's hyperparameters and the **stochasticity** of its search can promote the research of the MARL algorithms.
    2) Besides, the learned policies are **useful** for multi-objective optimization, which will facilitate the application of MARL.

Coauthors

NEURAL INFORMATION PROCESSING SYSTEMS

Spotlight (~3.4%)

Ke Xue[1]*, Jiacheng Xu[1]*, Lei Yuan[1],

Miqing Li[2], Chao Qian[1]†, Zongzhang Zhang[1], Yang Yu[1]
Nanjing University[1]
University of Birmingham[2]

Our code is available at
https://github.com/lamda-bbo/madac

Thank you!

xuek@lamda.nju.edu.cn