Evolutionary Transfer Optimization

Prof. Kay Chen TAN (IEEE Fellow)

Department of Computer Science

City University of Hong Kong

Contact: kaytan@cityu.edu.hk

IEEE WCCI 2020

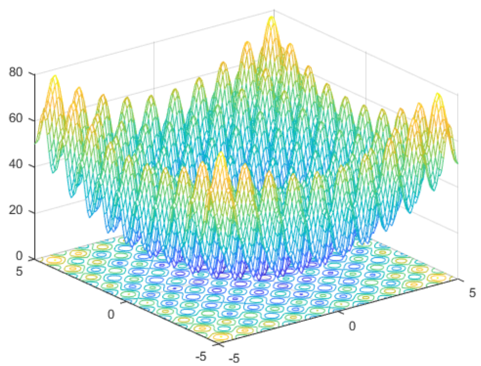
IEEE World Congress on Computational Intelligence

Glasgow, Scotland, United Kingdom – July 19-24, 2020

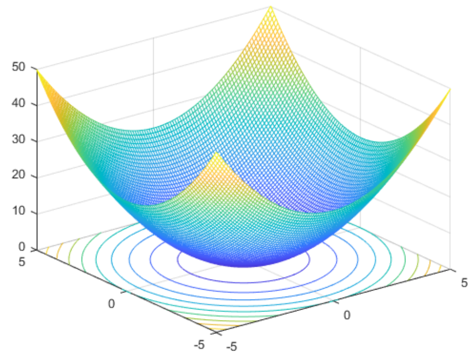
OUTLINE

Present Day Evolutionary Optimizers

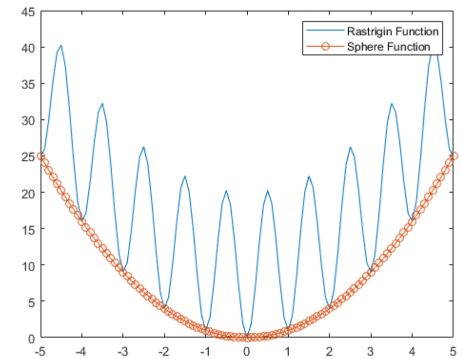
- Existing EAs often start a search from scratch or at “Ground Zero” knowledge state.
- It assumes all search problems are independent and so search capability does not grow or evolve along with the problem to be solved.
- But problems seldom exist in isolation and hence humans do not search from scratch.
- Common information exist between tasks/problems which can be effective for problem-solving when they are properly harnessed.



(a) Landscape of Rastrigin’s function (Task 1)



(b) Landscape of Sphere’s function (Task 2)



(c) Task 1 and Task 2 in 1 dimension

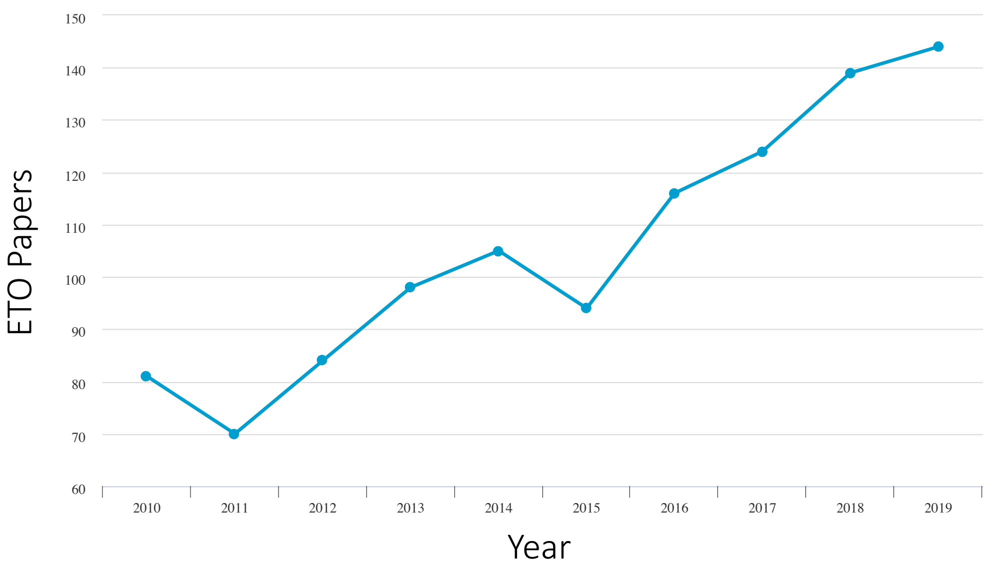
The solutions found along the optimization process of Sphere function can potentially be utilized to aid the optimization of the more complex Rastrigin's function.

Evolutionary Optimization + Transfer Learning

- Evolutionary Transfer Optimization (ETO)

  A paradigm that integrates EA solvers with knowledge learning and transfer across related domains for better optimization performance.
- The design of new knowledge learning and transfer approaches is necessary for developing advanced ETO algorithms.
- There are three issues to be considered in ETO, e.g., transferability, transfer component and transfer algorithm.

Scopus - Search with "Evolutionary Transfer Optimization".



Evolutionary Transfer Optimization (ETO)

Evolutionary Transfer Optimization (ETO)

1. Optimization in Uncertain Environment
2. Multi-Task Optimization
3. Complex Optimization Applications
4. Multi/Many-Objective Optimization
5. Machine Learning Applications

Homogeneous ETO

Knowledge transfer across problems sharing the same search space

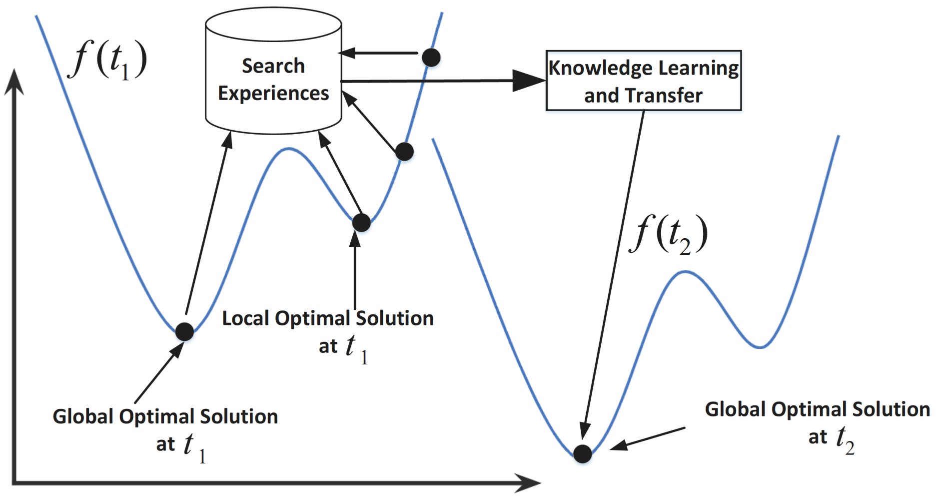
Heterogenous ETO

Knowledge transfer across problems possessing different search spaces

- K. C. Tan, L. Feng and M. Jiang, “Evolutionary Transfer Optimization - A New Frontier in Evolutionary Computation Research”, *IEEE Computational Intelligence Magazine*, submitted.

ETO for Optimization in Uncertain Environment

- The optimization problems may need to be solved in the presence of uncertainties, such as noise or approximations in function evaluation, dynamic changes of decision variables and/or fitness functions, and robustness.



ETO for solving a problem in uncertain environment

- How to design robust and incremental transfer learning methods for positive knowledge transfer while the evolutionary search progresses online?
- How to design ETO approaches considering data imbalance in knowledge learning and transfer?

- M. Jiang, Z. Huang, L. Qiu, W. Huang, and G. Yen, "Transfer Learning-Based Dynamic Multiobjective Optimization Algorithms", *IEEE Trans. on Evolutionary Computation*, 22(4), pp. 501–514, 2018.
- A. Simoes and E. Costa, "Improving Memory Usage in Evolutionary Algorithms for Changing Environments", *IEEE Congress on Evolutionary Computation*, pp. 276–283, 2007.
- I. Hatzakis and D. Wallace, "Dynamic Multi-Objective Optimization with Evolutionary Algorithms: A Forward-Looking Approach", *The 8th Annual Conf. on Genetic and Evolutionary Computation*, pp. 1201–1208, 2006.

ETO for Multi-task Optimization

- Multi-task optimization focuses on solving multiple self-contained tasks simultaneously. By transferring useful knowledge across tasks online, the solving of one problem may lead to the related problem being solved automatically.

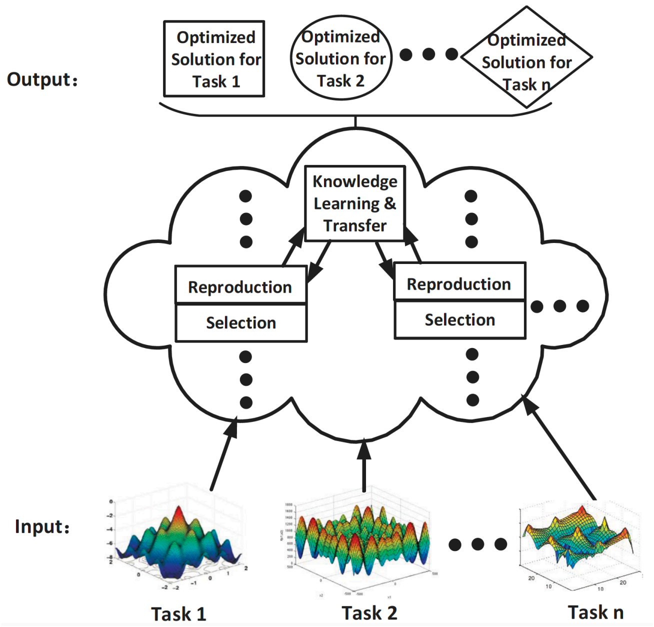


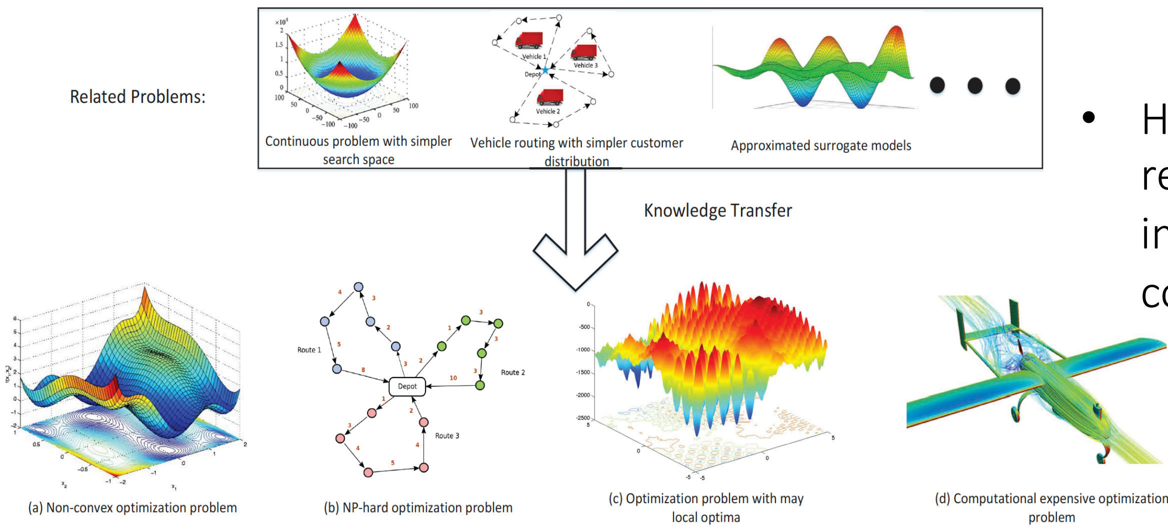
Illustration of Multi-task optimization

- How to evaluate correlation between tasks to ensure positive knowledge transfer in ETO?
- How to design ETO methods capable of solving many tasks simultaneously (with better performance and yet faster)?

- M. Gong, Z. Tang, H. Li, and J. Zhang, “Evolutionary Multitasking with Dynamic Resource Allocating Strategy”, *IEEE Trans. on Evolutionary Computation*, 23(5), pp. 858-869, 2019.
- J. Ding, C. Yang, Y. Jin, and T. Chai, “Generalized Multitasking for Evolutionary Optimization of Expensive Problems”, *IEEE Trans. on Evolutionary Computation*, 23(1), pp. 44-58, 2019.
- A. Gupta, Y. Ong, and L. Feng, “Multifactorial Evolution: Toward Evolutionary Multitasking”, *IEEE Trans. on Evolutionary Computation*, 20(3), pp. 343-357, 2016.

ETO for Complex Optimization Applications

- Many real-world applications involve complex optimization problems. By learning and transferring useful knowledge from related and simpler problem domains, ETO can help to deal with complex optimization problems.



- How to construct simple and related problems in ETO to improve the solving of a given complex problem?

Examples of complex optimization problems

- A. Chaabani and L. Said, "Transfer of Learning with the Coevolutionary Decomposition-Based Algorithm-II: A Realization on the Bi-level Production-distribution Planning System", *Applied Intelligence*, 49(3), pp. 963-982, 2019.
- L. Feng, Y. Ong, M. Lim, and I. Tsang, "Memetic Search with Interdomain Learning: A Realization Between CVRP and CARP", *IEEE Trans. on Evolutionary Computation*, 19(5), pp. 644-658, 2015.
- R. Santana, A. Mendiburu, and J. Lozano, "Structural Transfer Using EDAS: An Application to Multi-Marker Tagging SNP Selection", *IEEE Congress on Evolutionary Computation*, pp. 1-8, 2012.

ETO for Multi/Many-Objective Optimization

- Multi-objective problem (MOP) involves more than one objective function to be optimized simultaneously, e.g., when optimal decisions need to be taken in the presence of trade-offs between two or more conflicting objectives.

- ETO can be applied to solve multi- and many-objective optimization problems by transferring useful knowledge across the problems.

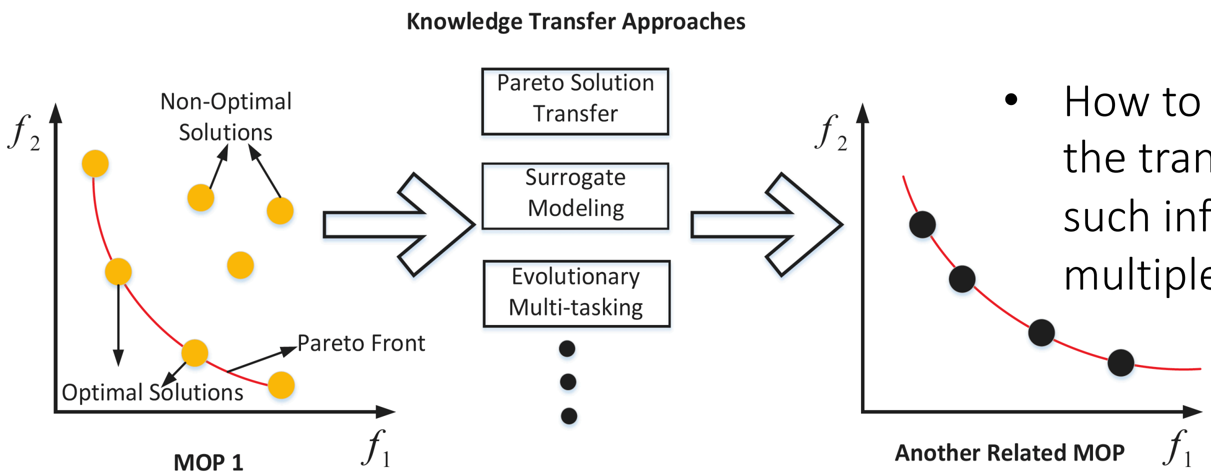


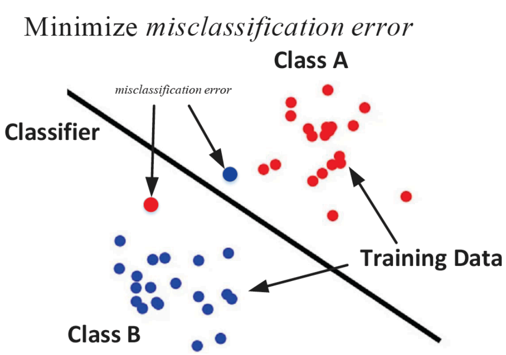
Illustration of knowledge transfer across MOPs

- How to define useful information in the transferred solutions and how such information are related to the multiple objectives across problems?

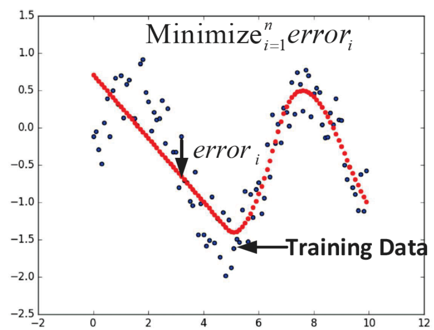
- J. Lin, H. Liu, K. C. Tan, and F. Gu, "An Effective Knowledge Transfer Approach for Multiobjective Multitasking Optimization", *IEEE Trans. on Cybernetics*, in press, 2020.
- C. Yang, J. Ding, Y. Jin, and T. Chai, "Offline Data-Driven Multiobjective Optimization: Knowledge Transfer Between Surrogates and Generation of Final Solutions," *IEEE Trans. on Evolutionary Computation*, 24(3), pp. 409-423, June 2020.

ETO for Machine Learning Applications

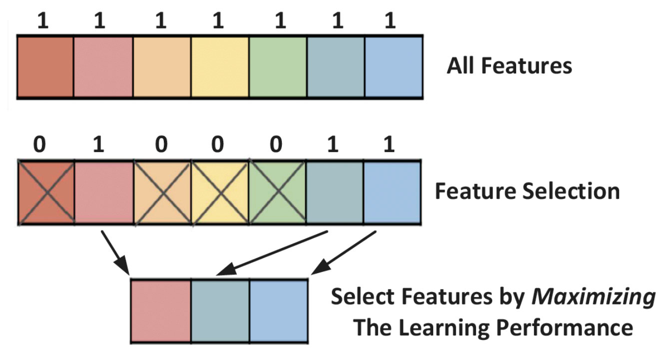
- ETO methods can be used in machine learning applications by leveraging on useful knowledge across learning problem domains, which can lead to more efficient performance of classification and feature selection etc.



(a) Classification problem



(b) Regression problem



(c) Feature selection

Examples of classification, regression, and feature selection problems

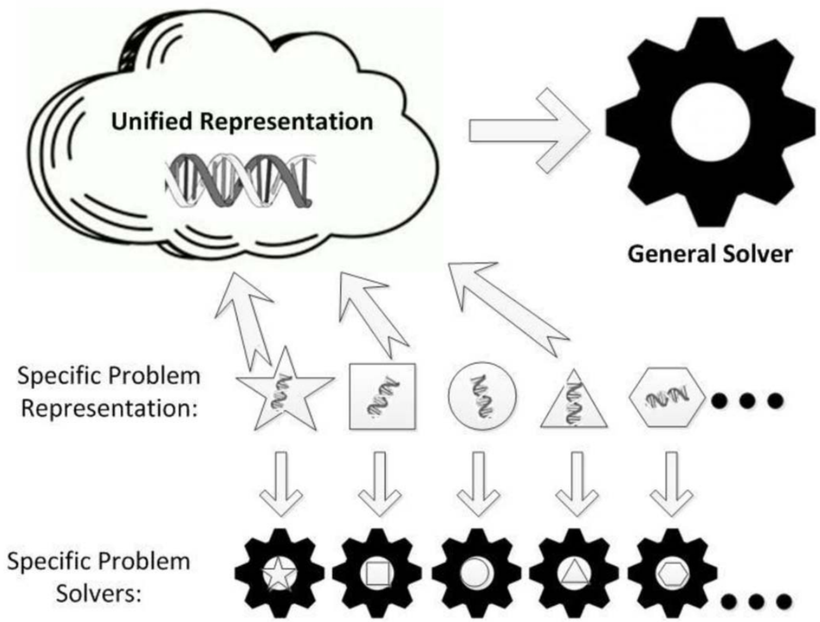
- How to design ETO algorithms capable of leveraging on big data technologies and advanced hardware (e.g., graphics processing units) to address today's ever-growing range and scale of demands in machine learning applications?

- T. Wei and J. Zhong, "A Preliminary Study of Knowledge Transfer in Multi-Classification Using Gene Expression Programming", *Frontiers in Neuroscience*, 13, pp.1396, 2020.
- B. Xue, M. Zhang, W. Browne, and X. Yao, "A Survey on Evolutionary Computation Approaches to Feature Selection", *IEEE Trans. on Evolutionary Computation*, 20(4), pp. 606-626, 2016.

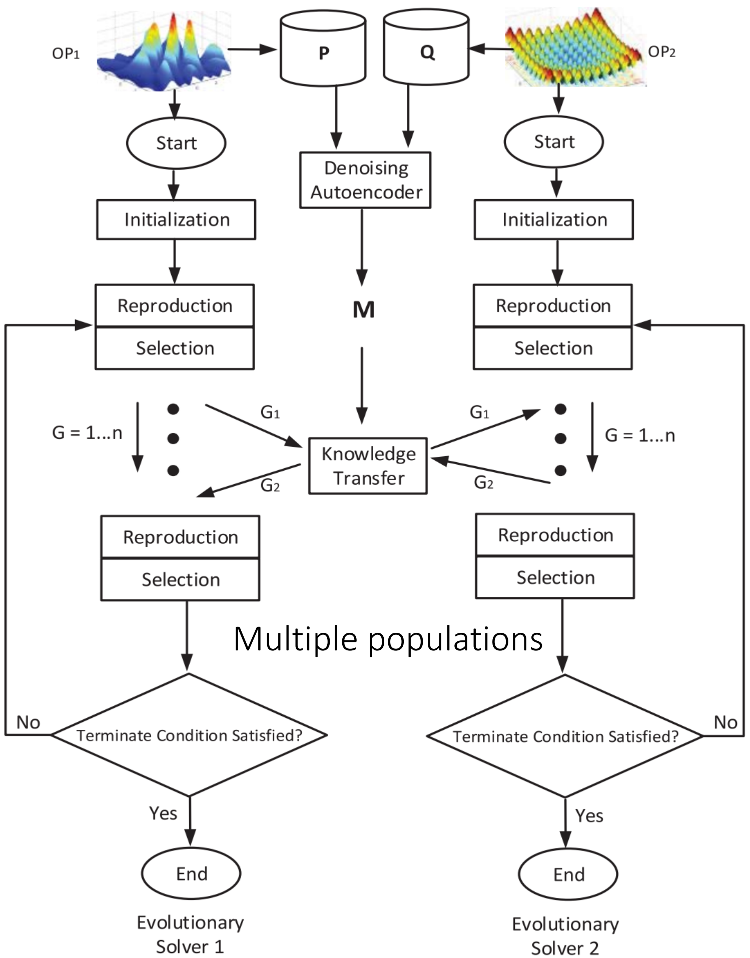
OUTLINE

- L. Zhou, L. Feng, **K. C. Tan**, J. Zhong, Z. Zhu, K. Liu, and C. Chen, "Towards Adaptive Knowledge Transfer in Multifactorial Evolutionary Computation", ***IEEE Trans. on Cybernetics***, in press, 2020.
- J. Lin, HL. Liu, **K. C. Tan**, and F. Gu, "An Effective Knowledge Transfer Approach for Multiobjective Multitasking Optimization", ***IEEE Trans. on Cybernetics***, in press, 2020.
- L. Feng, Y. Huang, L. Zhou, J. Zhong, A. Gupta, K. Tang, and **K. C. Tan**, "Explicit Evolutionary Multitasking for Combinatorial Optimization: A Case Study on Capacitated Vehicle Routing Problem", ***IEEE Trans. on Cybernetics***, in press, 2020.
- L. Feng, L. Zhou, A. Gupta, J. Zhong, Z. Zhu, **K. C. Tan**, and K. Qin, "Solving Generalized Vehicle Routing Problem With Occasional Drivers via Evolutionary Multitasking", ***IEEE Trans. on Cybernetics***, in press, 2019.
- L. Feng, L. Zhou, J. Zhong, A. Gupta, Y. Ong, **K. C. Tan**, and A. Qin, “Evolutionary Multitasking via Explicit Autoencoding”, ***IEEE Trans. on Cybernetics***, 49(9), pp. 3457-3470, 2018.
- A. Gupta, Y. Ong, L. Feng, and **K. C. Tan**, "Multiobjective Multifactorial Optimization in Evolutionary Multitasking", *IEEE Trans. on Cybernetics*, 47(7), pp. 1652-1665, 2016.

Implicit & Explicit Evolutionary Multitasking



(a) Knowledge transfer occurs implicitly via genetic crossover.



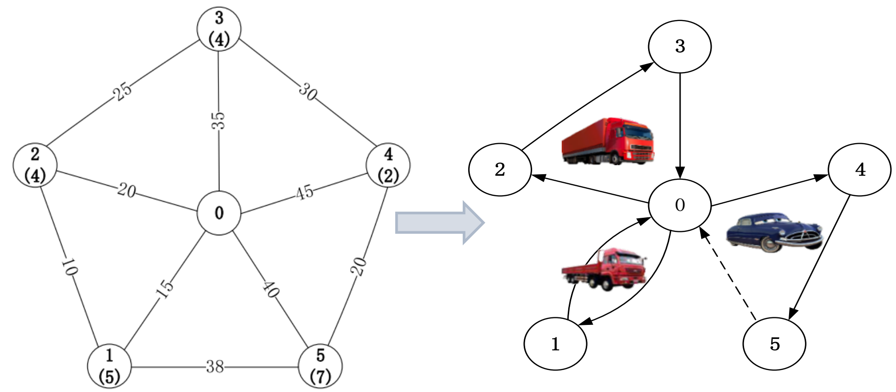
(b) Knowledge transfer occurs explicitly via additional transfer approach.

Implicit Knowledge Transfer

Regular drivers

Occasional drivers

Each type of vehicle has a distinct capacity, cost (fix cost + variable cost) and vehicle time window.



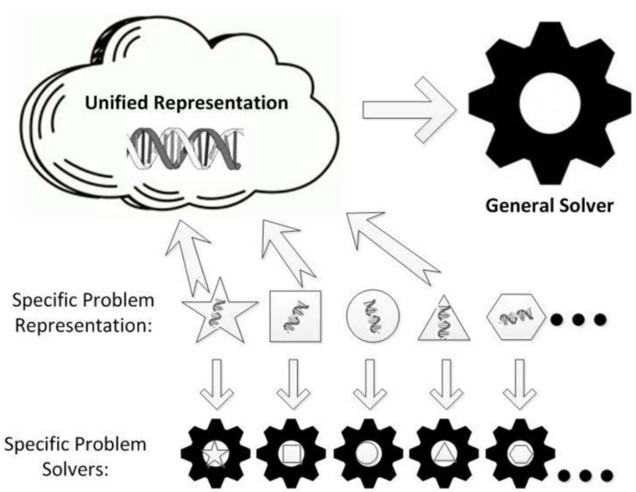
- Case study on Vehicle Routing Problems (VRPs) with heterogeneous capacity, time window and occasional driver (VRPHTO).
- As both regular drivers and occasional drivers are considered for providing services, VRPHTO contains more constraints than a capacitated VRP.
- The objective of VRPHTO: Minimize the total cost involved without violating any constraint.

- L. Feng, L. Zhou, A. Gupta, J. Zhong, Z. Zhu, **K. C. Tan**, and K. Qin, "Solving Generalized Vehicle Routing Problem With Occasional Drivers via Evolutionary Multitasking", ***IEEE Trans. on Cybernetics***, in press, 2019.

Implicit Knowledge Transfer

The Algorithm

- To solve VRPHTOs via evolutionary multi-tasking with a single solver, the following issues should be considered:
  - How to integrate multiple VRPHTOs with different properties into a unified search space?
  - How to evaluate a chromosome in the unified search space for a specific task?
  - How to determine the elitism of chromosomes whose performance vary among different VRPHTO tasks?



Algorithm 1: Pseudo Code of the proposed *EMA*.

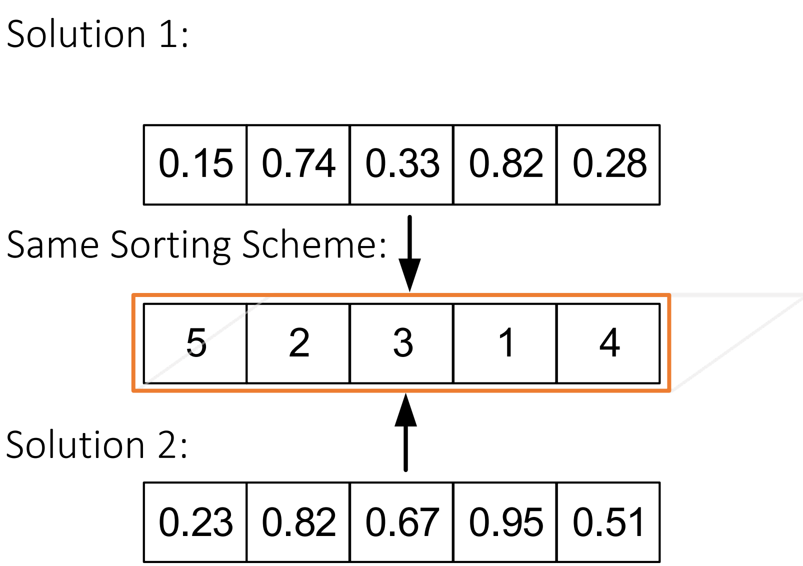
Initialize a population P of size N_p with a ***permutation-based unified representation***;
Evaluate each chromosome i on all the tasks, and obtain its *factorial cost* f_i, *factorial rank* r_i, *scalar fitness* φ_i and *skill factor* τ_i accordingly;
for *Restart:= 1* ***to*** N_{re} **do**
    while *(Number of task evaluations < TE)* **do**
        Select two parents p_a and p_b in P via binary tournament;
        Generate two offspring c_a and c_b, where **routing information exchange** is performed with a predefined probability.
        Conduct **split procedure** on c and evaluate c on a particular task.
        Add c_a and c_b to the offspring population C;
        if *the number of solutions in C, i.e.,* n_c*, is equal to* N_p **then**
            Concatenate P and C to form an intermediate population I;
            Update the *scalar fitness* φ_i and *skill factor* τ_i of every chromosome in I;
            Preserve the fittest N_p chromosomes in I that are selected by ***chromosome evaluation***, as the next generation in P;
    if *Restart* $\leq N_{re}$ **then**
        Restart by regenerating all the chromosomes in P, while preserving the N_{be} best ones.

Implicit Knowledge Transfer

Permutation-based Representation

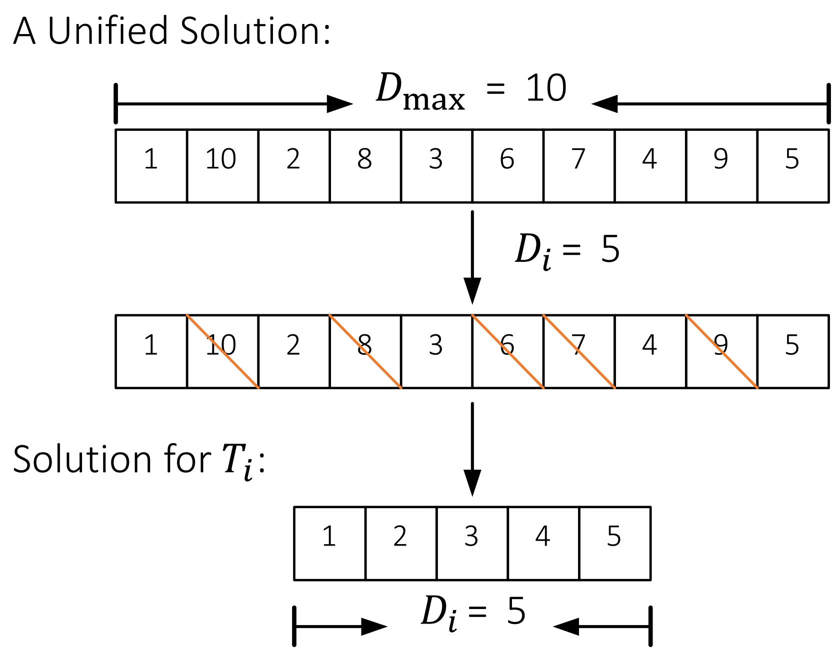
Random-key representation:

- Based on sorting scheme
- Cannot represent VRP solutions effectively



Permutation-based Representation:

- D_i : Dimension of the i^{th} task
- D_{max} : Dimension of the unified search space
- $D_{max} = \max\{D_i\}, i = 1, \ldots, K$



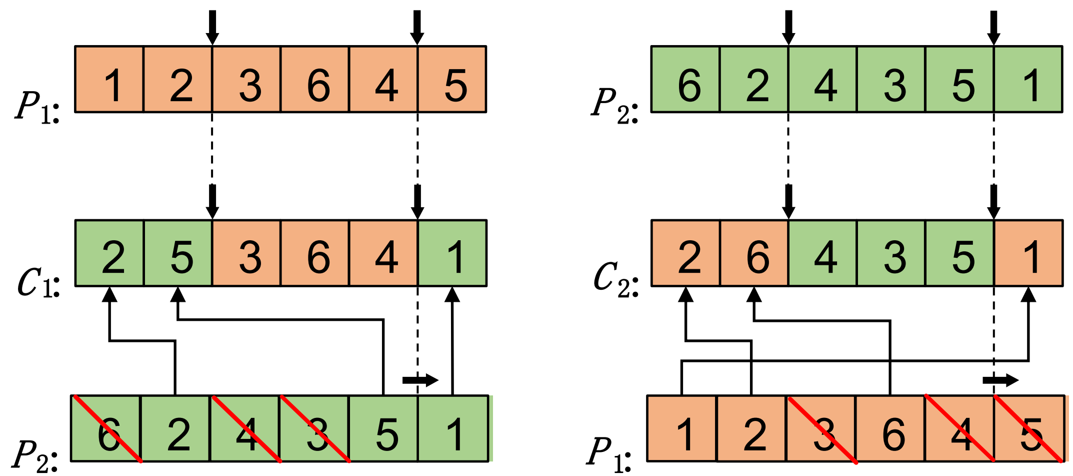
Implicit Knowledge Transfer

Routing information exchange

Routing Information Exchange

---

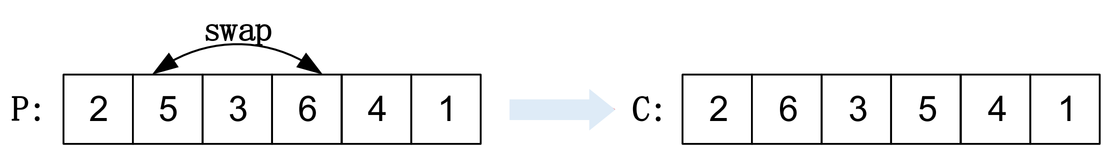
1. if $(\boldsymbol{\tau_a} == \boldsymbol{\tau_b})$ or $\boldsymbol{rand} < \boldsymbol{rmp}$ then
2. $\quad$ $\mathbf{Crossover}(\boldsymbol{p_a}, \boldsymbol{p_b}) \rightarrow$ Offspring $\boldsymbol{c_a}$ and $\boldsymbol{c_b}$;
3. $\quad$ else
4. $\quad$ $\mathbf{Mutate}(\boldsymbol{p_a}) \rightarrow$ Offspring $\boldsymbol{c_a}$;
5. $\quad$ $\mathbf{Mutate}(\boldsymbol{p_b}) \rightarrow$ Offspring $\boldsymbol{c_b}$;
6. endif



An example of order crossover (OX)

Offspring Generation

- Crossover operator: Order crossover
- Mutation operator: Swap mutation



An example of swap mutation (SW)

Implicit Knowledge Transfer

Empirical Study

Generation of VRPHTO Benchmarks

- Denote the n regular types as $RT_1, RT_2, \cdots, RT_n$ with the capacity arranged in an ascending order.
- Randomly select $\left\lfloor \frac{n}{2} \right\rfloor$ regular types to generate occasional types.
- The time window of occasional driver is randomly generated within the length of $\left[\frac{L-E}{\alpha}, \frac{L-E}{\beta}\right]$ that lies in the time window of the depot, i.e., [E, L].

R1A (12 Problems)						R2A (11 Problems)					
V-Type	Capacity	Fix Cost	Variable Cost	VTW		V-Type	Capacity	Fix Cost	Variable Cost	VTW	
				Ve	Vl					Ve	Vl
A	30	50	1	0	230	A	300	450	1	0	1000
B	50	80	1	0	230	B	400	700	1	0	1000
C	80	140	1	0	230	C	600	1200	1	0	1000
D	120	250	1	0	230	D	1000	2500	1	0	1000
E	200	500	1	0	230	A1	300	225	1.5	186	596
B1	50	65	1.5	80	191	B1	400	575	1.5	554	917
C1	80	110	1.5	29	113	/	/	/	/	/	/
Time window of the depot				[0, 230]		Time window of the depot				[0, 1000]	

Implicit Knowledge Transfer

Empirical Study

Generation of Multi-tasking VRPHTO Benchmarks

Multi-tasking Problems	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
VRPHTO Instances	R101A	R103A	R105A	R107A	R109A	R111A	C101A	C103A	C105A	C107A	RC101A	RC103A	RC105A	RC107A
	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	R102A	R104A	R106A	R108A	R110A	R112A	C102A	C104A	C106A	C108A	RC102A	RC104A	RC106A	RC108A
Multi-tasking Problems	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	
VRPHTO Instances	R201A	R203A	R205A	R207A	R209A	C201A	C203A	C205A	C207A	RC201A	RC203A	RC205A	RC207A	
	+	+	+	+	+	+	+	+	+	+	+	+	+	
	R202A	R204A	R206A	R208A	R210A	C202A	C204A	C206A	C208A	RC202A	RC204A	RC206A	RC208A	

- 27 multi-tasking VRPHTO problem sets are obtained by pairing the instances in order within the same VRPHTO category.

CityU
香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative For The World

Implicit Knowledge Transfer

Results and Analysis

Problem	HVRPTWOD	EMA			SEA		
		Ave.Cost	B.Cost	Std.Dev	Ave.Cost	B.Cost	Std.Dev
P1	R101A	**4313.16** ≈	4293.64	10.10	4316.88	4293.64	12.30
	R102A	**4100.83** ≈	**4081.19**	13.50	4104.31	4086.61	11.80
P4	R107A	**3856.19** ≈	**3835.71**	12.00	3858.50	3839.70	12.40
	R108A	**3736.52** ≈	**3714.18**	10.90	3741.90	3720.39	12.10
P8	C103A	5543.79 ≈	**5488.79**	47.90	**5539.57**	5490.09	37.80
	C104A	**5098.19** ≈	**5034.15**	36.90	5108.92	5039.83	29.70
P12	RC104A	**4553.24** ≈	4527.84	12.90	4560.20	**4525.66**	20.40
P16	C203A	**5210.46** ≈	5210.45	0.00	5210.47	5210.45	0.10
	C204A	**5204.95** ≈	5204.86	0.20	5205.28	5204.86	0.60
P20	R203A	**2949.74** ≈	2864.29	63.10	2966.87	**2861.57**	55.50
	R204A	**2505.06** ≈	**2412.65**	32.80	2519.87	2502.35	7.50
P23	R209A	**2801.59** +	**2615.27**	65.00	2833.46	2808.98	12.60
	R210A	2877.76 ≈	2843.57	13.50	**2872.50**	**2735.90**	45.10
P25	RC203A	**3307.27** ≈	**3270.96**	21.40	3311.12	3278.12	18.30
	RC204A	**3042.94** ≈	3030.65	10.70	3045.65	**3029.99**	15.20
P27	RC207A	**3067.49** ≈	**3050.27**	8.10	3069.79	3054.51	12.40
	RC208A	**2722.10** ≈	2716.96	2.90	2722.11	2716.96	2.20

EMA achieved superior or competitive performance against SEA on 37 out of a total of 54 VRPHTO instances in terms of the averaged cost.

Implicit Knowledge Transfer

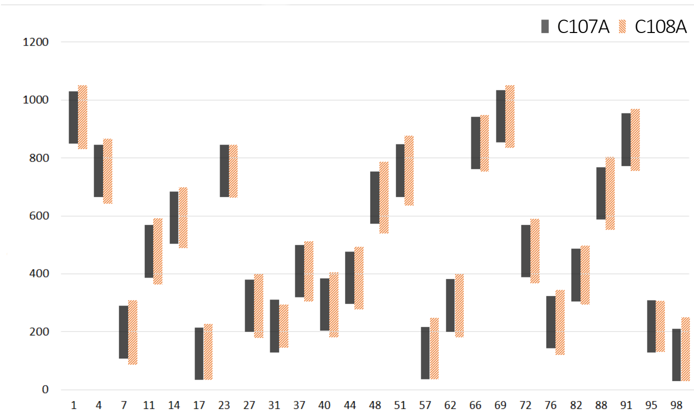
Results and Analysis

Instances	Speed up (Fitness) $SpeedUp = \frac{SEA^{i}_{Task\ Evalution}}{EMA^{i}_{Task\ Evalution}}$							
	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8
R112A	1.64 (3772.40)	1.86 (3754.48)	**2.09** (3736.56)	1.77 (3718.64)	1.72 (3700.72)	1.63 (3682.79)	1.48 (3664.87)	**2.00** (3646.95)
C108A	1.56 (6144.[illegible])	1.66 [illegible]	1.94 [illegible]	1.57 [illegible]	**2.13** [illegible]	**2.25** [illegible]	**2.46** ([illegible]30.94)	1.31 (5895.39)
RC107A	1.64 (4927.[illegible])	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible].38 ([illegible]14.92)	1.80 (4679.53)
R204A	**2.47** (2785.61)	1.06 (2752.39)	1.39 (2719.17)	1.00 (2685.95)	1.29 (2652.74)	1.14 (2619.52)	1.18 (2586.30)	0.88 (2553.09)
C202A	1.57 (5517.12)	1.09 (5479.11)	1.00 (5441.09)	1.42 (5403.08)	**3.08** (5365.06)	**3.31** (5327.05)	**3.91** (5289.03)	**2.13** (5251.02)
RC203A	1.48 (3505.92)	1.20 (3481.57)	1.53 (3457.22)	1.44 (3432.87)	1.16 (3408.52)	1.98 (3384.17)	1.93 (3359.82)	**2.87** (3335.47)

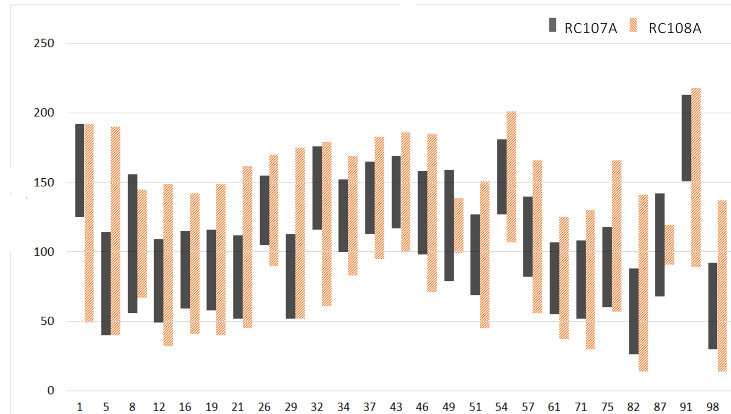
EMA achieved a faster convergence for most stages of the evolution in different instances.

Implicit Knowledge Transfer

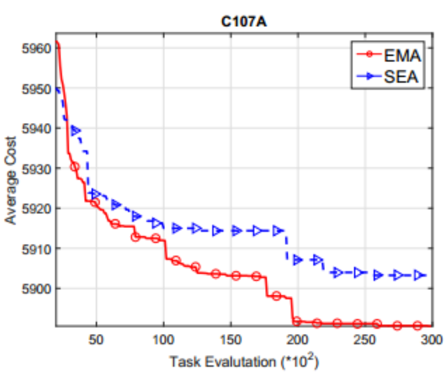
Results and Analysis

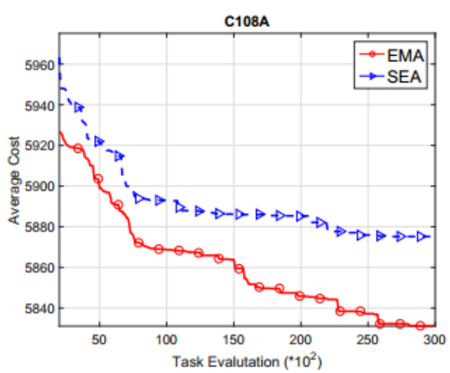


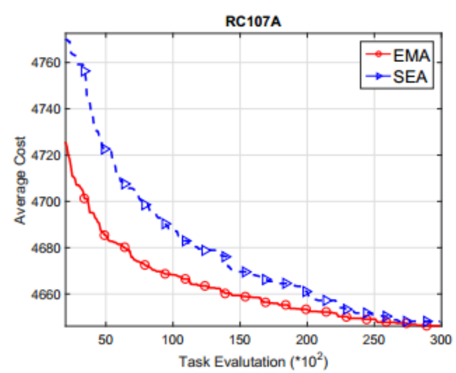
(a)

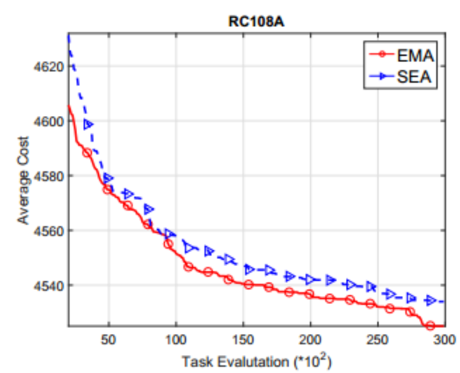


(b)









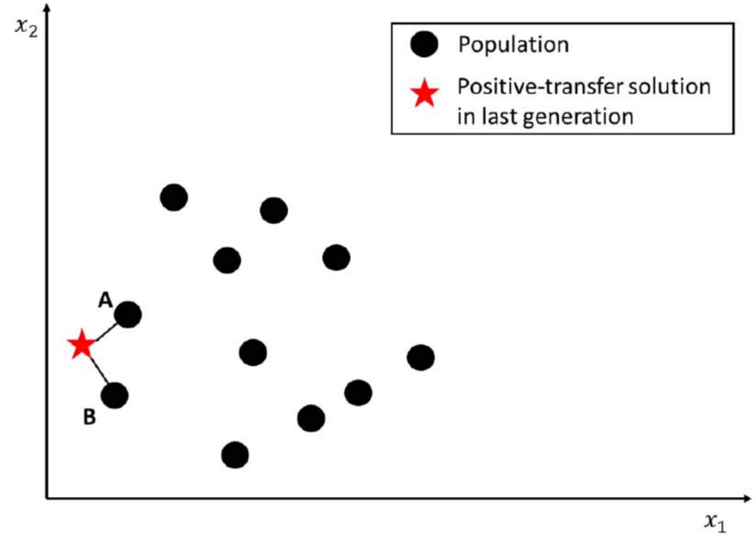
- Better optimization performance can be obtained by EMT as compared to SEA if similar problems are paired to be solved simultaneously.

Effective Knowledge Transfer

EMT/ET

- Each solution in EMT possesses the same probability of undergoing knowledge transfer. Poor solutions may lead to negative transfer in ETO.
- How to identify useful solutions for positive knowledge transfer?

  - Black dots denote the population at the current generation.
  - Red star represents a solution which achieved positive transfer in the last generation.
  - The neighbors of the red star (denoted as A and B) will be selected as the transferred solutions at the current generation.



The transferred solutions across tasks are selected based on those solutions achieving a positive transfer in the last generation.

- J. Lin, HL. Liu, **K. C. Tan**, and F. Gu, "An Effective Knowledge Transfer Approach for Multiobjective Multitasking Optimization", ***IEEE Trans. on Cybernetics***, in press, 2020.

Effective Knowledge Transfer

Results and Analysis

- Baseline solvers: NSGAII, SPEA2, two recent multi-tasking algorithms (EMEA and MFEA).
- Baseline strategies:
  - Selecting some non-dominated solutions in each task as the transferred solutions.
  - Randomly selecting the transferred solutions in each task.

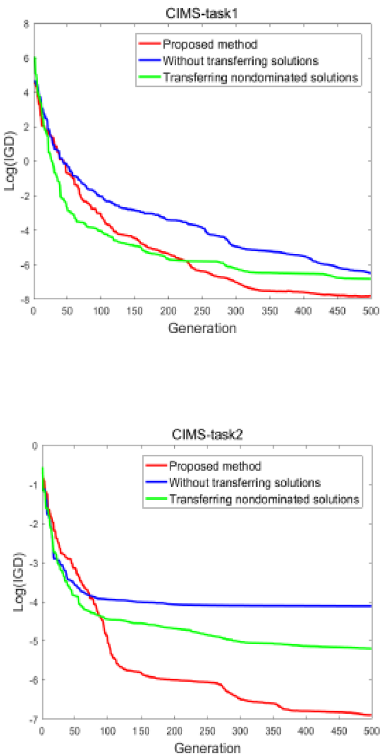
Problem	Task	EMT/ET	EMEA	MFEA	SPEA2	NSGA-II
CIHS	T1	1.90E-04(2.11E-04)	**1.67E-04(1.85E-04)**	1.60E-03(3.50E-03)	3.59E-03(6.18E-03)	-
	T2	1.77E-04(2.04E-04)	**1.80E-04(1.94E-04)**	5.90E-03(9.10E-03)	-	3.90E-03(5.20E-03)
CIMS	T1	**1.73E-04(1.94E-04)**	1.73E-04(4.41E-04)	1.78E-04(4.72E-04)	9.64E-04(4.63E-02)	-
	T2	**1.70E-04(3.16E-04)**	9.97E-04(0.0148)	2.04E-04(6.02E-04)	-	9.00E-04(1.08E-02)
CILS	T1	**1.70E-04(1.88E-04)**	1.73E-04(1.92E-04)	1.89E-04(1.93E-04)	1.17E-02(7.35E-02)	-
	T2	1.74E-04(1.92E-04)	**1.72E-04(1.90E-04)**	1.84E-04(1.96E-04)	-	8.96E-04(1.03E-03)
PIHS	T1	**1.70E-04(1.88E-04)**	7.12E-04(2.60E-03)	2.59E-02(5.06E-02)	2.58E-03(3.96E-03)	-
	T2	**1.67E-04(7.21E-04)**	3.12E-02(1.57E-01)	5.84E-01(1.12)	-	1.05E-01(1.94E-01)
PIMS	T1	**1.74E-04(2.81E-04)**	8.82E-04(2.80E-03)	2.40E-03(4.70E-03)	3.78E-02(7.15E-02)	-
	T2	**1.71E-04(2.06E-04)**	6.8369(7.527)	7(10.461)	-	2.14E+00(4.37E+00)
PILS	T1	**1.71E-04(1.89E-04)**	2.32E-04(3.44E-04)	9.91E-04(1.60E-03)	7.80E-04(1.98E-03)	-
	T2	**3.11E-04(1.34E-03)**	3.53E-02(6.01E-02)	6.27E-02(8.21E-02)	-	2.00E-01(2.01E-01)
NIHS	T1	**1.47E+00(1.48E+00)**	1.49E+00(1.49E+00)	2.03E+00(2.83E+00)	2.81E+00(9.19E+00)	-
	T2	1.77E-04(1.88E-04)	**1.72E-04(1.86E-04)**	3.20E-03(8.30E-03)	-	1.60E-03(2.60E-03)
NIMS	T1	1.48E-01(1.55E-01)	1.60E-01(1.64E-01)	**9.90E-02(1.53E-01)**	7.06E-02(6.22E-01)	-
	T2	**1.41E-04(1.50E-04)**	1.42E-04(1.55E-04)	3.31E-04(1.80E-02)	-	0.300E-03(4.29E-02)
NILS	T1	6.43E-04(7.53E-04)	6.83E-04(7.56E-04)	9.27E-04(1.6E-03)	**5.26E-04(5.60E-04)**	-
	T2	**4.51E-04(1.11E-02)**	6.42E-01(6.42E-01)	6.43E-01(6.44E-01)	-	2.02E-01(2.03E-01)

The Smallest and average values (shown in the brackets) of IGD obtained by EMT/ET, EMEA, MFEA, SPEA2, NSGA-II on the nine MTO benchmarks.

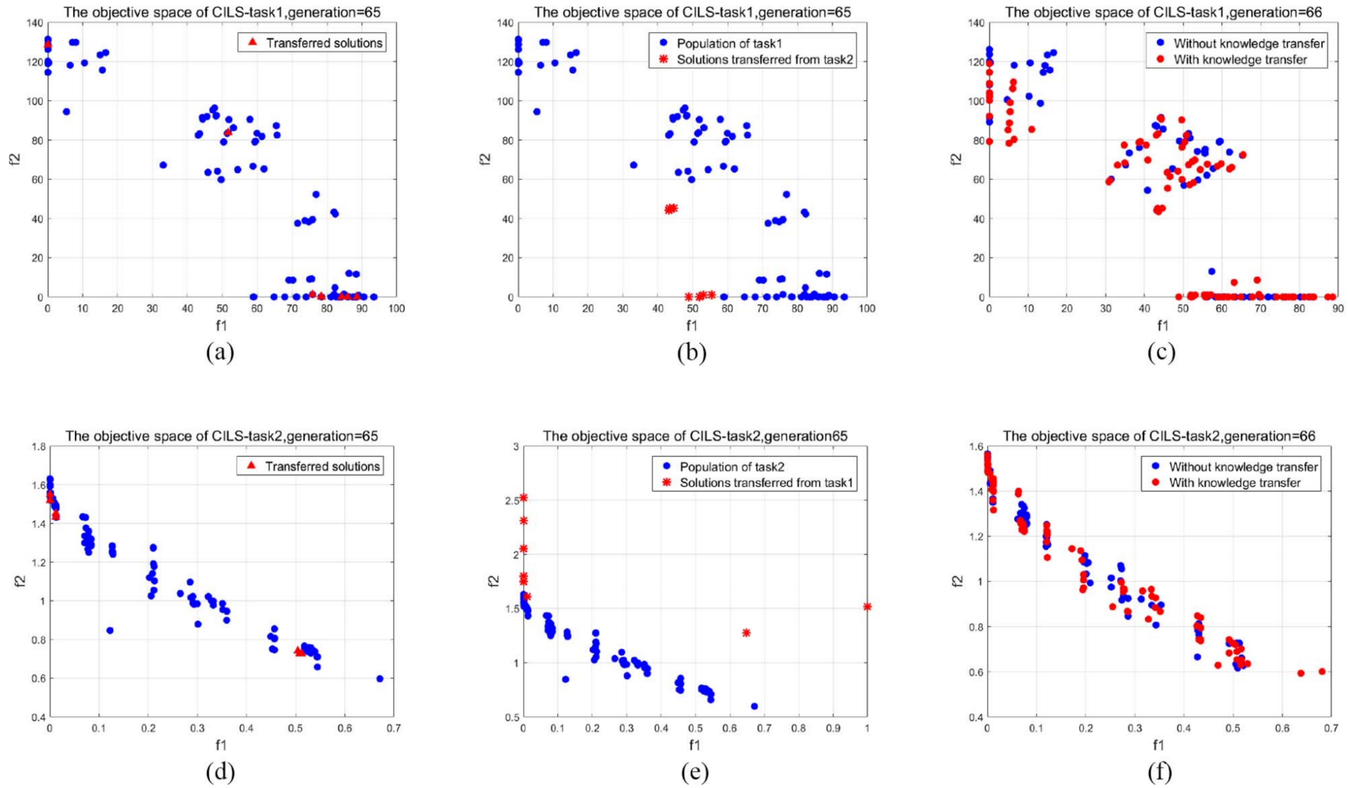
- Y. Yuan, Y. Ong, L. Feng, A. Qin, A. Gupta., B. Da, Q. Zhang, **K. C. Tan**, Y. Jin, and H. Ishibuchi, "Evolutionary Multitasking for Multiobjective Continuous Optimization: Benchmark Problems, Performance Metrics and Baseline Results", *Technical Report*, 2016.

Effective Knowledge Transfer

Results and Analysis



(a) Convergence trace

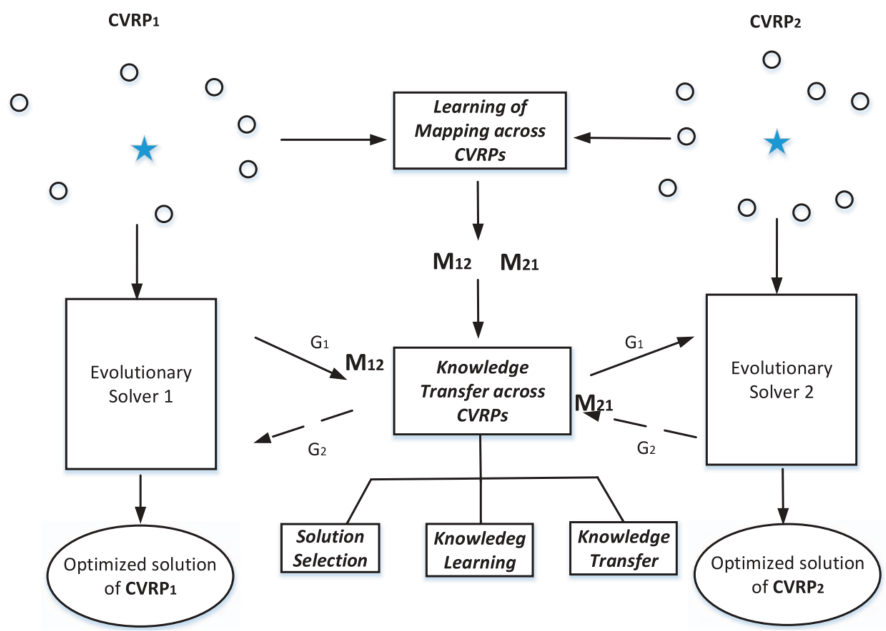


(b) Effects of transferred solutions for CILS

Explicit Knowledge Transfer

Explicit Multi-tasking for CVRP

- **Learning of Mapping** across Capacitated Vehicle Routing Problems (CVRPs)
  - Represent two CVRPs by two matrices, i.e., P_s and P_t ($d \times n_s$ and $d \times n_t$ matrix; where d is the number of features for representing the location of a customer, and n_s and n_t is the number of customers in P_s and P_t, respectively).



Explicit Knowledge Transfer for solving CVRPs

- The problem of finding customers from P_s to represent customers in P_t can be formulated as the learning of an $n_s \times n_t$ transformation matrix M, so that $P_s \times M = P_t$,

$$\min_{M} \|P_s \times M - P_t\|_F + \|D \odot M\|_{l_1},$$

where the first term is the reconstruction error and the second term is the weighted l_1 norm-based regularization.

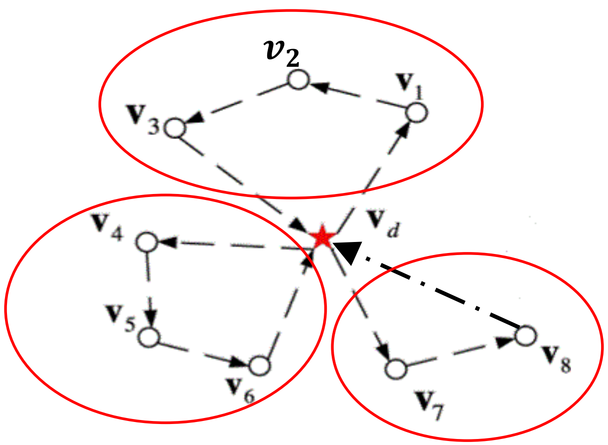
- L. Feng, Y. Huang, L. Zhou, J. Zhong, A. Gupta, K. Tang, and **K. C. Tan**, "Explicit Evolutionary Multitasking for Combinatorial Optimization: A Case Study on Capacitated Vehicle Routing Problem", *IEEE Trans. on Cybernetics*, in press, 2020.

Explicit Knowledge Transfer

Explicit Multi-tasking for CVRP

- **Solution Selection:** Select the best Q number of optimized solutions from the source CVRP domain based on the objective value to be transferred to the target CVRP domain.

- **Knowledge Learning**: To capture the useful information embedded in each of the selected solutions, which can be transferred across different CVRPs.
  - Construct an $n_s \times n_s$ distance matrix DM.
  - The new estimated customer representations P_s^{new} can be obtained via multi-dimensional scaling with DM.



Optimized CVRP solution:

$\{0, \boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, 0, \boldsymbol{v_4}, \boldsymbol{v_5}, \boldsymbol{v_6}, 0, \boldsymbol{v_7}, \boldsymbol{v_8}, 0\}$

$$\begin{pmatrix} 0 & \alpha & 2\alpha & \beta & \beta & \beta & \beta & \beta \\ \alpha & 0 & \alpha & \beta & \beta & \beta & \beta & \beta \\ 2\alpha & \alpha & 0 & \beta & \beta & \beta & \beta & \beta \\ \beta & \beta & \beta & 0 & \alpha & 2\alpha & \beta & \beta \\ \beta & \beta & \beta & \alpha & 0 & \alpha & \beta & \beta \\ \beta & \beta & \beta & 2\alpha & \alpha & 0 & \beta & \beta \\ \beta & \beta & \beta & \beta & \beta & \beta & 0 & \alpha \\ \beta & \beta & \beta & \beta & \beta & \beta & \alpha & 0 \end{pmatrix}$$

The rule for setting α and β ($\alpha << \beta$) is that the vehicle assignment and service order in the selected solution can be accurately obtained when applying clustering and pair-wise distance sorting with DM

Explicit Knowledge Transfer

Explicit Multi-tasking for CVRP

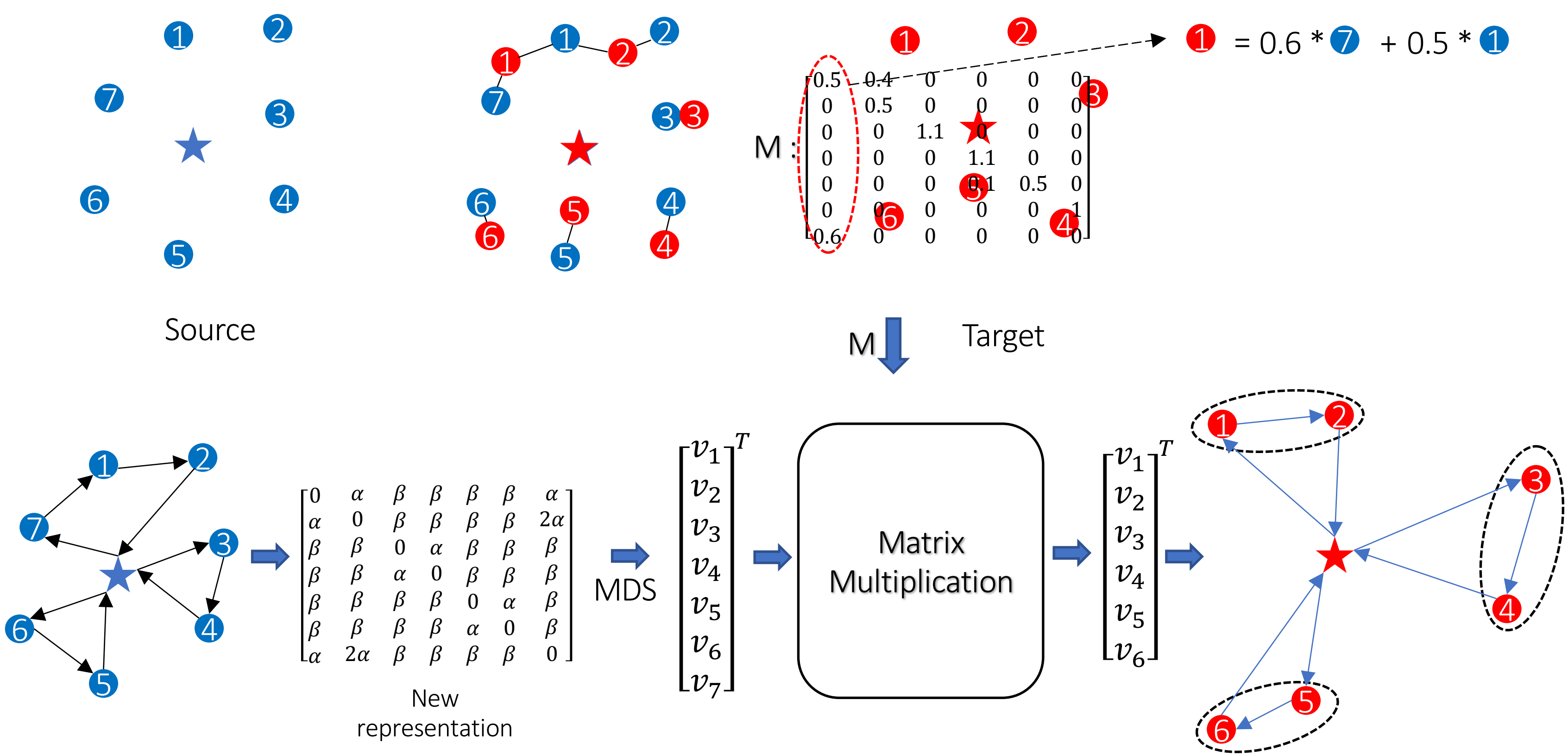
- Knowledge Transfer:
  - With the learned sparse customer mapping of M_{12} and M_{21} across CVRP domains and the new CVRP customer representation, the knowledge transfer across CVRPs can be performed by simple operation of matrix multiplication.
  - The approximated customers of P_t is obtained via $P_t^{new} = P_s^{new} \times M_{12}$. To obtain the transferred CVRP solution for P_t, *K*-means clustering with random initialization is conducted on P_t^{new} to derive the customer assignments of vehicles.

for $i \leq Q$ **do**
- **Set** $\mathbf{s}_s = \mathbf{s}_i$, and **Estimate** $\mathbf{P}_s^{\text{new}}$ as discussed in Section III-B2;
- **Obtain** $\boxed{\mathbf{P}_t^{\text{new}} \text{ via } \mathbf{P}_s^{\text{new}} \times \mathbf{M}_{12};}$
- **Perform** K-means and pair-wise distance sorting with $\mathbf{P}_t^{\text{new}}$ to generate CVRP solution for $\mathbf{p}_t$;
- **Insert** the generated solution into the population to undergo natural selection;
- $i = i + 1$;

for $i \leq Q$ **do**
- **Set** $\mathbf{s}_s = \mathbf{s}_i$, and **Estimate** $\mathbf{P}_s^{\text{new}}$ as discussed in Section III-B2;
- **Obtain** $\boxed{\mathbf{P}_t^{\text{new}} \text{ via } \mathbf{P}_s^{\text{new}} \times \mathbf{M}_{21};}$
- **Perform** K-means and pair-wise distance sorting with $\mathbf{P}_t^{\text{new}}$ to generate CVRP solution for $\mathbf{p}_t$;
- **Insert** the generated solution into the population to undergo natural selection;
- $i = i + 1$;

Explicit Knowledge Transfer

Illustrating Example



Explicit Knowledge Transfer

Empirical Study

- Comparison Algorithms
  - Two memetic algorithms (with different local search) as single task VRP solvers, i.e., *EA1* and *EA2*
  - Single task solver with random solution injection, i.e., *EA1+R* and *EA2+R*
  - Proposed explicit evolutionary multi-tasking algorithm, i.e., *EEMTA*
  - Existing evolutionary multi-tasking algorithm, i.e., *PMFEA*
- Benchmarks

Instance	Customer Number	Vehicle Capacity	Vehicle Number
A-n54-k7	54	100	7
A-n62-k8	62	100	8
A-n80-k10	80	100	10
B-n50-k7	50	100	7
B-n64-k9	64	100	9
B-n78-k10	78	100	10
P-n50-k8	50	120	8
P-n60-k10	60	120	10
P-n76-k5	76	280	5

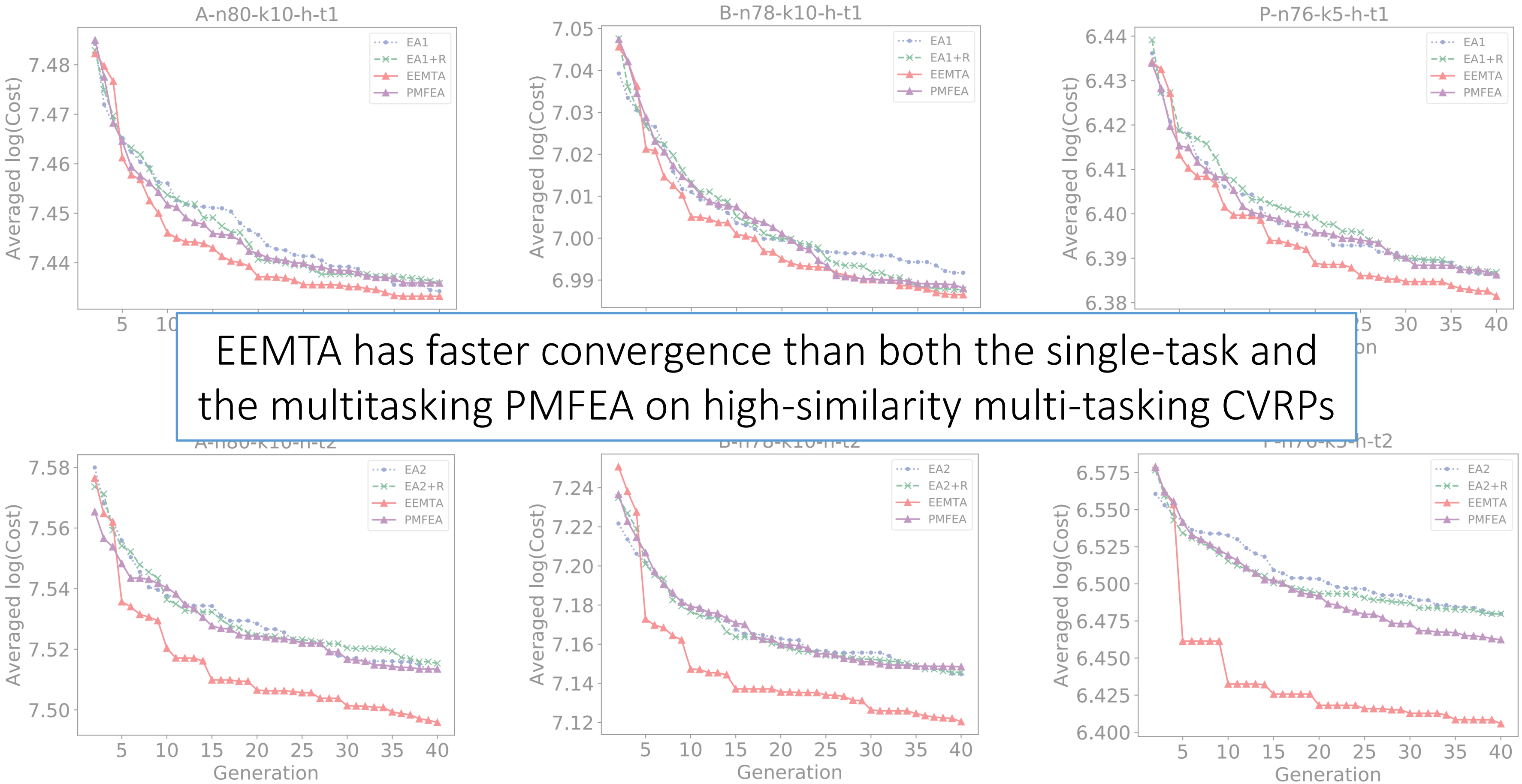
High-similarity, medium-similarity, and low-similarity multi-tasking CVRP pairs are constructed by randomly and independently deleting 10%, 30%, and 50% customers from the CVRP instances, respectively.

- http://neo.lcc.uma.es/vrp/known-best-results/

- Configurations
  - Parameters for the proposed EEMTA:
    - α and β in *DM*: $\alpha = 10$ and $\beta = 1000$
    - Number of solutions for transfer: $Q = 5$
    - Gen. interval for knowledge transfer: $G = 5$
  - Population size:
    - EA1, EA2, EA1+R, EA2+R, and EEMTA: 50
    - PMFEA: 100
  - Maximum generations: 100
  - Independent runs: 20
  - Local search settings:
    - Local search in EA1 and EA1+R: Replace, single-insertion, and two-swap.
    - Local search in EA2 and EA2+R: Replace.

Explicit Knowledge Transfer

Results and Analysis

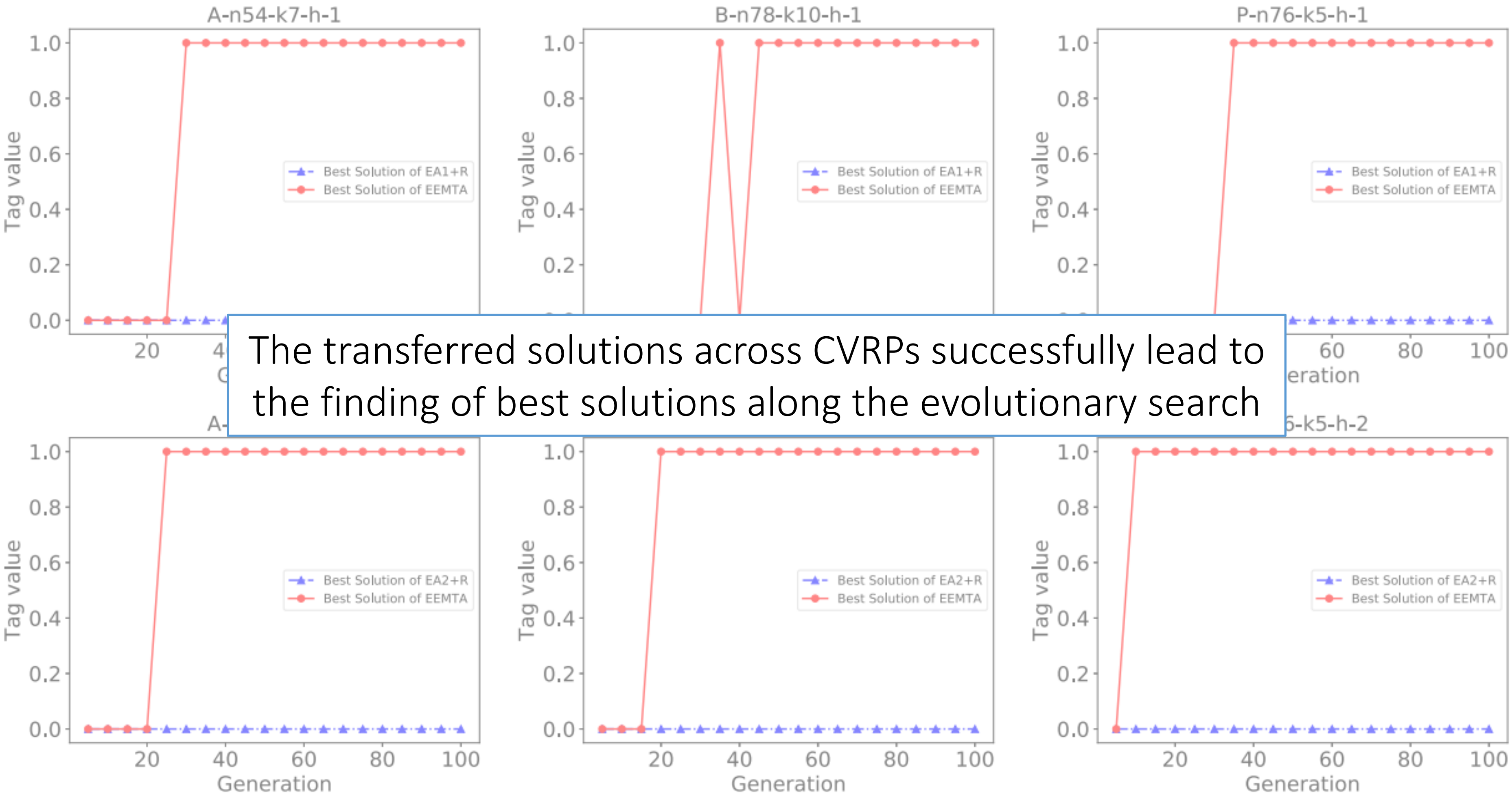


EEMTA has faster convergence than both the single-task and the multitasking PMFEA on high-similarity multi-tasking CVRPs

Convergence trace of EEMTA versus PMFEA and single-task EAs on representative high-similarity multitasking CVRPs.

Explicit Knowledge Transfer

Results and Analysis



The transferred solutions across CVRPs successfully lead to the finding of best solutions along the evolutionary search

Tracking of best solutions along the evolutionary search. If the best solution is generated from the transferred solutions, it will be given tag value 1, otherwise 0.

OUTLINE

01 ▶ Introduction

02 ▶ Evolutionary Transfer Multi-task Optimization

03 ▶ Evolutionary Transfer Dynamic Optimization

04 ▶ Evolutionary Transfer Optimization Application

05 ▶ Future Research Topics

- M. Jiang, Z. Wang, L. Qiu, S. Guo, X. Gao, and **K. C. Tan**, "A Fast Dynamic Evolutionary Multiobjective Algorithm via Manifold Transfer Learning", ***IEEE Trans. on Cybernetics***, in press, 2020.
- L. Feng, W. Zhou, W. Liu, Y. Ong, and **K. C. Tan**, “Solving Dynamic Multi-objective Problem via Autoencoding Evolutionary Search”, ***IEEE Trans. on Cybernetics***, submitted, 2020.
- M. Jiang, Z. Wang, S. Guo, X. Gao, and **K. C. Tan**, “Individual-based Transfer Learning for Dynamic Multi-objective Optimization”, ***IEEE Trans. on Cybernetics***, submitted.
- M. Jiang, G. Chi, G. Pan, S. Guo, and **K. C. Tan**, “Evolutionary Gait Transfer of Multi-Legged Robots in Complex Terrains”, ***IEEE Trans. on Robotics***, submitted.
- W. Zhou, L. Feng, **K. C. Tan**, M. Jiang, and Y. Liu, “Evolutionary Search with Multi-View Prediction for Dynamic Multi-objective Optimization”, ***IEEE Trans. on Cybernetics***, submitted.

Evolutionary Transfer Dynamic Optimization

- Dynamic multi-objective optimization problems (DMOPs):

$$\text{Minimize } F(\boldsymbol{x}, t) = [f_1(\boldsymbol{x}, t), \ldots, f_m(\boldsymbol{x}, t)]^{\mathrm{T}}, \ \text{subject to } \boldsymbol{x} \in X$$

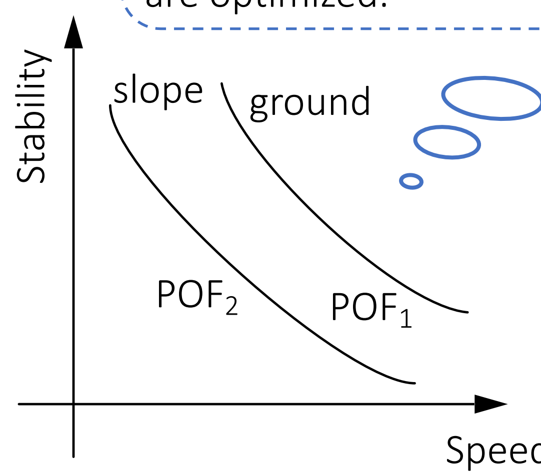
- t represents time index:

$$t = \frac{1}{n_t}\left\lfloor \frac{\tau}{\tau_t} \right\rfloor$$

where τ is the generation counter, τ_T is the number of generations for which t remains the same, and n_t is the number of distinct steps in t. τ_T determines the frequency of change and n_t determines the severity of change in a problem. A smaller value of n_t means larger change, whereas a smaller value of τ_T means more frequent occurrence of changes.

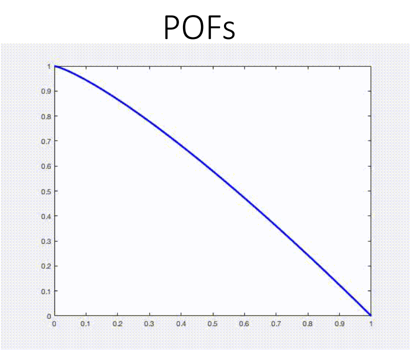
Robot gait is optimized based on the changing terrain so that objective functions such as stability and speed are optimized.

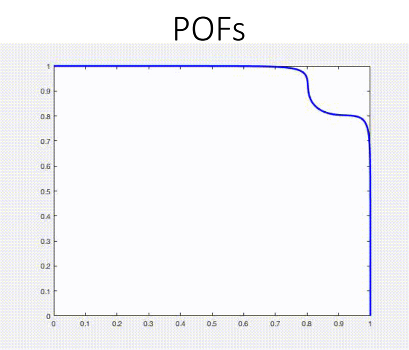
Design example of robot gait



Evolutionary Transfer Dynamic Optimization

- Evolutionary algorithms (EAs) have been widely applied to solve dynamic multi-objective optimization problems.
- Prediction-based EA methods estimate some solutions based on change patterns observed in past search experience to guide the search in subsequent environments.
- In DMOPs, the changing POFs may lead to different distributions of the training and predicted samples. The POFs may be Non-IID within two consecutive environments.
- Transfer learning (TF) can help to address the Non-IID issue in DMOPs by exploiting knowledge in past environment to generate a good initial population for Prediction-based EAs.





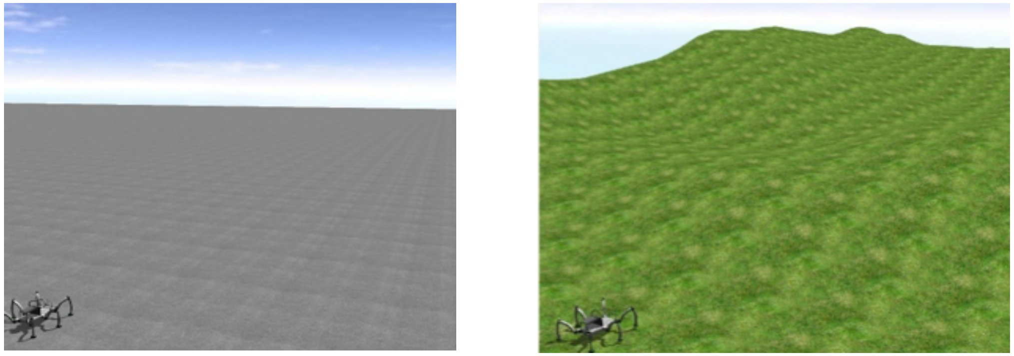
Non Independent and Identical Distributed (Non-IID) problem: Training samples and test samples subject to different distributions.

- S. Jiang, S. Yang, X. Yao, **K. C. Tan**, M. Kaiser, and N. Krasnogor, “Benchmark Problems for CEC’2018 Competition on Dynamic Multiobjective Optimisation”, *IEEE CEC’2018 Competition*, 2018.

Individual-based Transfer Learning

- Individual-based Transfer Learning for Dynamic Multi-objective Optimization (IT-DMOEA): Exploring the current environment to obtain a set of good solutions for knowledge transfer.
  - **Pre-search Strategy:** Obtain a high-quality population as training samples to guide the search direction to reduce effect of negative transfer.
  - **Individual Transfer:** Reuse information from past environment and the guided population to produce a good initial population.

Example of negative transfer

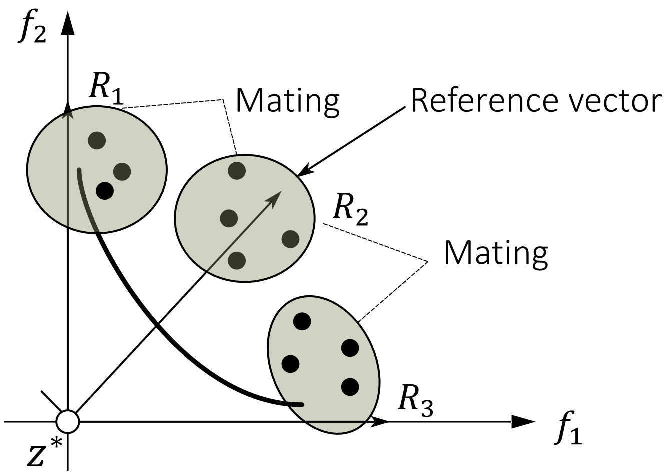


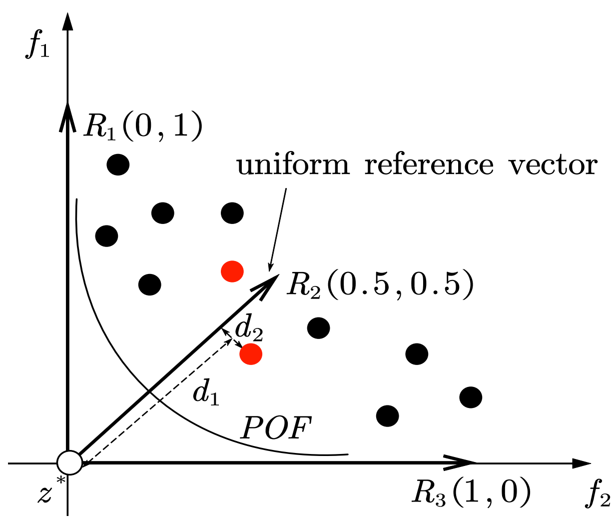
(a) Environment 1 (b) Environment 2

- M. Jiang, Z. Wang, S. Guo, X. Gao, and **K. C. Tan**, "Individual-based Transfer Learning for Dynamic Multi-objective Optimization", ***IEEE Trans. on Cybernetics***, submitted.

IT-DMOEA: Pre-Search Strategy

- Pre-search strategy obtains a high-quality population (guided population) as training samples to reduce the possibility of negative transfer.
  - **Reference vectors**: A set of uniform reference vectors R_i are generated. For each reference vector, two individuals (in red) with the minimum PBI values are selected.
  - **Crossover**: Individuals on neighboring reference vectors are mated to form a population pool.
  - **Mutation**: Individuals are mutated with Gaussian variation and added to the population pool.
  - For each reference vector, two individuals are selected from the population pool to generate a guided population.





- Penalty boundary intersection (PBI) is used for the selection of two individuals (for each reference vector) in order to maintain a good balance of convergence and diversity in the population.

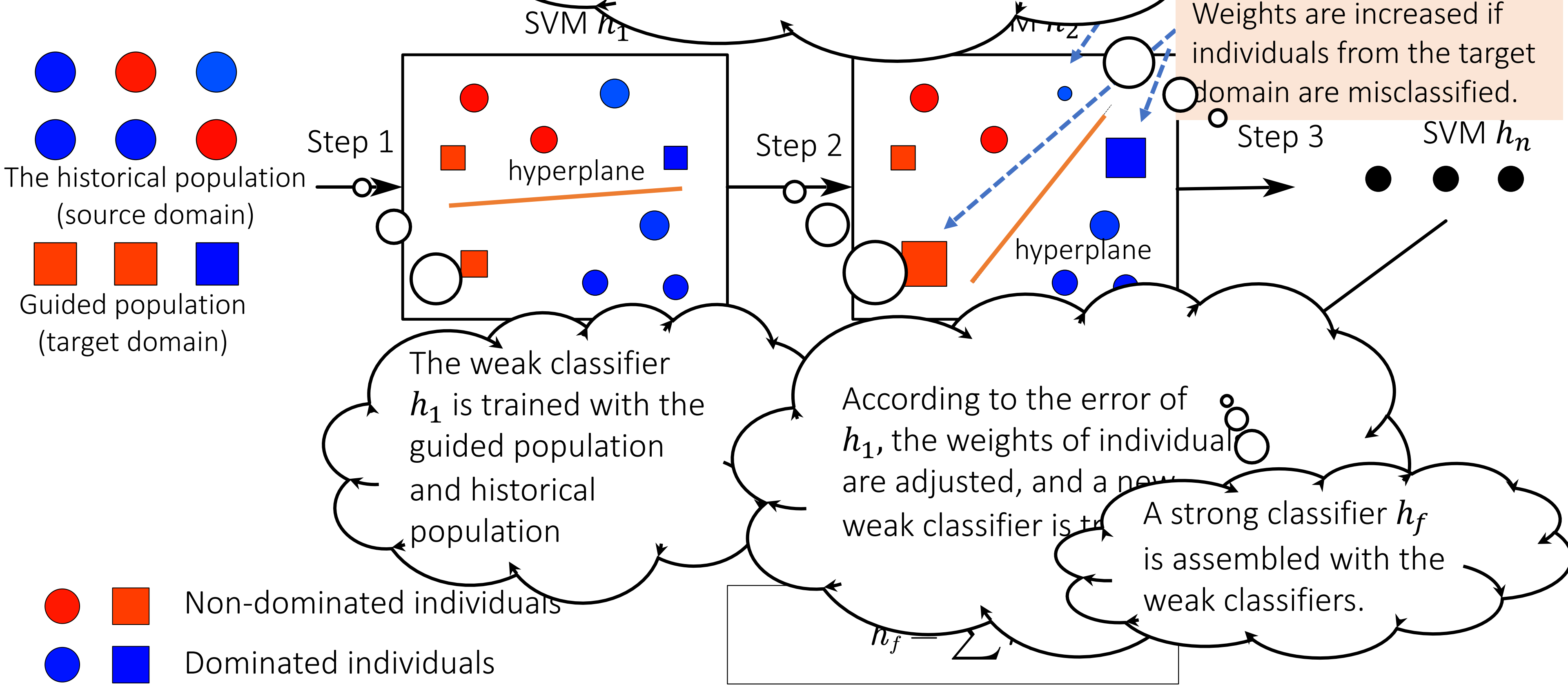
IT-DMOEA: Individual Transfer

- The idea of individual transfer ... individual is a strong indivi...
  - Sample-based classifier.
  - Ensemble of weak classifi...

Adjusting weights repeatedly can gradually make weak classifiers more accurate in classifying solutions of the target domain.

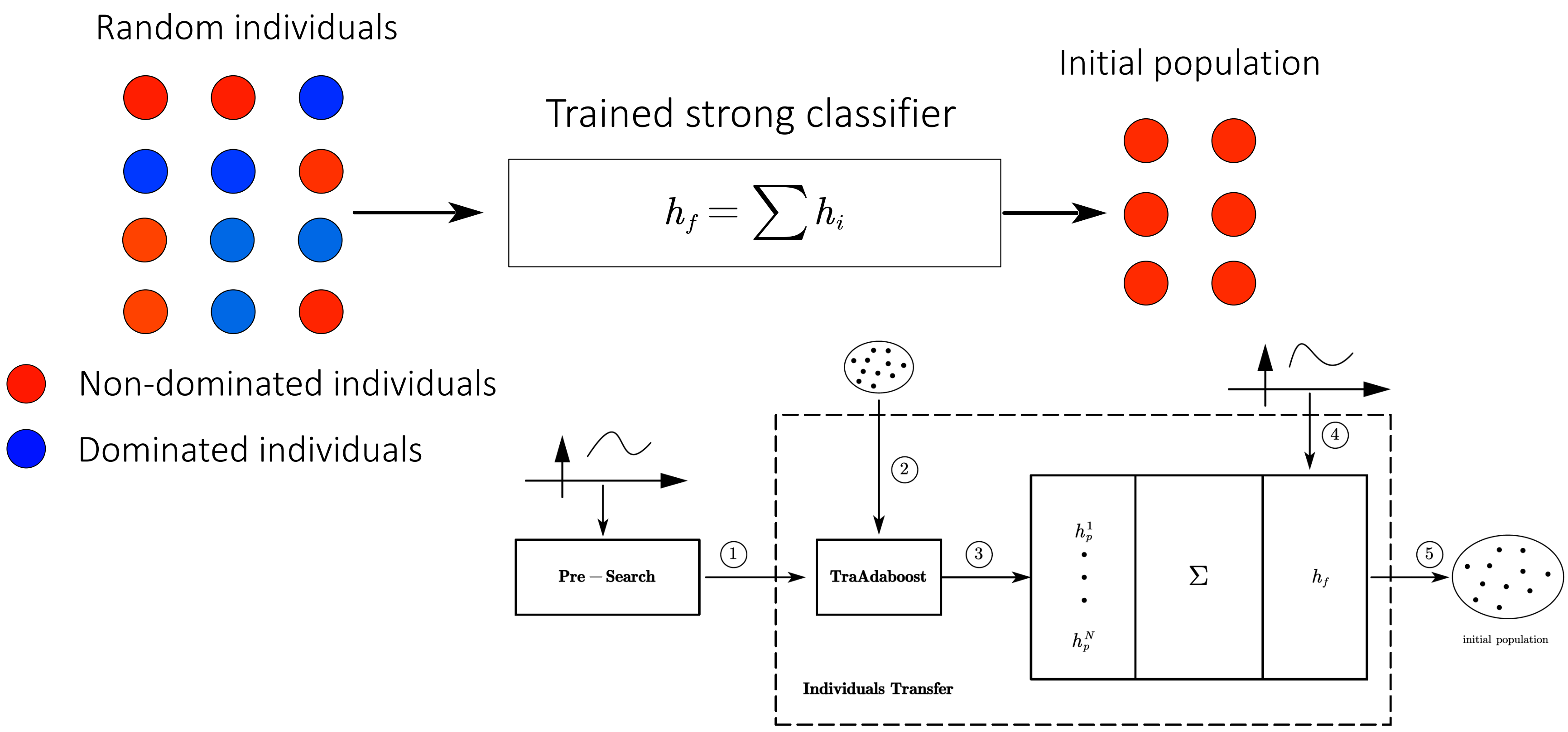
...ts are reduced if ...uals from the source ...n are misclassified.

Weights are increased if individuals from the target domain are misclassified.



IT-DMOEA: Forming Initial Population

- The strong classifier h_f then identifies individuals that are randomly generated in the current environment as strong or weak individuals. Those Individuals that have been identified as non-dominated will be selected to form an initial population to drive the evolution towards the newly changed Pareto front.



IT-DMOEA: Experimental Study

MIGD

MEAN AND STANDARD DEVIATION VALUES OF MIGD METRIC OBTAINED BY COMPARED ALGORITHMS FOR DIFFERENT DYNAMIC TEST SETTINGS

Problems	τ_t, n_t	PPS	MDP	MOEA/D-SVR	Tr-DMOEA	MOEA/D-KF	SGEA	T-DMOEA
DF11	10,10	0.0711±5.68E-3(+)	0.0773±6.79E-4(+)	0.0469±5.23E-3(=)	0.3331±2.19E-2(+)	0.0483±3.58E-3(+)	0.0886±1.08E-2(+)	**0.0461±4.68E-3**
	10,5	0.0733±1.90E-3(+)	0.0787±7.61E-4(+)	**0.0451±3.27E-3(-)**	0.3797±5.13E-2(+)	0.0511±2.94E-3(=)	0.0977±1.45E-2(+)	0.0506±3.09E-3
	5,10	0.0749±9.11E-3(+)	0.0849±1.01E-3(+)	**0.0453±3.66E-3(-)**	0.4178±2.94E-2(+)	0.0604±2.91E-2(+)	0.1192±1.37E-2(+)	0.0495±3.58E-3
DF12	10,10	0.8513±8.37E-2(-)	0.6769±1.23E-2(+)	0.5851±5.26E-2(-)	1.1900±1.64E-5(=)	0.6468±5.90E-2(-)	**0.2480±2.98E-2(-)**	1.1915±2.66E-1
	10,5	0.8193±9.67E-2(-)	0.6766±1.06E-2(+)	0.5444±5.19E-2(-)	1.1933±1.29E-5(=)	0.5809±2.49E-2(-)	**0.2895±3.06E-2(-)**	1.1974±2.75E-1
	5,10	0.8618±6.61E-2(-)	**0.3552±1.05E-2(-)**	0.5927±3.45E-2(-)	1.1923±1.65E-5(=)	0.6866±4.74E-2(-)	0.3953±4.32E-2(-)	1.2037±2.72E-1
DF13	10,10	0.2414±2.49E-3(+)	1.3416±3.41E+0(+)	0.2345±2.87E-3(+)	2.7312±1.09E+1(+)	0.2364±5.20E-3(+)	0.1587±1.91E-2(+)	**0.0668±1.40E-3**
	10,5	0.2660±2.30E-3(+)	1.3526±3.64E+0(+)	0.2351±2.57E-3(+)	2.6262±1.08E+1(+)	0.2478±3.42E-3(+)	0.1900±1.49E-2(+)	**0.0702±1.25E-3**
	5,10	0.2338±2.52E-3(+)	1.3573±3.50E+0(+)	0.2454±3.96E-3(+)	2.8032±1.14E+1(+)	0.2414±2.13E-3(+)	0.2735±3.40E-2(+)	**0.0693±1.13E-3**
DF14	10,10	0.0329±6.76E-3(+)	0.9623±2.05E+0(+)	0.0332±6.65E-3(+)	1.8257±6.69E+0(+)	0.0358±5.93E-3(+)	0.0555±6.89E-3(+)	**0.0150±2.00E-3**
	10,5	0.0339±2.95E-3(+)	0.9556±1.98E+0(+)	0.0368±4.41E-3(+)	1.6727±5.57E+0(+)	0.0351±5.95E-3(+)	0.0781±1.01E-2(+)	**0.0198±2.35E-3**
	5,10	0.0400±6.41E-3(+)	0.9730±2.18E+0(+)	0.0361±9.98E-4(+)	1.8333±6.06E+0(+)	0.0466±3.19E-3(+)	0.0838±1.01E-2(+)	**0.0276±3.36E-3**
F5	10,10	2.3272±8.15E+0(+)	**0.0768±1.99E-2(-)**	2.1466±1.80E+0(+)	2.6592±1.98E+0(+)	2.7516±4.27E+0(+)	2.5827±1.88E+0(+)	0.1092±7.19E-2
	10,5	2.4299±1.39E+1(+)	**0.1252±1.82E-2(-)**	2.3276±1.31E+0(+)	2.8026±3.26E+0(+)	2.4003±2.52E+0(+)	2.4366±1.85E+0(+)	0.1888±3.34E-2
	5,10	3.8499±4.96E+0(+)	0.6815±3.37E-1(+)	2.6935±1.35E+0(+)	3.6919±1.63E+0(+)	3.0477±2.21E+0(+)	4.5352±8.15E+0(+)	**0.0884±8.25E-3**
F6	10,10	1.4723±8.97E-1(+)	**0.0509±1.06E-3(-)**	1.8716±9.12E-1(+)	1.3095±8.35E-1(+)	1.3247±3.86E-1(+)	1.2826±6.51E-1(+)	0.2804±5.10E-2
	10,5	1.5399±2.44E+0(+)	0.4274±1.04E-1(+)	1.4738±4.62E+0(+)	1.2349±5.27E-1(+)	1.9116±2.30+00(+)	1.2529±3.994E-1(+)	**0.2251±8.25E-2**
	5,10	1.4189±8.39E-1(+)	1.2831±5.07E-1(+)	2.0589±2.38E+0(+)	2.4094±2.71E+0(+)	2.2788±1.67E+0(+)	2.3230±8.53E-1(+)	**0.3039±3.04E-2**
F7	10,10	1.2089±2.03E+0(+)	**0.0734±8.87E-3(-)**	1.3380±7.74E-1(+)	1.3270±5.77E-1(+)	1.6307±1.05E+0(+)	1.1449±1.98E-1(+)	0.2436±7.19E-2
	10,5	1.4417±5.84E+0(+)	0.2224±4.75E-2(+)	1.4501±2.21E+0(+)	1.4295±3.32E-1(+)	1.5664±1.77E+0(+)	1.2566±3.59E-1(+)	**0.0766±2.46E-3**
	5,10	1.7897±2.02E+0(+)	1.2676±4.51E-1(+)	1.7304±1.55E+0(+)	3.1593±8.77E-1(+)	1.9415±8.07E-1(+)	2.6158±1.19E+0(+)	**0.0941±2.06E-3**
F8	10,10	0.3216±6.11E-2(+)	0.4479±2.47E-2(+)	0.5578±1.01E-1(+)	0.7875±8.01E-2(+)	**0.2171±5.94E-3(-)**	0.8432±6.46E-2(+)	0.3157±7.93E-2
	10,5	0.4917±1.21E-2(+)	0.4334±1.93E-2(+)	0.7866±3.72E-1(+)	0.7331±8.94E-2(+)	**0.2393±2.18E-2(-)**	0.8125±2.51E-2(+)	0.3076±3.30E-2
	5,10	0.5349±2.52E-2(+)	0.5831±7.23E-2(+)	0.6813±1.05E-1(+)	1.0615±4.13E-1(+)	**0.3266±1.16E-2(-)**	1.3390±1.69E-1(+)	0.3536±1.38E-3
F9	10,10	0.9658±6.26E-1(+)	**0.1478±3.61E-2(-)**	1.5411±8.39E-1(+)	1.4721±5.94E-1(+)	0.8209±5.96E-1(+)	1.3766±6.30E-1(+)	0.2221±1.27E-3
	10,5	0.9949±2.62E+0(+)	**0.1769±2.62E-2(-)**	1.6188±5.61E-1(+)	1.4594±4.75E-1(+)	0.7225±2.34E-1(+)	1.3579±3.91E-1(+)	0.2079±3.30E-3
	5,10	1.7503±1.29E+0(+)	0.6782±3.47E-1(+)	2.7898±4.17E+0(+)	2.6079±1.14E+0(+)	1.7066±6.89E-1(+)	3.0071±2.05E+0(+)	**0.1797±3.34E-3**
F10	10,10	3.5221±7.01E+0(+)	1.1779±1.96E+0(+)	2.1921±1.87E+0(+)	2.7327±2.60E+0(+)	2.9650±5.78E+0(+)	4.1148±1.84E+1(+)	**0.1252±1.56E-3**
	10,5	1.8889±2.72E+0(+)	**0.1690±2.10E-2(-)**	1.9292±9.45E-1(+)	1.4407±6.08E-1(+)	0.6933±1.57E-1(-)	1.6451±9.44E-1(+)	0.9813±2.10E-3
	5,10	3.5543±7.53E+0(+)	0.6153±3.01E-1(+)	4.5335±2.21E+1(+)	2.7845±3.20E+0(+)	1.9263±1.46E+0(+)	3.1591±7.60E+0(+)	**0.0785±4.45E-3**
+/=/-		47/0/13	47/1/12	36/4/20	43/4/3	42/2/16	46/0/14	

- Negative IGD (*NIGD*)

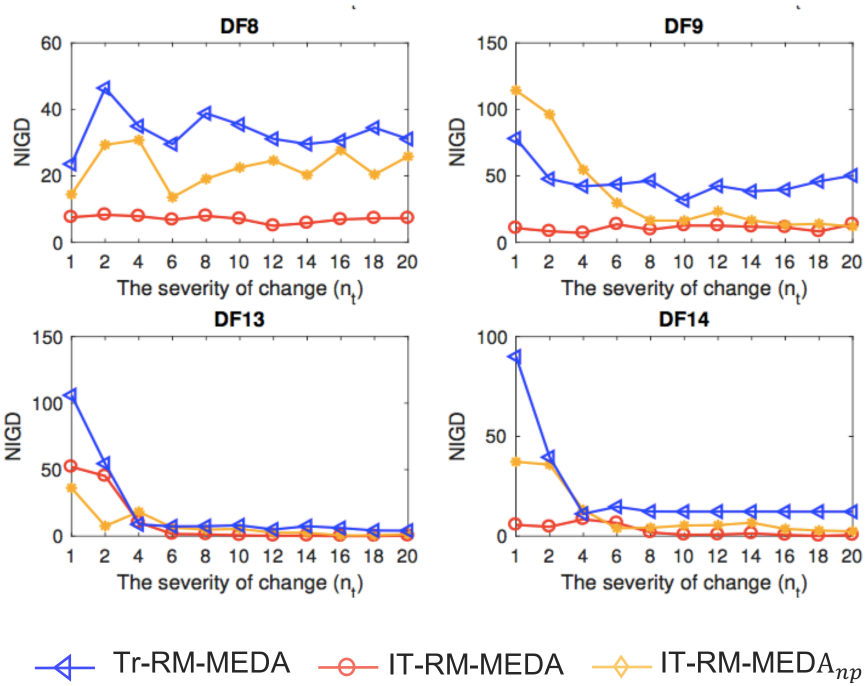
  Evaluate the degree of negative transfer:

$$NIGD = \sum_{t\in T}\left(IGD_t^{tr} - IGD_t^{rnd}\right)$$
$$s.t.\ IGD_t^{tr} > IGD_t^{rnd},$$

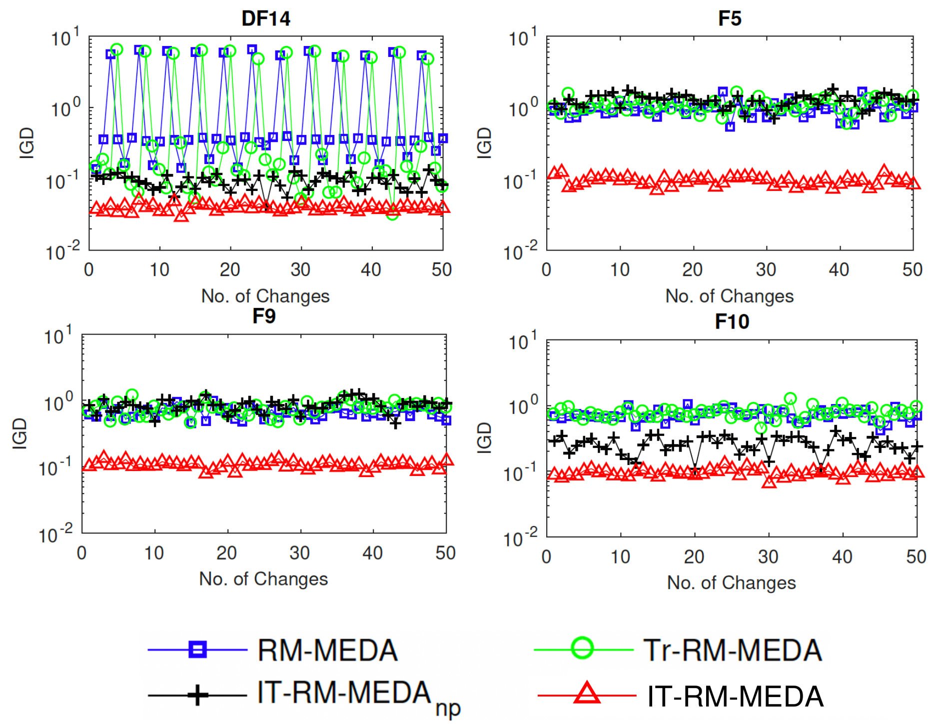
where IGD_t^{tr} and IGD_t^{rnd} is the IGD value obtained by the transferred population and the random population at time t, respectively.

- The experimental results suggest that IT-DMOEA can obtain solutions with good convergence and diversity for many of the benchmark functions.

IT-DMOEA: Experimental Study



Relationship between the severity of change and NIGD in the initial populations obtained by various methods.



Examples of averaged IGD values for various methods.

- The proposed IT-RM-MEDA (IT-DMOEA with RM-MEDA method) tends to have the smallest NIGD values which suggests that it is effective in preventing negative transfer generally.

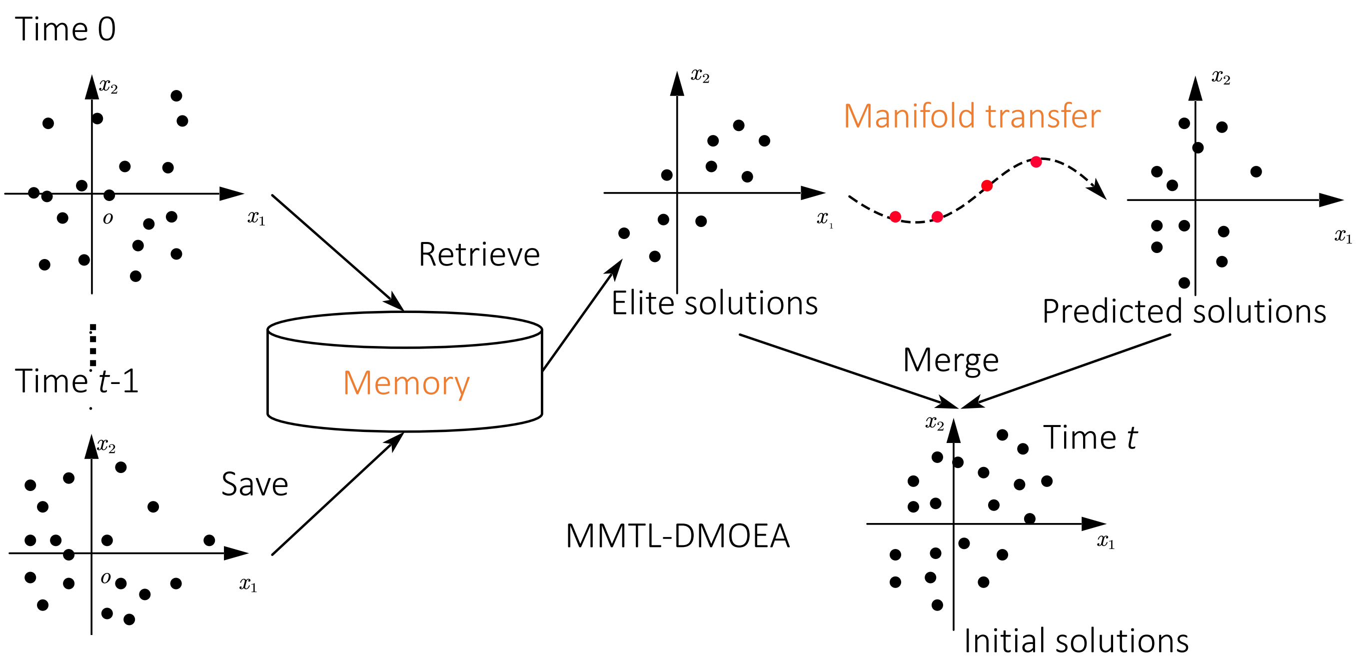
IT-DMOEA: Experimental Study on Run Time

Problems	IT-DMOEA	MDP	MOEA/D-SVR	Tr-DMOEA	MOEA/D-KF	SGEA	PPS
DF1	0.7999	0.9580	3.4562	46.1883	3.0723	**0.5219**	3.0559
DF2	0.8076	**0.4117**	3.3490	36.3649	3.2793	0.4849	2.9185
DF3	0.7375	0.5536	2.9751	45.9701	2.7094	**0.5013**	1.8899
DF4	0.6477	**0.4030**	2.3864	16.5985	2.3663	0.4174	1.7276
DF5	0.7324	0.4159	3.4126	39.9392	3.2197	**0.3346**	3.0275
DF9	**0.5523**	0.5732	2.6720	48.4561	2.5312	0.6145	2.0859
DF10	1.8495	2.4438	10.4172	86.6931	9.7104	**0.9557**	8.9799
DF11	**0.9427**	1.2538	8.9900	89.1901	9.1663	0.9535	8.1743
DF12	**0.8792**	6.2164	5.3619	97.0358	4.7006	3.5782	4.9663
DF13	1.2273	1.1792	10.5268	109.3813	9.8379	**0.9079**	8.5382
DF14	**1.2543**	1.6758	10.0918	90.8033	9.6494	2.4510	8.5602
F5	**1.3396**	10.6630	7.2160	51.3163	5.9335	4.9360	4.8181
F6	**1.3208**	10.5513	6.4391	54.5093	6.0874	4.4336	5.0109
F7	**1.3236**	9.7266	6.4625	57.1244	6.1244	4.3350	5.9100
F8	1.3874	3.2069	7.2878	126.9918	8.6568	**1.1961**	7.6505
F9	**1.1304**	9.8454	6.4259	97.0314	6.3404	4.6409	5.2085
F10	**1.1389**	10.4060	7.0571	59.7421	5.2330	4.3505	4.9030

The results show that the runtime of the proposed IT-DMOEA is the fastest on most of the benchmark functions.

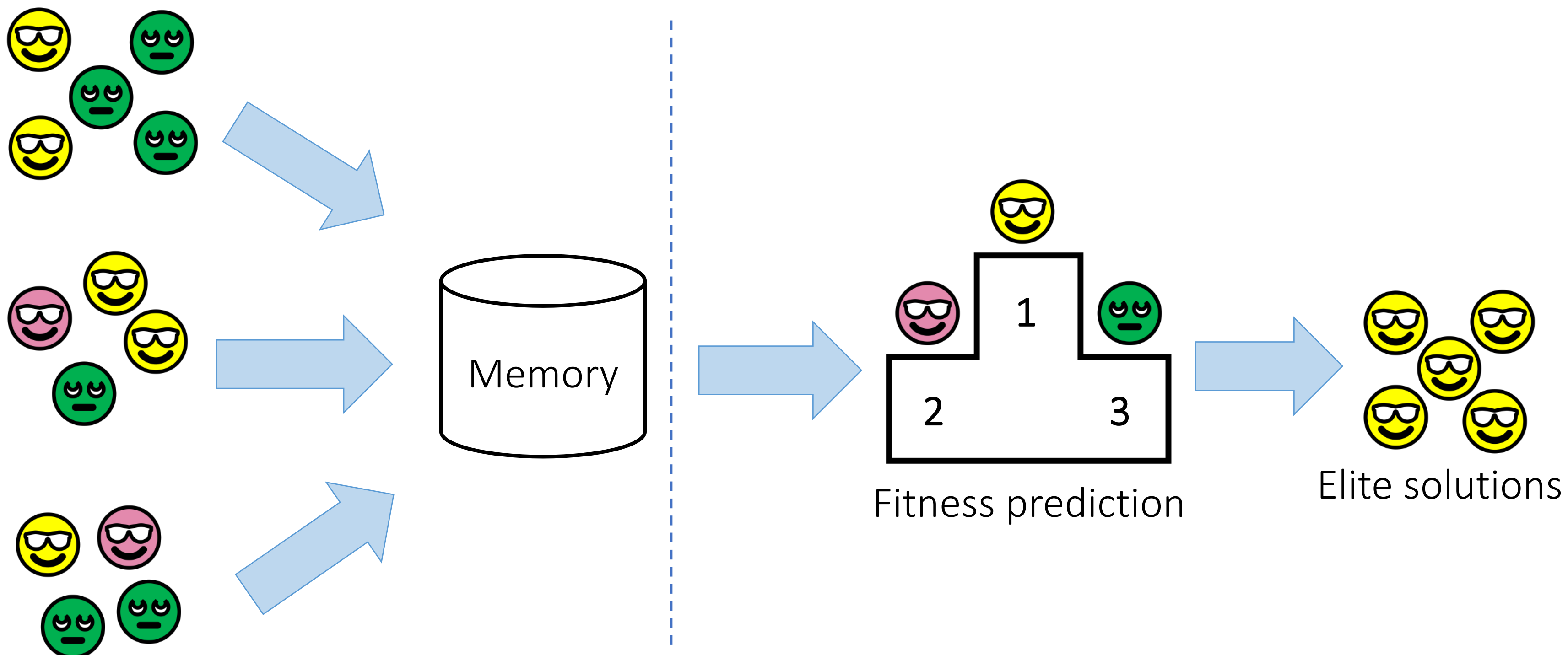
MMTL-DMOEA

- It is often desired to find and track the Pareto optimal solutions rapidly in DMOPs.
- Dynamic evolutionary multi-objective algorithm via manifold transfer learning (MMTL-DMOEA) combines a memory mechanism with a manifold transfer learning of sample geodesic flow (SGF).



- M. Jiang, Z. Wang, L. Qiu, S. Guo, X. Gao, and **K. C. Tan**, "A Fast Dynamic Evolutionary Multiobjective Algorithm via Manifold Transfer Learning", *IEEE Transacts on Cybernetics*, in press, 2020.

MMTL-DMOEA: Memory



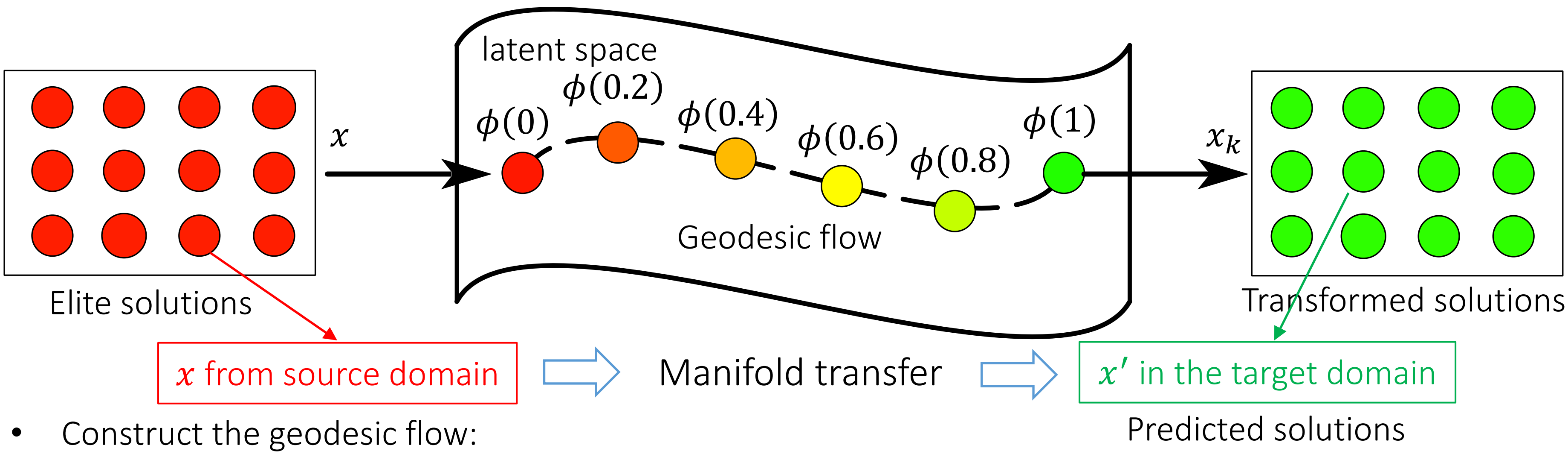
Saving of Solutions

- Solutions from past environments are stored in external memory.
- When the external memory overflows, the earliest stored individuals are replaced.

Retrieving of Solutions

- Once any environment change is detected, fitness values of individuals from memory are predicted.
- Non-dominated solutions are selected as the elite solutions.

MMTL-DMOEA: Manifold Transfer



- Construct the geodesic flow:

$$\phi(k) = P_s U_1 \Gamma(k) - R_s U_2 \Sigma(k)$$

$\phi(0) = P_s$ and $\phi(1) = P_T$, P_s, P_T is the covariance matrices (produced by PCA) of elite solutions and good random solutions, respectively; U_1, U_2 are orthogonal matrices; $\Gamma(k)$ and $\Sigma(k)$ are diagonal matrices; R_s is the orthogonal complement of P_s; k determines the number of points on the manifold.

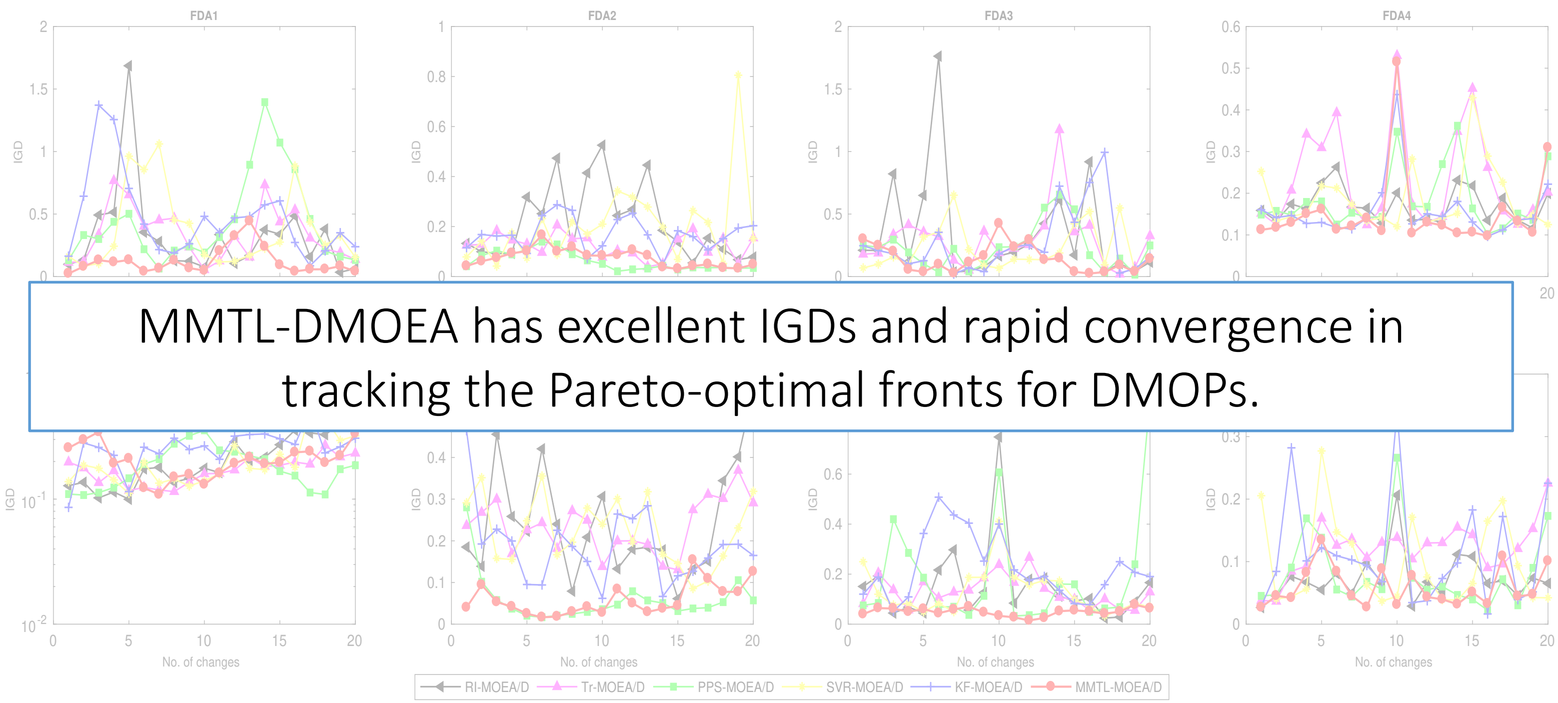
- Transform elite solutions:

$$x_k = x^T \phi(k), k \in (0,1)$$

- Find predicted solutions: The transformed latent space is like the target domain, e.g., the predicted solution x' is like x_k on the geodesic flow $\phi(\cdot)$,

$$x' = argmin\ ||x'^T \phi(\cdot) - x_k||$$

MMTL-DMOEA: Experimental Study



MMTL-DMOEA has excellent IGDs and rapid convergence in tracking the Pareto-optimal fronts for DMOPs.

Run Time Performance (Unit: Seconds)

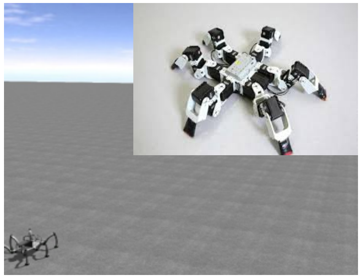
	FDA1	FDA2	FDA3	FDA4	FDA5	dMOP1	dMOP2	dMOP3
Tr-MOEA/D	42.54	45.27	57.71	132.59	115.52	80.23	73.53	75.55
MMTL-MOEA/D	6.23	5.4	5.09	10.06	9.94	7.05	7.74	5.63
Speed Improvement	6.82	8.38	11.33	13.18	11.62	11.38	95.00	13.42

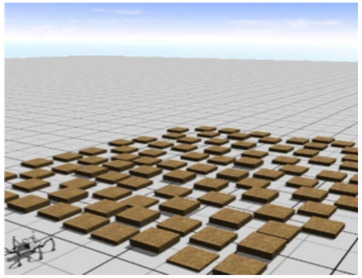
OUTLINE

Gait Generation & Optimization

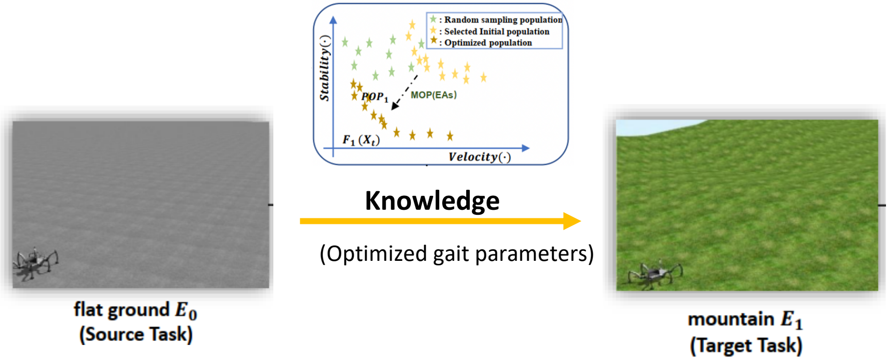
- Robot gait refers to the periodic movement of legged robot joints. Gait optimization is to generate the optimal control trajectory of a robot under different internal and external constraints.

- It can be solved as a multi-objective problem by defining multiple designated objective functions, e.g. speed, stability, energy efficiency etc.





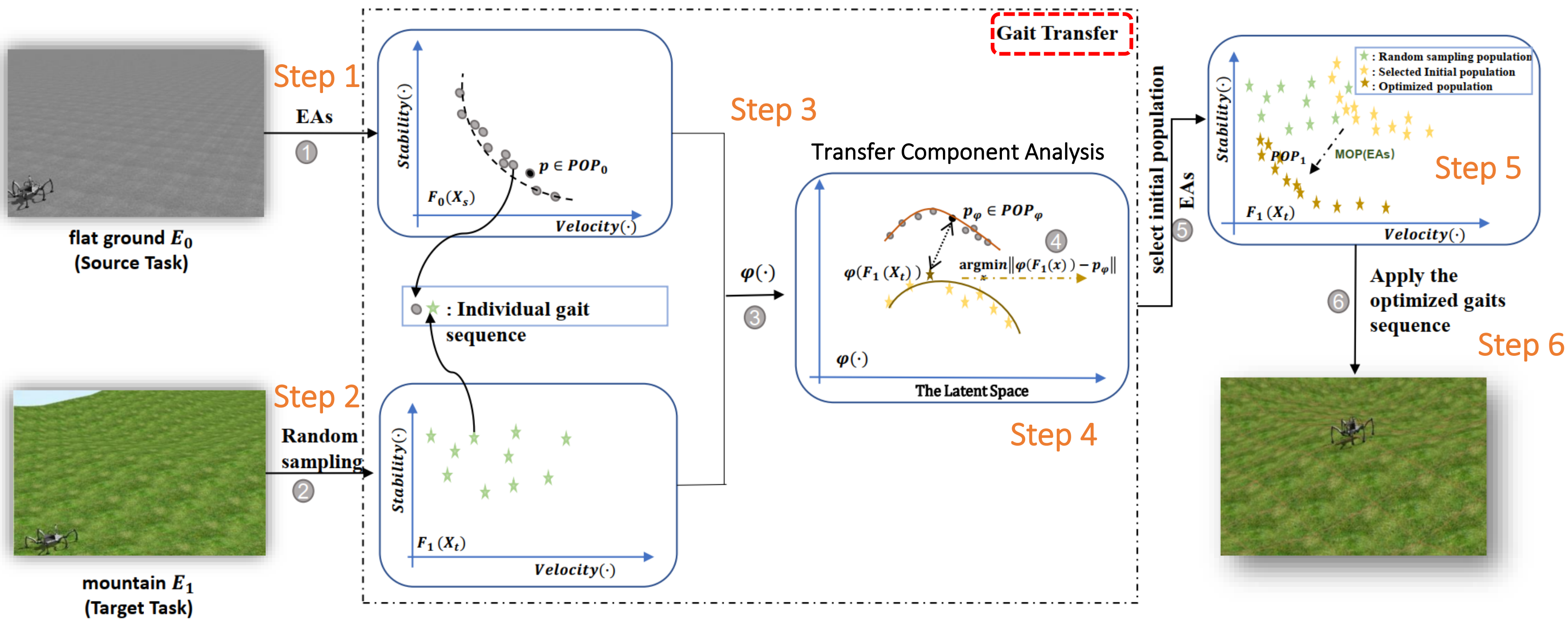
- In many situations, different gait optimization is needed under different external constraints and environments.



Transfer learning can help to obtain high-quality gait movements efficiently in a new environment by re-using/ transfer knowledge from optimal gaits in previous environments.

Gait Generation & Optimization

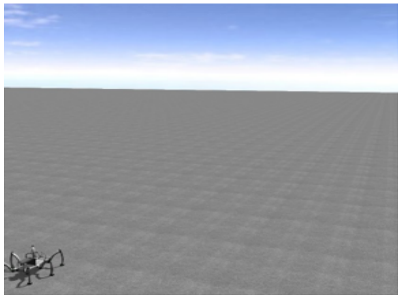
- From the perspective of transfer learning, the gaits trained in simple and complex environments can be regarded as the source and target domain, respectively. These two tasks are 'similar' in that they both attempt to generate a gait that optimizes performance in a specific environment.



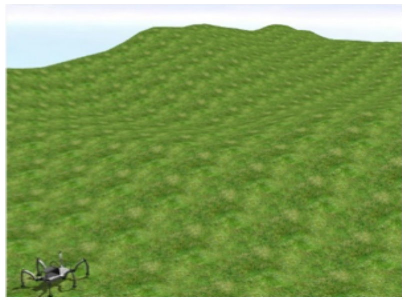
Tr-GO

Experimental Results

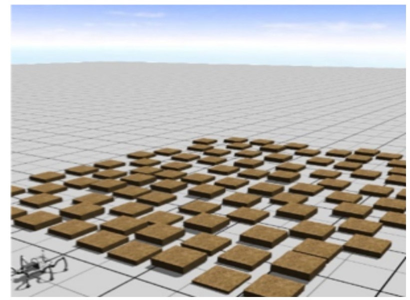
- Four different experimental environments are studied.



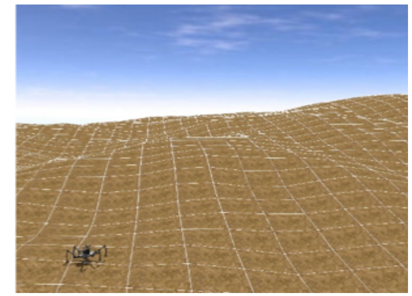
(a) Source environment E0



(b) Target environment E1



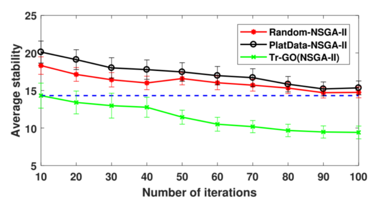
(c) Target environment E2

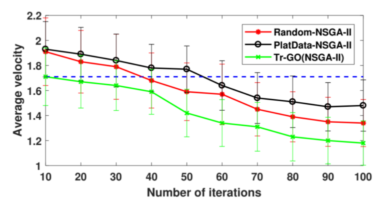


(d) Target environment E3

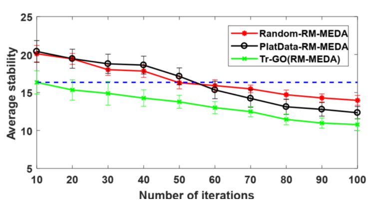
- Convergence trace

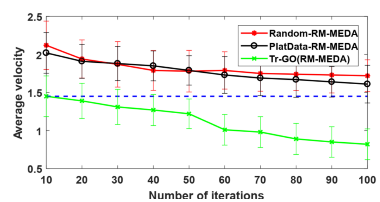
E1



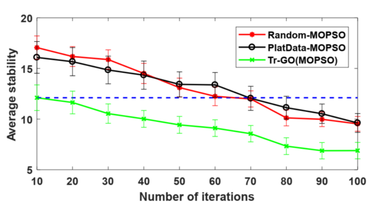


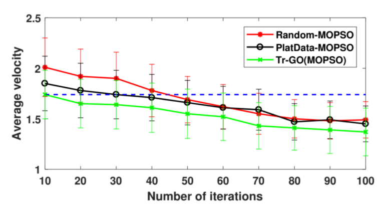
E2





E3

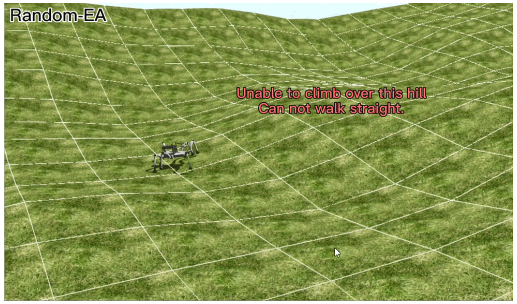




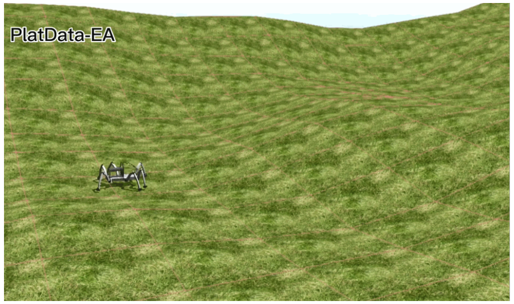
- Convergence traces of different algorithms under various environments. The blue dashed line (benchmark) is the result achieved by Tr-GO after the first 10 generations of evolution.

Experimental Results

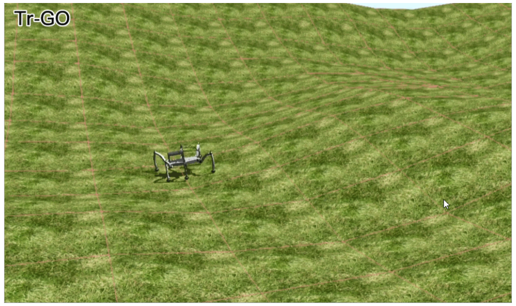
- The robot walking in E1 environment



(a) Random-EA



(b) PlatData-EA

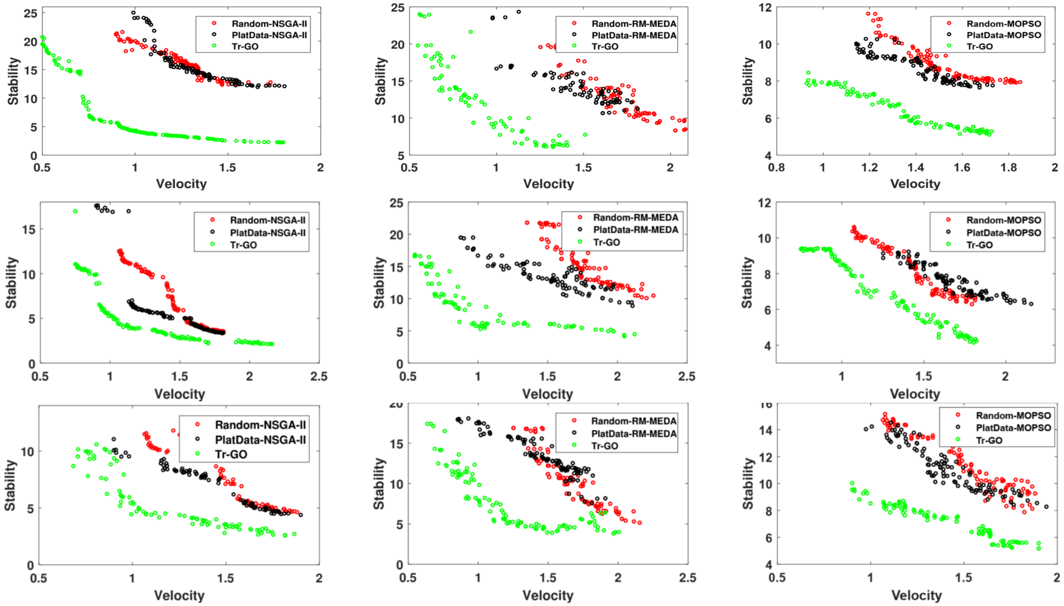


(c) Tr-Go

- The trade-off graphs

NSGA-II RM-MEDA MOPSO

E1 E2 E3



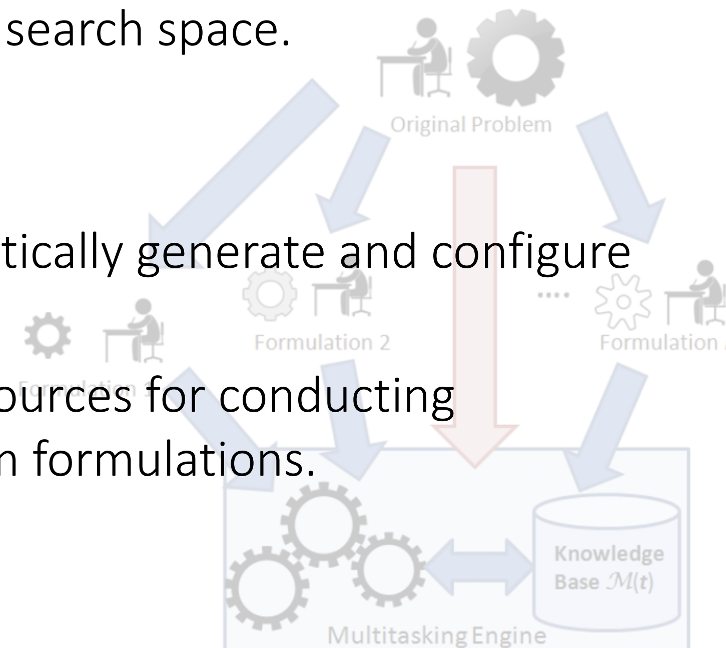
- The trade-off graphs under different environment of E1, E2, and E3. Each algorithm is evaluated under three settings: random sampling (in red), direct use of samples from E0 (in black) and the proposed Tr-Go algorithm (in green).

OUTLINE

Future Research Topics

- ETO for Large-scale Optimization
  - Solving large-scale optimization problems via transfer and learning from related/smaller and/or solved problems (such as via cooperative coevolution etc.).
  - Designing ETO algorithms capable of solving large-scale problems in a knowledge space having a much smaller search space.

- ETO for Multi-form Optimization
  - Designing algorithms capable of automatically generate and configure different formulations of a problem.
  - Efficient allocation of computational resources for conducting evolutionary search on different problem formulations.



Future Research Topics

- ETO in Complex Data Environment
  - Designing ETO algorithms capable of positive knowledge transfer from from noisy data and across problem domains where the data appears in a sequential order.
  - Designing ETO algorithms for problems having imbalance data (with/without labels) or data with property changes rapidly in uncertain environments.

- Theoretical Study of ETO
  - Study on how and when knowledge in source problem can help to improve the search in a target task.
  - Defining useful representation of knowledge that can be transferred across heterogeneous problem domains.

Conclusions

- ETO is an emerging paradigm that integrates EA solvers with knowledge learning and transfer across related domains to achieve efficient and better optimization performance.

- ETO has been applied to multi-task and dynamic optimization. Enhanced optimization performance can be achieved with knowledge learning and transfer across problems.

- ETO has been applied to robot gait optimization by transferring knowledge across environments. Knowledge transfer along the evolutionary search can further improve the gait optimization performance in dynamic environments.

- As one of the emerging research areas in Computational Intelligence, there are many challenges and open research questions in ETO.

Acknowledgements

- Research Collaborators
    - Prof. Liang Feng, Chongqing University, China
    - Prof. Min Jiang, Xiamen University, China
    - Prof. Hai-Lin Liu, Guangdong University of Technology, China
    - Dr. Ye Tian, Anhui University, China (Hong Kong Scholars Program)

- PhD Students (City University of Hong Kong)
    - Zhian Huang
    - Xuezhou Fan
    - Songbai Liu
    - Yinglan Feng
    - Xun Zhou
    - Weibo Shu
    - Yao Hu
    - Xiaoming Xue
    - Lyuyang Tong (Joint PhD Program)
    - Rui Liu (MEng)

THANKS.

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional · Creative
For The World