

To Preserve or Not to Preserve Invalid Solutions in Search-Based Software Engineering: A Case Study in Cloud Cost Optimization

Jianmei Guo

Alibaba Group

2018.11.17 @CSBSE

Search-based software engineering (SBSE)

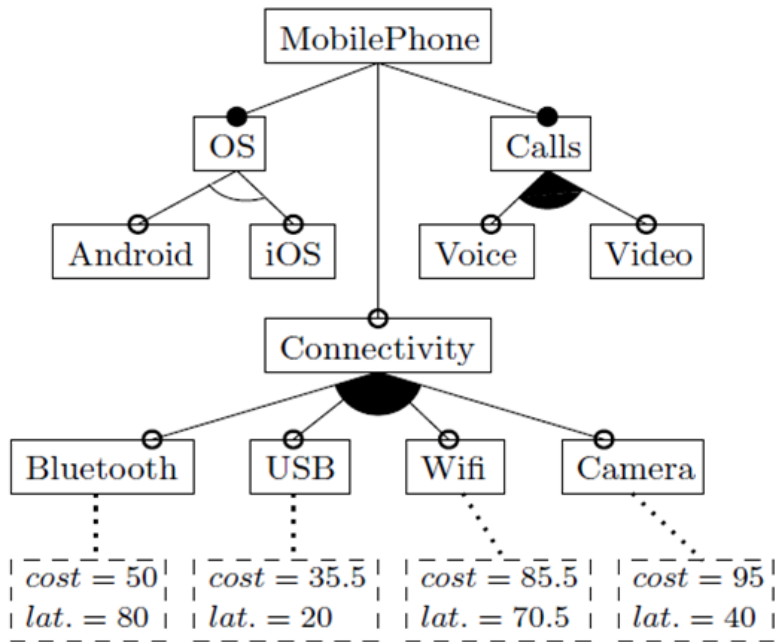
Search-based software engineering (SBSE)

**Should we preserve invalid solutions
during the search?**

Outline

- A case study in software produce lines (SPLs)
 - SPL optimization problem
 - Five algorithm variants based on the state-of-the-art
 - Seven real-world subjects (*a particular one in cloud cost optimization*)
- Three potential limitations of preserving invalid solutions
- Empirical guidance on how to preserve valid solutions
 - How to choose the population for performance evaluation?
 - Should we keep valid solutions in the initial population?
 - Should we fix invalid solutions after mutation?
 - Should we preserve valid solutions all along the way?
 - Should we preserve invalid solutions during the search?

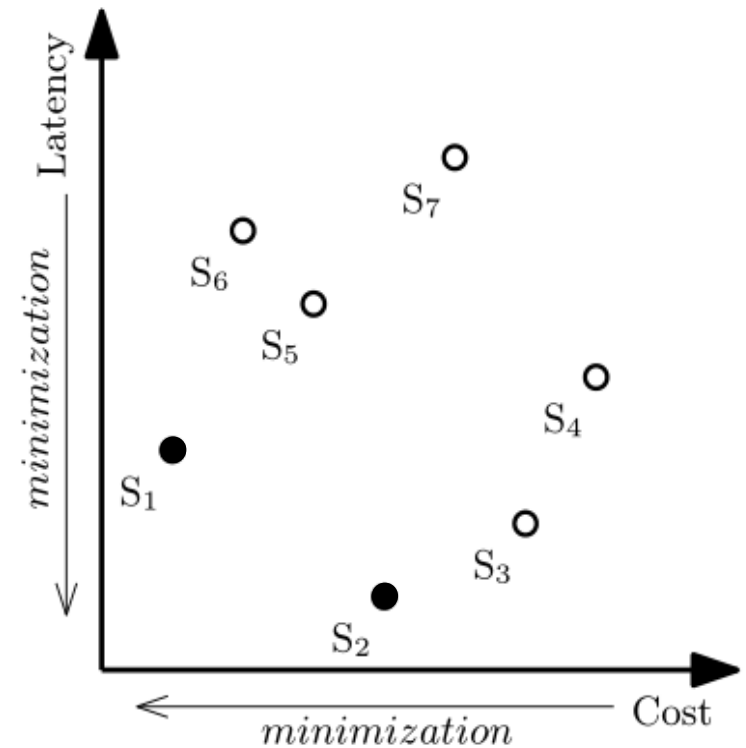
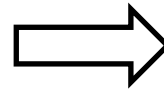
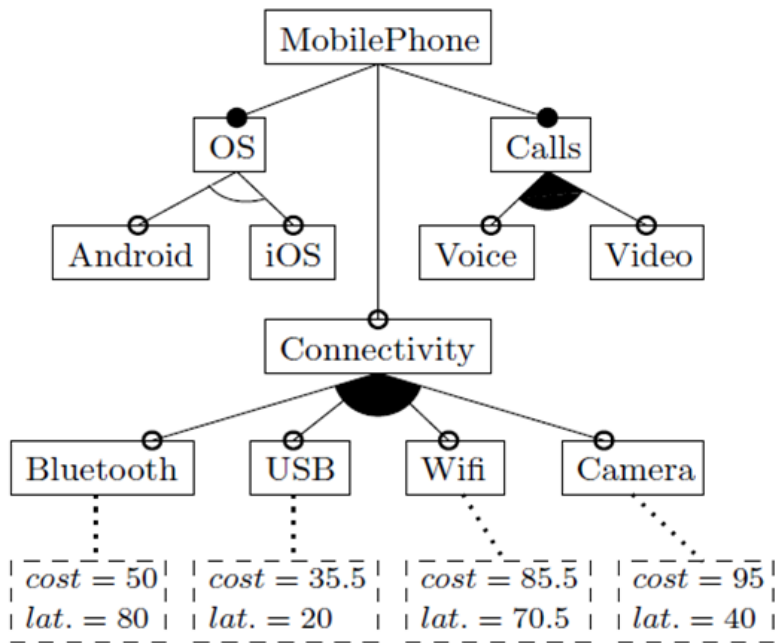
SPL Optimization Problem



Cross-tree constraints: Video requires Camera

Objectives: minimizing cost, minimizing latency

SPL Optimization Problem



Cross-tree constraints: Video requires Camera

Objectives: minimizing cost, minimizing latency

A multi-objective combinatorial optimization problem

Five Algorithm Variants

- Multi-objective evolutionary algorithms (MOEAs)
 - IBEA [Sayyad et al., ICSE'13, ASE'13]
 - SATIBEA [Henard et al., ICSE'15]

Five Algorithm Variants

- Multi-objective evolutionary algorithms (MOEAs)
 - IBEA [Sayyad et al., ICSE'13, ASE'13]
 - SATIBEA [Henard et al., ICSE'15]

Algorithm	λ	δ	TIP
SATIBEA	0.001	0.02	random
SATIBEA _{v1}	0.001	0.02	valid
SATIBEA _{v2}	0.001	0.02	random + rich seed
SATIBEA _{v3}	1	1	random
SATIBEA _{v4}	0.001	1	random
SATIBEA _{v5}	1	1	valid

Mutation
rate

Probability of using SAT
solving for mutation

Type of initial
population

Five Algorithm Variants

- Multi-objective evolutionary algorithms (MOEAs)

Algorithm	λ	δ	TIP
SATIBEA	0.001	0.02	random
SATIBEA _{v1}	0.001	0.02	valid
SATIBEA _{v2}	0.001	0.02	random + rich seed
SATIBEA _{v3}	1	1	random
SATIBEA _{v4}	0.001	1	random
SATIBEA _{v5}	1	1	valid

Preserving validity “**all along the way**” by incorporating a subroutine that resolves **non-Boolean** constraints over integer or real-number variables together with arithmetic or relational operators

Seven Subjects

- Five largest SPLs hitherto reported in the literature
- Two SPLs with realistic values and constraints of attributes

SPL	Version	#Features	#CTC	#Attr	Type of Attr Values	#Attr Constraints
ECOS	3.0	1,244	3,146	3	synthetic	N/A
FREEBSD	8.0.0	1,396	62,183	3	synthetic	N/A
FIASCO	2011081207	1,638	5,228	3	synthetic	N/A
UCLINUX	20100825	1,850	2,468	3	synthetic	N/A
LINUX	2.6.28.6	6,888	343,944	3	synthetic	N/A
DRUPAL	7.23	47	23	22	realistic	N/A
AMAZONEC2	12 June 2014	81	3,859	17	realistic	17,049

[Henard et al., ICSE'15]

[Sánchez et al., Software & Systems Modeling, 2017]

[García-Galán et al., Future Generation Comp. Syst., 2016]

Potential Limitations – 1/3

- Ratios of invalid solutions in the final populations produced by state-of-the-art approaches

SPL	IBEA	SATIBEA
ECOS	0%	4.94%
FREEBSD	2%	89.02%
FIASCO	0%	63.86%
UCLINUX	69%	0.02%
LINUX	100%	81.83%

IBEA [Sayyad et al., ICSE'13, ASE'13]

SATIBEA [Henard et al., ICSE'15]

Potential Limitations – 2/3

Performance evaluation depends mainly on the metrics calculated on the **population** produced

➤ Including invalid solutions

➤ Excluding invalid solutions

SATIBEA_{v1}

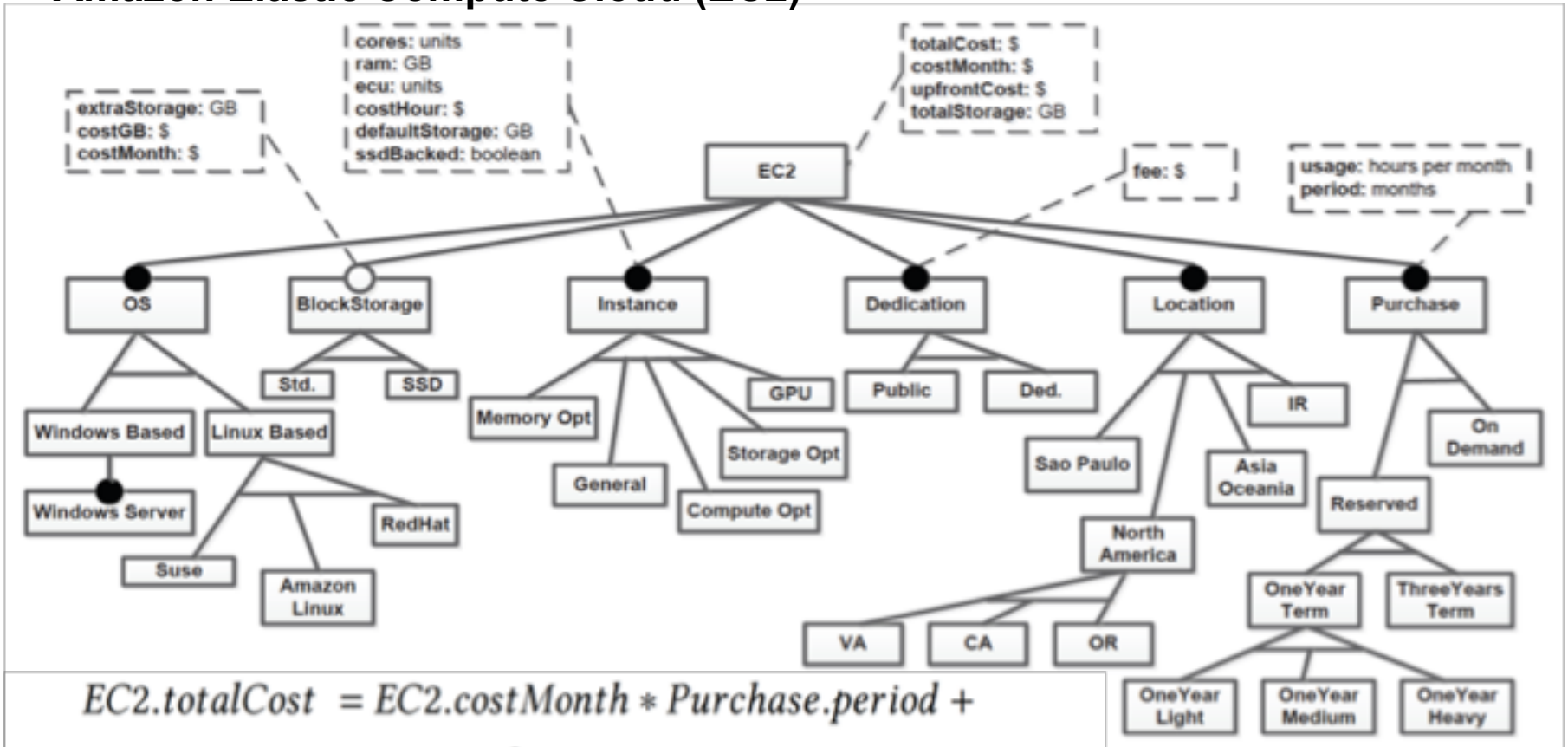
SATIBEA

SPL	Metric	SATIBEA (v0)		SATIBEA (v1 (v))		SATIBEA (v2 (v))		v0 VS v1		v0 VS v2		v1 VS v2	
		Median	Mean	Median	Mean	Median	Mean	p-value	$\hat{\Delta}_{12}$	p-value	$\hat{\Delta}_{12}$	p-value	$\hat{\Delta}_{12}$
vCos	HV	2.8196e-1	2.8172e-1	2.8119e-1	2.8111e-1	2.8111e-1	2.8111e-1	0.978	1.6276e-1	0.486	6.1284e-1	0.534	
	σ	6.7436e-2	6.9626e-2	7.1173e-2	7.3257e-2	7.4465e-2	7.6994e-2	9.8816e-2	0.372	3.9514e-1	0.436	5.1945e-1	0.530
	IGD	5.1243e-4	5.0866e-4	4.9895e-4	4.9895e-4	5.0793e-4	5.0576e-4	6.3032e-2	0.640	6.2969e-1	0.530	2.2257e-1	0.448
	ER	2.3037e-1	2.4739e-1	2.4213e-1	2.3367e-1	2.2967e-1	2.4213e-1	4.9312e-1	0.443	8.0723e-1	0.389	2.9756e-1	0.579
	GS	8.7903e-1	8.9232e-1	9.0625e-1	9.0491e-1	8.8484e-1	8.8439e-1	3.2533e-1	0.426	9.1217e-1	0.389	1.1799e-1	0.620
	PF5	168	167.76	168	168.47	167	168.4	8.8229e-1	0.485	7.6723e-1	0.477	9.7946e-1	0.497
FaceBND	HV	2.6291e-1	2.6299e-1	2.6526e-1	2.6497e-1	2.6366e-1	2.6577e-1	6.2827e-4	0.262	1.7145e-1	0.397	3.2814e-2	0.683
	σ	6.6625e-2	6.8796e-2	7.0466e-2	7.0977e-2	6.9907e-2	6.8553e-2	2.8926e-1	0.419	7.6769e-1	0.477	6.3122e-1	0.539
	IGD	4.5249e-4	4.5296e-4	4.5713e-4	4.5446e-4	4.5975e-4	4.4813e-4	6.7356e-1	0.446	5.1877e-1	0.447	6.7692e-1	0.638
	ER	1.6839e-1	1.6883e-1	1.6863e-1	1.5227e-1	1.6484e-1	1.9427e-1	3.7273e-1	0.718	1.7894e-1	0.390	1.9394e-1	0.526
	GS	8.7736e-1	9.0226e-1	8.9631e-1	9.3391e-1	9.3451e-1	9.2891e-1	2.1792e-1	0.467	6.3886e-2	0.348	4.8252e-1	0.447
	PF5	239	218.83	219	215.6	228.5	218.9	4.4136e-1	0.538	7.8992e-1	0.479	2.8391e-1	0.619
Faceco	HV	2.4839e-1	2.4839e-1	2.4413e-1	2.4396e-1	2.4489e-1	2.4489e-1	3.4829e-1	0.372	2.8501e-1	0.388	8.7631e-1	0.512
	σ	1.8437e-1	1.8596e-1	1.8963e-1	1.9424e-1	1.8946e-1	1.8783e-1	2.4422e-1	0.331	8.3026e-1	0.483	7.6262e-1	0.634
	IGD	7.8366e-4	7.6173e-4	8.1268e-4	8.1795e-4	7.8914e-4	7.8796e-4	3.6482e-2	0.342	5.3953e-1	0.453	7.4827e-1	0.634
	ER	2.5833e-1	2.5676e-1	2.7193e-1	2.7906e-1	2.6899e-1	2.7424e-1	4.1179e-1	0.438	5.7963e-1	0.458	8.6977e-1	0.513
	GS	6.8637e-1	6.1292e-1	6.2199e-1	6.2696e-1	6.2803e-1	6.3633e-1	3.6875e-1	0.396	5.6723e-1	0.472	5.9991e-1	0.640
	PF5	115	119.67	112	111.93	115	115.57	7.8386e-2	0.635	3.9827e-1	0.577	5.6823e-1	0.629
vCircex	HV	2.9399e-1	2.9325e-1	2.9593e-1	2.9429e-1	2.9466e-1	2.9199e-1	4.2691e-1	0.439	2.8129e-2	0.686	1.6796e-3	0.737
	σ	6.4749e-2	6.4403e-2	6.3492e-2	6.4177e-2	6.2796e-2	6.2299e-2	8.2415e-1	0.317	2.3466e-1	0.386	2.7631e-1	0.582
	IGD	5.6187e-4	5.8934e-4	5.9923e-4	5.7262e-4	5.7796e-4	5.7793e-4	9.7602e-1	0.497	9.8291e-1	0.498	9.8331e-1	0.582
	ER	3.2637e-1	3.2826e-1	3.3506e-1	3.3779e-1	3.4099e-1	4.0522e-1	4.9346e-1	0.448	6.4722e-1	0.362	2.4477e-1	0.616
	GS	8.5866e-1	8.5499e-1	8.5920e-1	8.5721e-1	8.5439e-1	8.5739e-1	4.8149e-1	0.516	6.3886e-1	0.463	5.4971e-1	0.649
	PF5	106	107.33	107	108.2	109	108.57	5.8866e-1	0.439	6.0469e-1	0.443	8.3761e-1	0.683
Leaves	HV	2.2424e-1	2.2495e-1	2.2664e-1	2.2739e-1	2.2499e-1	2.2455e-1	4.8198e-3	0.283	7.9466e-1	0.521	3.5636e-2	0.682
	σ	1.1175e-1	1.1899e-1	1.1125e-1	1.0886e-1	1.1335e-1	1.1335e-1	9.3319e-1	0.367	6.6273e-1	0.467	6.3987e-1	0.643
	IGD	5.5243e-4	5.3939e-4	5.3563e-4	5.7242e-4	5.6975e-4	5.3502e-4	8.4188e-1	0.484	7.9951e-1	0.526	6.9751e-1	0.512
	ER	2.4239e-1	2.4889e-1	1.7891e-1	1.6372e-1	2.3743e-1	2.4899e-1	3.8274e-1	0.868	8.2889e-1	0.357	3.6896e-1	0.511
	GS	9.4874e-1	9.5833e-1	9.3312e-1	9.8842e-1	9.3216e-1	9.6783e-1	4.2039e-1	0.439	8.9999e-1	0.538	2.8376e-1	0.581
	PF5	290	290.83	289	288.2	289.8	288.73	3.3395e-1	0.613	2.4899e-1	0.387	6.3967e-1	0.644

SPL	Metric	SATIBEA (v0)		SATIBEA (v1 (v))		SATIBEA (v2 (v))		v0 VS v1		v0 VS v2		v1 VS v2	
		Median	Mean	Median	Mean	Median	Mean	p-value	$\hat{\Delta}_{12}$	p-value	$\hat{\Delta}_{12}$	p-value	$\hat{\Delta}_{12}$
vCos	HV	2.8894e-1	2.8823e-1	2.7893e-1	2.7929e-1	2.7905e-1	2.7905e-1	7.0225e-2	0.637	1.3732e-1	0.612	5.9969e-1	0.460
	σ	6.7036e-2	6.9033e-2	7.1173e-2	7.3229e-2	7.0879e-2	7.0879e-2	8.4933e-2	0.370	3.9926e-1	0.436	4.9630e-1	0.552
	IGD	5.7999e-4	5.7939e-4	5.6992e-4	5.7366e-4	5.7972e-4	5.7667e-4	9.4869e-1	0.494	5.7964e-1	0.433	4.5330e-1	0.443
	ER	2.3863e-1	2.5229e-1	2.9222e-1	2.5127e-1	2.3355e-1	2.4203e-1	4.7793e-1	0.446	7.2826e-1	0.527	2.8846e-1	0.582
	GS	8.1646e-1	8.2489e-1	8.1389e-1	8.2515e-1	8.0883e-1	8.0523e-1	7.3699e-1	0.476	4.8432e-1	0.358	2.5895e-1	0.586
	PF5	168	168.2	167.5	168.79	168	168.23	3.1013e-1	0.376	7.8842e-1	0.529	6.2314e-1	0.603
FaceBND	HV	7.0375e-1	8.1413e-1	9.3394e-1	8.3262e-1	5.6687e-1	8.6644e-1	8.1875e-1	0.482	9.9439e-1	0.361	7.8617e-1	0.529
	σ	3.5237e-1	3.5546e-1	3.6446e-1	3.4277e-1	3.3386e-1	3.5494e-1	8.3025e-1	0.483	7.8446e-1	0.321	5.8996e-1	0.546
	IGD	7.7393e-4	8.6763e-4	9.2703e-4	8.6299e-4	7.8816e-4	8.9323e-4	7.1719e-1	0.472	5.0134e-1	0.351	5.4829e-1	0.572
	ER	3.8599e-1	3.7636e-1	1.9796e-1	3.8669e-1	4.5296e-1	4.2925e-1	2.9831e-1	0.579	4.1170e-1	0.438	8.3333e-1	0.369
	GS	1.0394e-1	1.1299e-1	1.0414e-1	1.1196e-1	1.2429e-1	1.2269e-1	7.9955e-1	0.480	7.9792e-1	0.368	9.6263e-1	0.474
	PF5	15	16.9	9.5	13.77	11.5	15.67	5.0575e-2	0.647	9.8843e-1	0.628	8.8217e-1	0.658
Faceco	HV	2.5546e-1	2.5599e-1	2.5364e-1	2.5386e-1	2.5286e-1	2.5293e-1	2.6433e-1	0.594	2.0681e-2	0.674	1.8233e-1	0.623
	σ	1.8812e-1	1.8896e-1	1.9375e-1	1.9237e-1	1.9218e-1	1.9185e-1	5.3049e-1	0.290	1.5312e-1	0.392	1.5793e-1	0.687
	IGD	1.9999e-3	1.9882e-3	2.0839e-3	2.0636e-3	1.9883e-3	1.9983e-3	1.5636e-2	0.318	3.6323e-1	0.431	5.3895e-1	0.646
	ER	3.6977e-1	2.9884e-1	3.1213e-1	3.3346e-1	3.4549e-1	3.4525e-1	4.0933e-1	0.437	2.1582e-1	0.527	1.8895e-1	0.399
	GS	7.8793e-1	7.9859e-1	7.2691e-1	7.3262e-1	6.7259e-1	6.8725e-1	7.0675e-1	0.471	5.9949e-1	0.540	1.6236e-1	0.686
	PF5	36	39.43	32	31.7	31	31.77	1.8996e-2	0.676	1.6666e-2	0.679	7.3769e-1	0.526
vCircex	HV	2.9399e-1	2.9323e-1	2.9393e-1	2.9429e-1	2.9466e-1	2.9199e-1	4.2691e-1	0.439	2.8129e-2	0.646	1.6796e-3	0.737
	σ	6.4749e-2	6.4403e-2	6.3492e-2	6.4177e-2	6.2796e-2	6.2299e-2	8.2415e-1	0.317	2.3466e-1	0.386	2.7631e-1	0.582
	IGD	5.6187e-4	5.6143e-4	5.2942e-4	5.4339e-4	5.4894e-4	5.4874e-4	9.8671e-1	0.504	9.2344e-1	0.368	9.8231e-1	0.582
	ER	3.2637e-1	3.2634e-1	3.3389e-1	3.3779e-1	3.4099e-1	4.0522e-1	4.9346e-1	0.448	6.4721e-1	0.362	2.4477e-1	0.616
	GS	8.5866e-1	8.5429e-1	8.4818e-1	8.5675e-1	8.4462e-1	8.4696e-1	9.3591e-1	0.507	5.6923e-1	0.437	5.4973e-1	0.654
	PF5	106	107.3	107	108.2	109	108.57	5.8866e-1	0.439	6.0469e-1	0.443	8.3761e-1	0.684
Leaves	HV	1.6479e-1	1.6397e-1	1.6423e-1	1.6475e-1	1.5979e-1	1.6268e-1	4.9176e-1	0.512	7.2446e-1	0.436	5.7964e-1	0.567
	σ	1.2193e-1	1.1796e-1	1.1897e-1	1.1781e-1	1.1416e-1	1.2199e-1	8.7863e-1	0.512	1.6895e-1	0.396	2.2823e-1	0.609
	IGD	4.7996e-4	4.7752e-4	4.7887e-4	5.1449e-4	4.8516e-4	4.7974e-4	2.7973e-1	0.417	6.8432e-1	0.469	4.9746e-1	0.539
	ER	3.6885e-1	3.6881e-1	2.3312e-1	2.3999e-1	2.8132e-1	2.8193e-1	6.1556e-1	0.505	1.1874e-1	0.618	1.8911e-1	0.396
	GS	9.8496e-1	9.8227e-1	1.0189e-1	1.0356e-1	9.9702e-1	9.7493e-1	9.5314e-1	0.373	3.6323e-1	0.569	2.3243e-2	0.671
	PF5	47.3	49.73	49	49.65	49	48.73	4.0275e-1	0.563	9.953e-1	0.509	5.7346e-1	0.457

Potential Limitations – 3/3

Amazon Elastic Compute Cloud (EC2)



$$EC2.totalCost = EC2.costMonth * Purchase.period + EC2.upfrontCost$$

$$EC2.costMonth = (Inst.costHour + Inst.dedicatedFee) * Purchase.signedUsage + BlockS.costMonth$$

$$BlockS.costMonth = BlockS.extraSpace * BlockS.costGB$$

Excerpted from
[García-Galán et al., 2016]

Empirical Guidance – 1/5

How to choose the population for performance evaluation?

- Inconsistent results when including and excluding invalid solutions in the population
- Invalid solutions are unbuildable and thus useless in practice

Considering only valid solutions, ***at least***,
in the final population

Empirical Guidance – 2/5

Should we keep valid solutions in the initial population?

- Empirical results cannot statistically distinguish three algorithms adopting different initial populations in terms of any quality metric
- SATIBEA is the least and the easiest to build

Using random strategy to generate the initial population

Empirical Guidance – 3/5

Should we fix invalid solutions after mutation?

- Experimental results on six subjects demonstrate the superiority of SATIBEA among three algorithms adopting different mutation operators.

Adopting a lower mutation rate (0.001) and a lower probability (0.02) to fix a solution to be valid after mutation

Empirical Guidance – 4/5

Should we preserve valid solutions all along the way?

- Valid solutions must be preserved all along the way for AmazonEC2
- For the other subjects without attribute constraints, SATIBEA works better for 4 out of 6 subjects.

Preserving valid solutions all along the way might not always be necessary.

Empirical Guidance – 5/5

Should we preserve invalid solutions during the search?

Whether or not to preserve invalid solutions deserves **more attention** in SBSE, and in some cases, we have to preserve valid solutions all along the way during the search.

Conclusion

- **Open question:** *Should we preserve invalid solutions during the search?*
- Three **potential limitations** of preserving invalid solutions
 - No guarantee to produce valid solutions
 - Evaluation results might be misleading
 - Complex attribute constraints break the fitness calculation based on simple aggregation
- **Empirical guidance** on how to preserve valid solutions
 - Considering only valid solutions, at least, in the final population for evaluation
 - Using random strategy to generate the initial population
 - Adopting a lower mutation rate and a lower probability to fix a solution to be valid
 - Preserving valid solutions all along the way might not always be necessary
- **Take-home message:** *Whether or not to preserve invalid solutions deserves more attention in SBSE, and in some cases, we have to preserve valid solutions all along the way during the search.*
- See paper for:
 - Detailed experimental setup and results
 - Threats to validity & Future work

Thanks for your attention!

<https://github.com/jmguo/balanceValidity/>