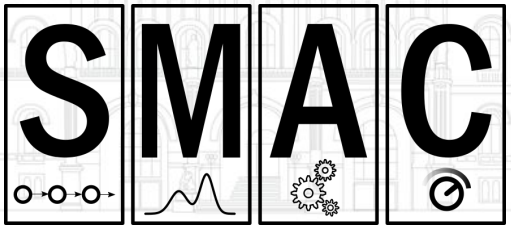
SMAC3: A Versatile Bayesian Optimization Package for Hyperparameter Optimization

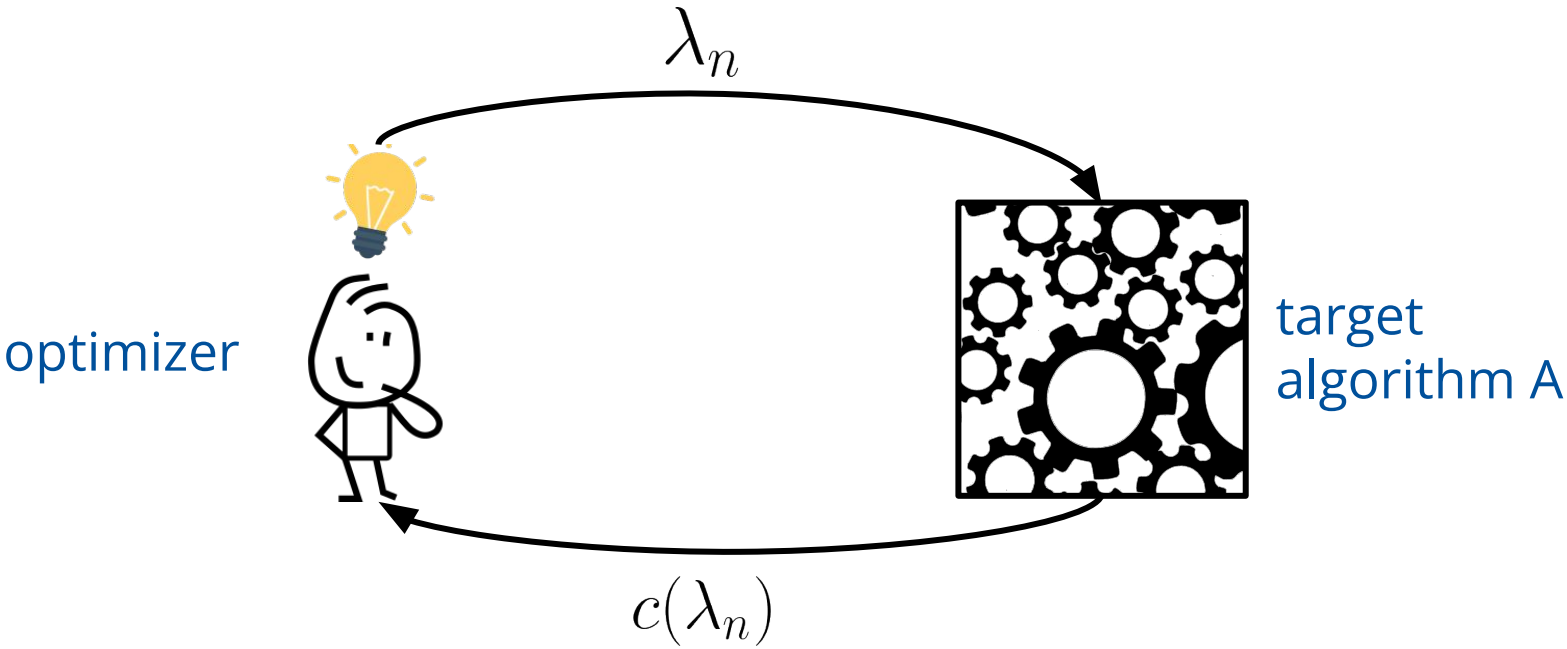
by Marius Lindauer, Katharina Eggensperger, Matthias Feurer, André Biedenkapp, Difan Deng, Carolin Benjamins, Tim Ruhkopf, René Sass & Frank Hutter

New Team Members: Alexander Tornede, Helena Graf, Sarah Segel, Tanja Tornede, Edward Bergmann

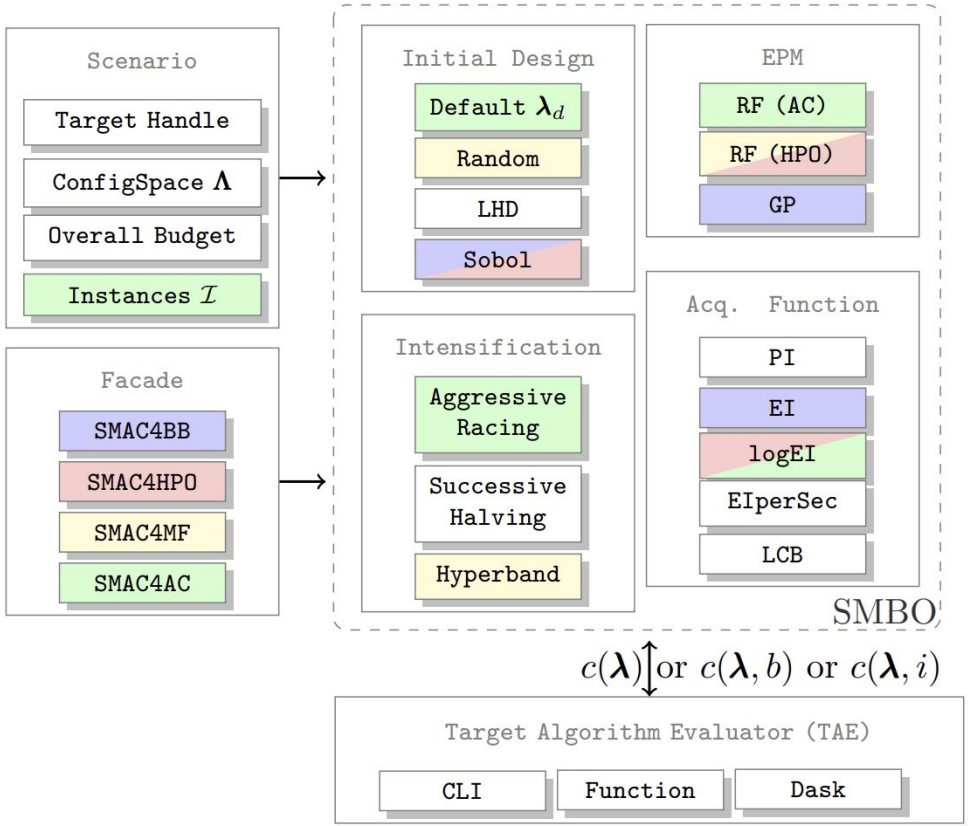


Hyperparameter Optimization

Goal: Find the best performing configuration: $\boldsymbol{\lambda}^* \in \arg\min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} c(\boldsymbol{\lambda}) = \arg\min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} \mathcal{L}(\mathcal{D}_{\text{train}}, \mathcal{D}_{\text{val}}; \boldsymbol{\lambda})$



Modular Design



SMAC for Black-Box Functions

$$\boldsymbol{\lambda}^* \in \arg\min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} c(\boldsymbol{\lambda})$$

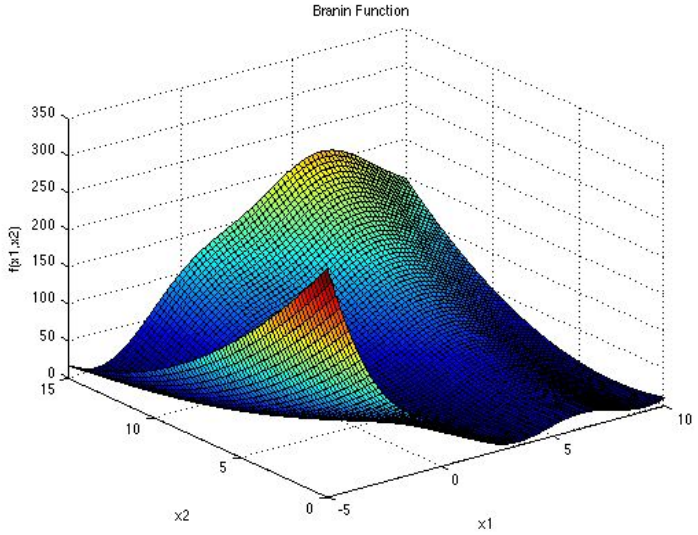
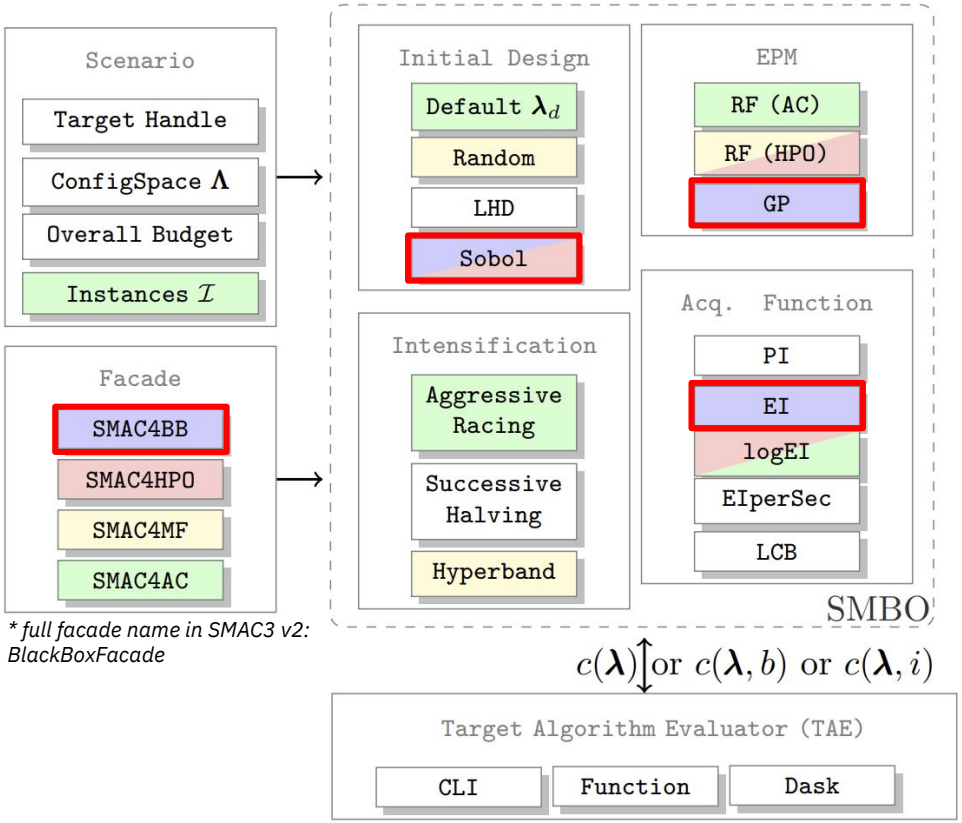


Image credit Derek Bingham

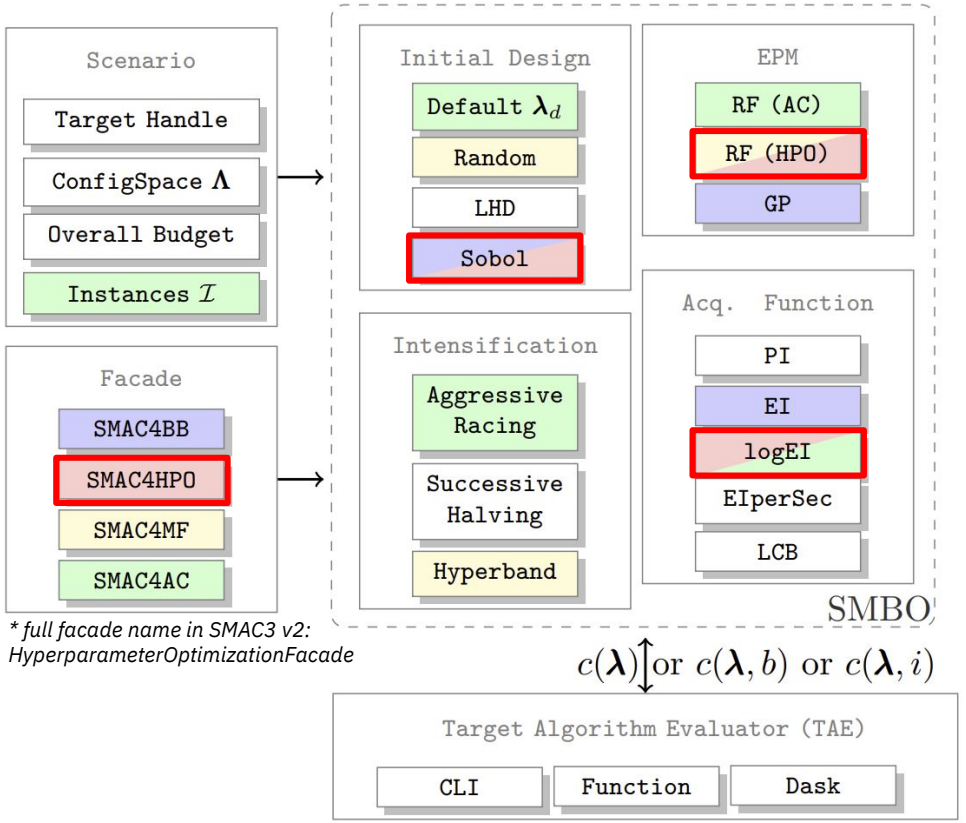


** full facade name in SMAC3 v2: BlackBoxFacade*

SMAC for CASH and Structured Hyperparameter Optimization

$$(A^*, \boldsymbol{\lambda}^*) \in \underset{A_i \in \mathbf{A}, \boldsymbol{\lambda} \in \boldsymbol{\Lambda}_i}{\arg\min}\ c(A_i, \boldsymbol{\lambda}) =$$

$$\underset{A_i \in \mathbf{A}, \boldsymbol{\lambda} \in \boldsymbol{\Lambda}_i}{\arg\min}\ \mathcal{L}(\mathcal{D}_{\text{train}}, \mathcal{D}_{\text{val}}; A_i(\boldsymbol{\lambda})).$$

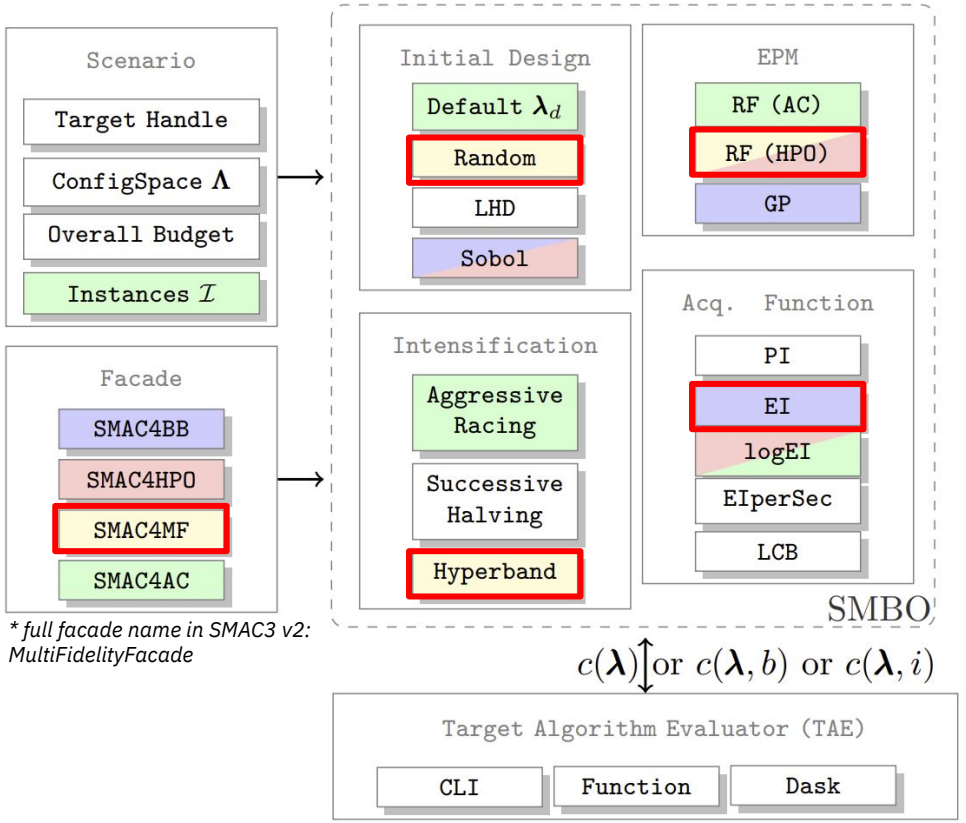


** full facade name in SMAC3 v2: HyperparameterOptimizationFacade*

SMAC for Expensive Tasks and Automated Deep Learning

$$\boldsymbol{\lambda}^* \in \arg\min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} c(\boldsymbol{\lambda}, b_{max}) =$$

$$\arg\min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} \mathcal{L}(\mathcal{D}_{\text{train}}, \mathcal{D}_{\text{val}}; \boldsymbol{\lambda}, b_{max}).$$

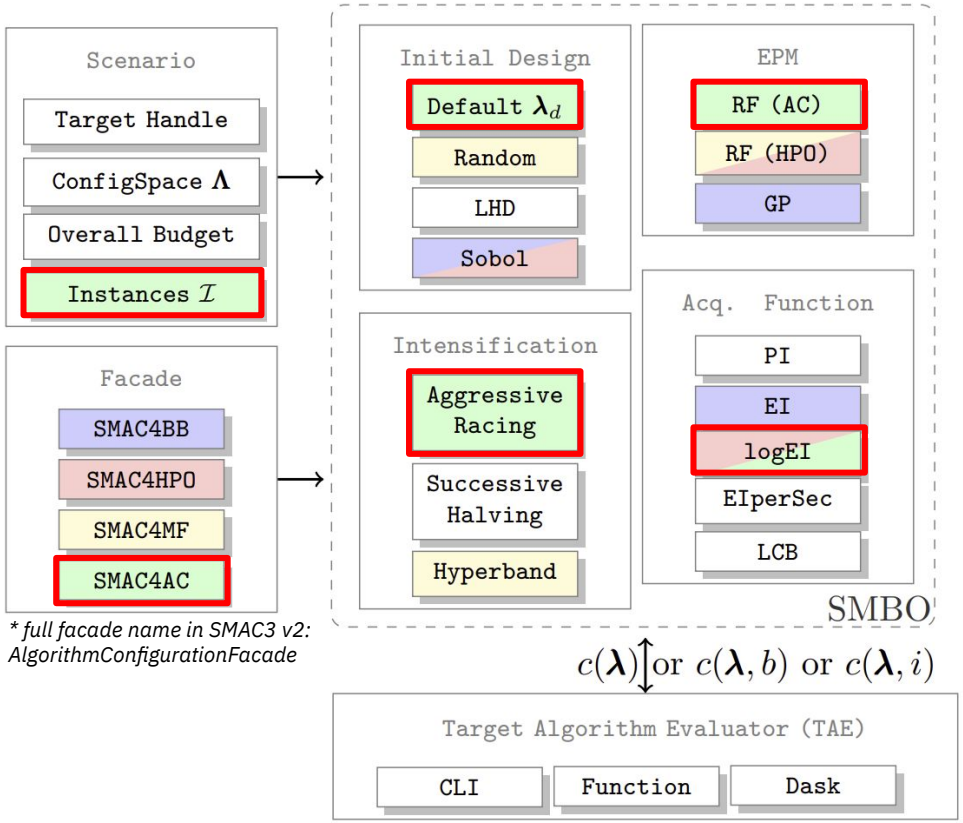


** full facade name in SMAC3 v2: MultiFidelityFacade*

SMAC for Algorithm Configuration

$$\boldsymbol{\lambda}^* \in \arg\min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} c(\boldsymbol{\lambda}) =$$

$$\arg\min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} \sum_{i \in \mathcal{I}} c'(\boldsymbol{\lambda}, i)$$



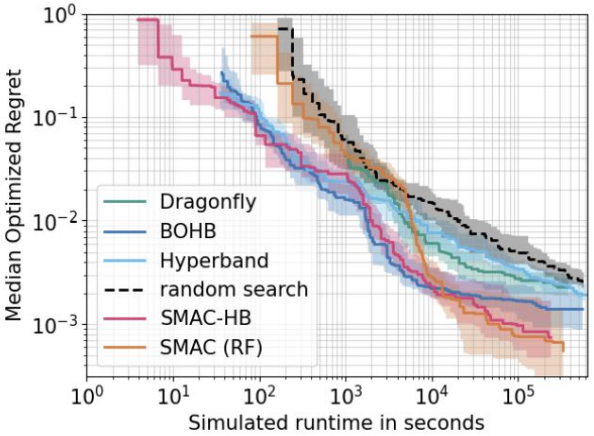
** full facade name in SMAC3 v2: AlgorithmConfigurationFacade*

Comparison to Other Packages

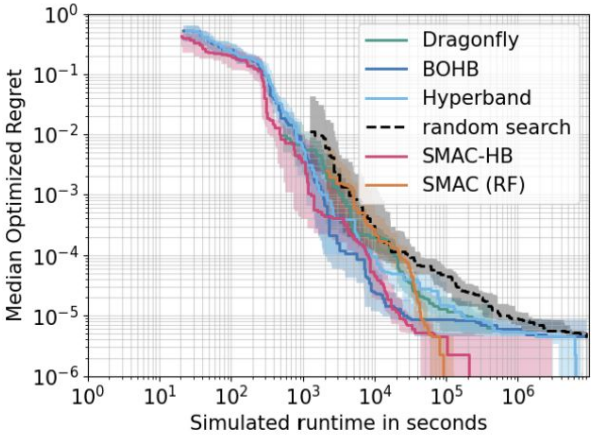
Package	Complex Hyperparameter Space	Multi-Objective	Multi-Fidelity	Instances	Command-Line Interface	Parallelism
HyperMapper	✅	✅	❌	❌	❌	❌
Optuna	✅	✅	✅	❌	✅	✅
Hyperopt	✅	❌	❌	❌	✅	✅
BoTorch	❌	✅	✅	❌	❌	✅
OpenBox	✅	✅	❌	❌	❌	✅
HpBandSter	✅	❌	✅	❌	❌	✅
SMAC	✅	✅	✅	✅	✅	✅

Exemplary Results

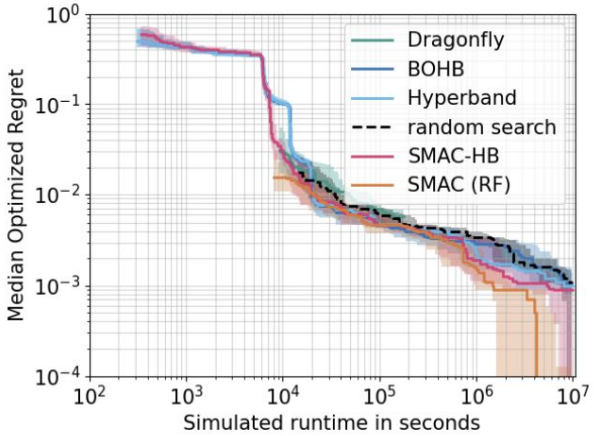
NetLetter (6D)



NBHPONaval (9D)



Nas1Shot1_2



Take-Aways:

1. SMAC with a RF as black-box HPO approach *"SMAC (RF)"* outperforms other approaches with TPE and GP models
2. SMAC's implementation of BOHB [Falkner et al. 2018] *"SMAC-HB" (also using a RF as surrogate) has a very strong any-time performance*

Find Us

AutoML.org

@AutoML_org

automl

@AutoML_org

Original authors:

New team members:

Funded by:

European Research Council
Established by the European Commission