Multi-Objective Coverage Bayesian Optimization (MOCOBO)

Natalie Maus, Kyurae Kim, Yimeng Zeng, Haydn T. Jones, Fangping Wan,
Marcelo Der Torossian Torres, Cesar de la Fuente-Nunez, Jacob R. Gardner

NEURAL INFORMATION PROCESSING SYSTEMS

Penn
University of Pennsylvania

Abstract

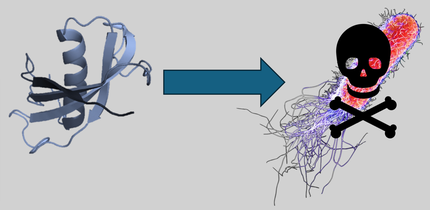
Traditional multi-objective optimization returns Pareto solutions balancing trade-offs—but when objectives conflict heavily, no single solution performs well across all.
Coverage optimization selects a small set of solutions such that each objective is well-addressed (i.e., "covered") by at least one.
We introduce MOCOBO, the first Bayesian optimization method for coverage optimization.

Drug Discovery as Black-Box Optimization

Goal: Find a new peptide that works as a drug against a target bacterial pathogen

Target: harmful bacterial pathogen

Drug: peptide that naturally inhibits target pathogen



Black-box objective function (real-valued):

$f(x)$ = effectiveness of peptide x in inhibiting the target

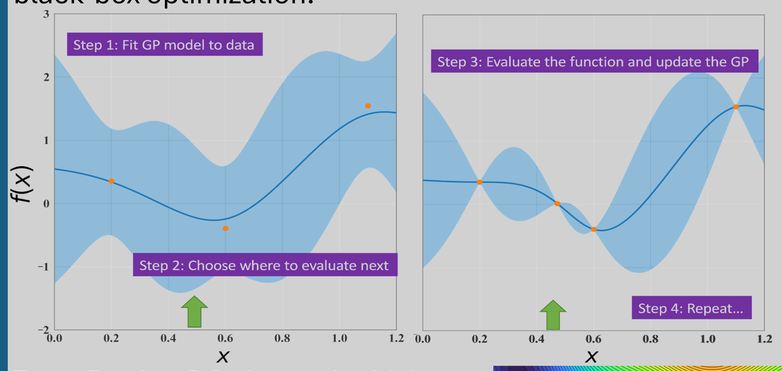
Goal: Find a peptide that optimizes $f(x)$, formally:

$$\underset{x \in \mathcal{X}}{argmax}\, f(x)$$

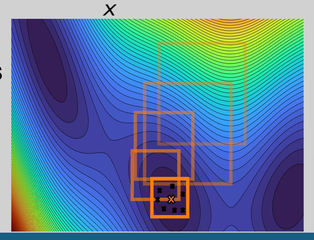
Multi-objective setting: Find a peptide x that optimizes multiple objectives $f_1, f_2, \ldots, f_T$ (e.g., multiple pathogens)

Bayesian Optimization (BO)

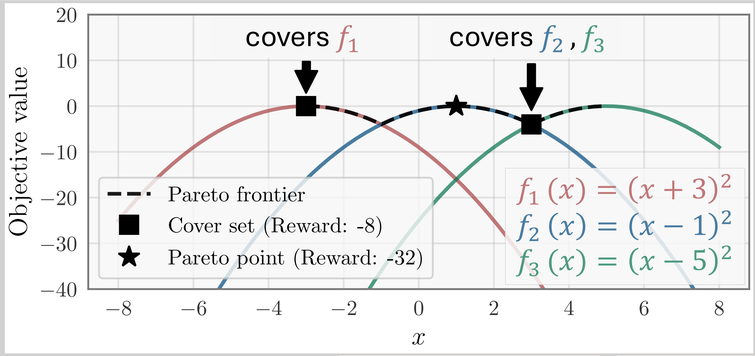
BO: A model-based global optimization technique for black-box optimization.



Trust Region BO: A popular high-dimensional BO method which limits the the large, high-dim search space to a hyper-rectangular trust region τ. The size and location of the trust region are updated dynamically.



Coverage Optimization



Goal: Cover T pathogens with a set of K<T peptides.

Idea: Each target pathogen is inhibited by at least one peptide in the "covering set".

Search Space X (All Peptides)

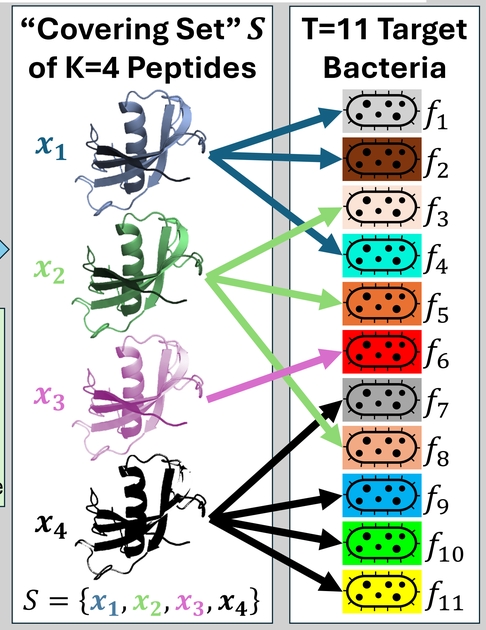
Coverage score:

$$c(S) = \sum_{i=1}^{T} \max_{x \in S} f_i(x)$$

The quantity we seek to optimize

Goal (formally):

$$S^* = \underset{S \subseteq X, |S|=K}{\text{argmax}}\ c(S)$$



Our Method: Multi Objective Coverage BO (MOCOBO)

MOCOBO is a Bayesian optimization algorithm for coverage optimization. To select new candidates for evaluation, MOCOBO introduces ECI, a new acquisition function trailered to the coverage optimization setting.

Expected Coverage Improvement (ECI):

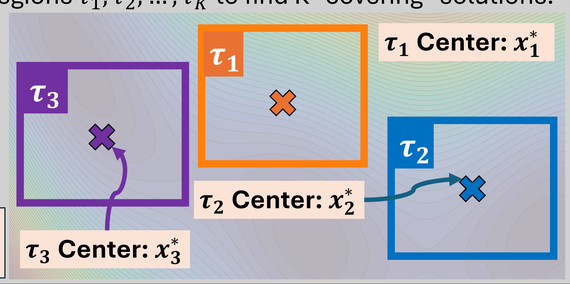
$$ECI(x) = E_{p(y|x,D)}[\max(0, c(S^*_{D \cup \{(x,y)\}}) - c(S^*_D))]$$

$c(S^*_{D \cup \{(x,y)\}})$: Best coverage score after observing (x, y)

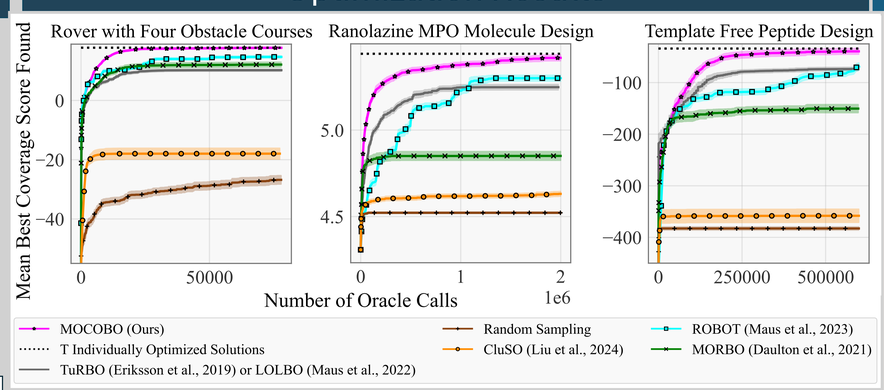
$c(S^*_D)$: Best coverage score so far

MOCOBO uses K trust regions $\tau_1, \tau_2, \ldots, \tau_K$ to find K "covering" solutions. On each step, trust regions are centered on $S^*_D = \{x^*_1, x^*_2, \ldots, x^*_K\}$: the best covering set of K solutions in the observed dataset D.

$$S^*_D = \underset{S \subseteq D, |S|=K}{\text{argmax}}\ c(S)$$

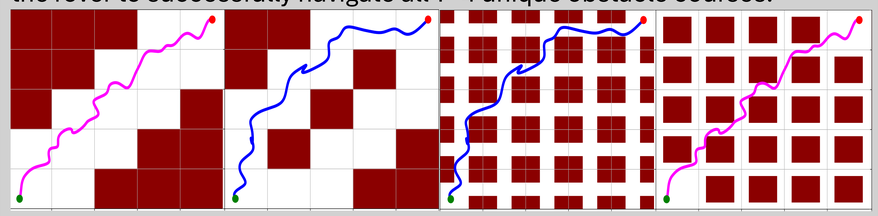


Optimization Results



Rover with Four Obstacle Courses Task

The K=2 trajectories (magenta, blue) found by one run of MOCOBO allow the rover to successfully navigate all T=4 unique obstacle courses.



Ranolazine MPO Molecule Design Task

T=6 objectives: add six target elements to Ranolazine (F, Cl, Br, Se, S, P). K=3 covering molecules found by one run of MOCOBO:

1. CC=C(C)C(OC(=O)C(O)CCCCCCC(=O)O)=CC=CCCCCCC[Se]CC(=O)NC1=CC=CC=C1C
2. CC=C(C)C(OC(=O)CCCCCCC(O)C(=S)Cl)=CC=CCOCCCCC(O)CC(=O)NC1=CC=CC=C1C
3. CC=C(C)C(OC(=O)C(O)CCCCCCC(=O)CBr)=CC=COCPCCCCN(C)CC(=O)[NH1]C1=CC=C(F)C=C1C

Antimicrobial Peptide Design Task

$\{x^*_1, x^*_2, x^*_3, x^*_4\}$: Covering set of K=4 peptides found by one run of MOCOBO
B1-B11: The T=11 target bacteria. (-): Gram-negative (+): Gram-positive
MIC (minimum inhibitory concentration): concentration of peptide needed to inhibit bacterial growth (lower is better)

***In silico* MICs:** MOCOBO successfully covers all T=11 target bacteria.

	B1(-)	B2(-)	B3(-)	B4(-)	B5(-)	B6(-)	B7(-)	B8(+)	B9(+)	B10(+)	B11(+)
x^*_1	1.017	1.040	1.893	**0.999**	8.613	**0.966**	**1.039**	65.999	38.361	338.692	1.393
x^*_2	0.999	15.565	1.860	1.952	404.254	486.860	406.034	**1.233**	**1.318**	**7.359**	**0.981**
x^*_3	2.654	3.268	3.113	4.854	**4.923**	12.967	14.610	22.631	29.685	254.306	3.947
x^*_4	**0.939**	**0.906**	**1.124**	1.310	10.909	1.384	1.711	12.776	32.884	434.193	1.037

***In vitro* MICs:** 9 out of 11 bacterial targets are covered *in vitro*.