Preferential Bayesian Optimization

Javier González, **Zhenwen Dai**, Andreas Damianou, Neil D. Lawrence

@ICML 2017, Sydney, Australia

June 26, 2019

My Colleagues

Javier González

Andreas Damianou

Neil D. Lawrence

Motivation

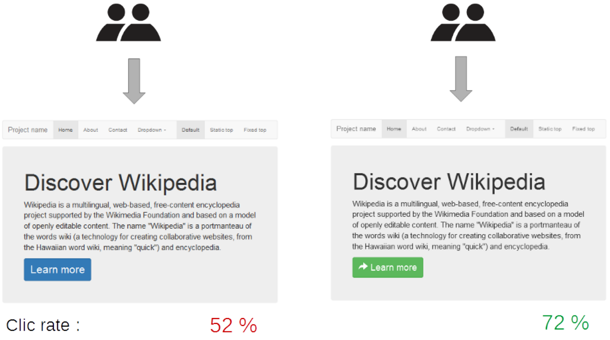
- Bayesian Optimization aims at searching for the global minimum of an expensive function g,

$$\mathbf{x}_{min} = \arg\min_{\mathbf{x} \in \mathcal{X}} g(\mathbf{x}).$$

- What if the function g is not directly measurable?

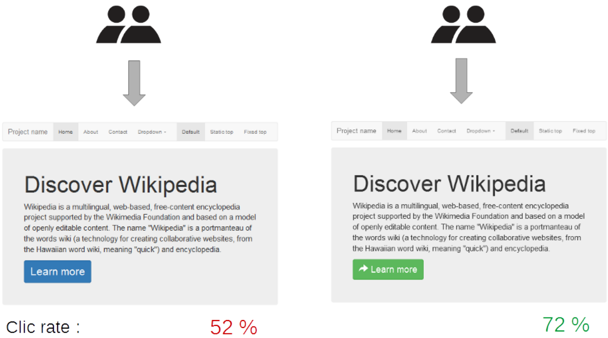
Preference vs. Rating

- The objective function of many tasks are difficult to precisely summarize into a single value.
- Comparison is almost always easier than rating for humans.
- Such observation has been exploited in A/B testing.



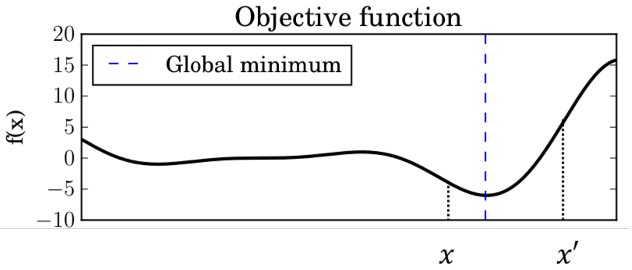
BO via Preference

- Beyond a single A/B testing.
- To optimize a system via tuning this configuration, e.g., the font size, background color of a website.
- The objective such as customer experience is not directly measurable
- Compare the objective with two different configurations.
- The task is to search for the best configuration by iteratively suggesting pairs of configurations and observing the results of comparisons.



Problem Definition

- To find the minimum of a latent function $g(x), x \in \mathcal{X}$.
- Observe only whether $g(\mathbf{x}) < g(\mathbf{x}')$ or not, for a *duel* $[\mathbf{x}, \mathbf{x}'] \in \mathcal{X} \times \mathcal{X}$.
- The outcomes are binary: *true* or *false*.
- The outcomes are *stochastic*.

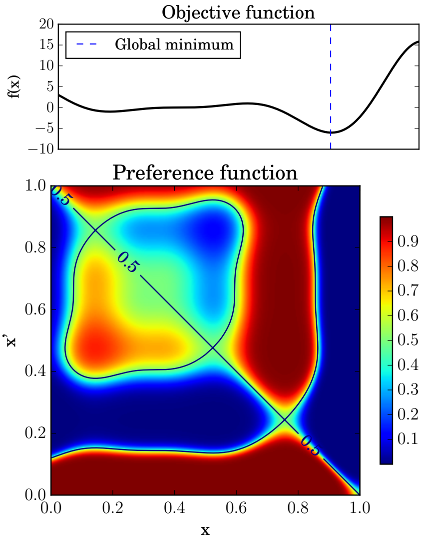


Preference Function

- In this work, the probabilistic distribution is assumed to Bernoulli:

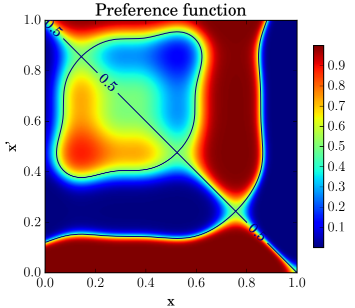
$$p(y \in \{0,1\} | [\mathbf{x}, \mathbf{x}']) = \pi^y (1-\pi)^{1-y},$$
$$\pi = \sigma\Big(g(\mathbf{x}') - g(\mathbf{x})\Big).$$

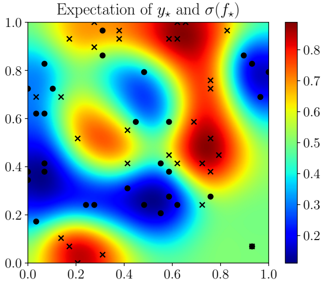
- π is referred to as a *preference function*.
- A Preferential Bayesian optimization algorithm will propose a sequence of *duels* that helps efficiently localize the minimum of a latent function $g(\mathbf{x})$.



A Surrogate Model

- The preference function is not observable.
- Only observe a few comparisons.
- Need a surrogate model to guide the search.
- Two choices:
  - a surrogate model for the *latent* function (like in standard BO). [Brochu, 2010, Guo et al., 2010]
  - a surrogate model for the preference function



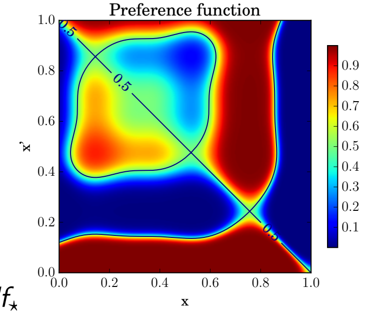


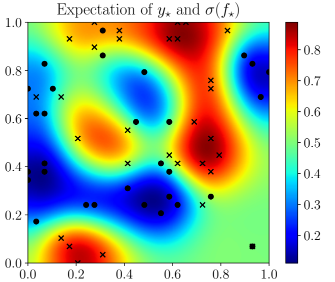
A Surrogate Model of Preference Model

- We propose to build a surrogate model for the preference function.
- Pros: easy to model (Gaussian process Binary Classification is used:)

$$p(y_\star = 1 | \mathcal{D}, [\mathbf{x}, \mathbf{x}'], \theta) = \int \sigma(f_\star) p(f_\star | \mathcal{D}, [\mathbf{x}_\star, \mathbf{x}'_\star], \theta) df_\star$$

- Pros: flexible latent function (e.g., non-stationality).
- Cons: the minimum of the latent function is not directly accessible





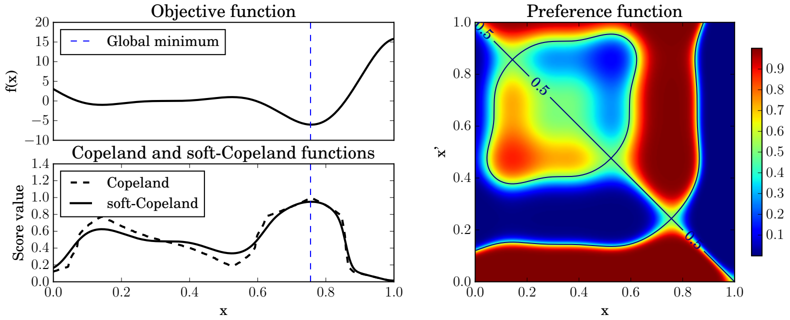
Who is the winner (the minimum)?

- The minimum beats *all* the other locations on average.
- Extending an idea from armed-bandits [Zoghi et al., 2015], we define the *soft-Copeland* score as, (the average winning probability),

$$C(\mathbf{x}) = \mathrm{Vol}(\mathcal{X})^{-1} \int_{\mathcal{X}} \pi_f([\mathbf{x}, \mathbf{x}'])d\mathbf{x}',$$

- The optimum of $g(\mathbf{x})$ can be estimated as, denoted as the *Condorcet* winner,

$$x_c = \arg\max_{\mathbf{x}\in\mathcal{X}} C(\mathbf{x}),$$

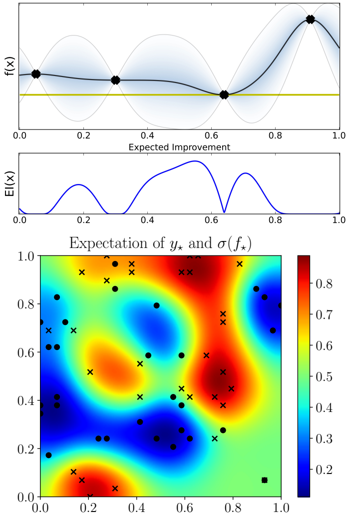


The current estimation of minimum

- Only have a surrogate model of preference function.
- Estimate the *soft-Copeland* score from the surrogate model and get an approximate *Condorcet* winner.
- Note that the approximated *Condorcet* winner may *not* be the optimum of $g(\mathbf{x})$.

Acquisition Function

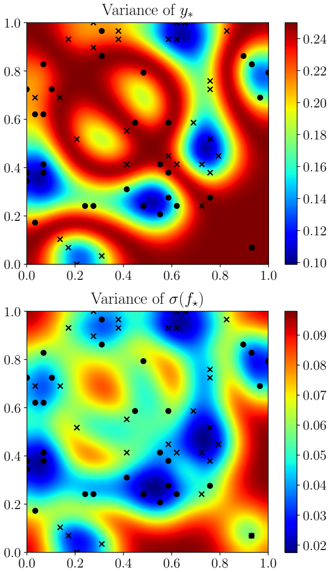
- Existing Acq. Func. are not *applicable*.
- They are designed to work with a surrogate model of the objective function.
- In PBO, the surrogate model does not directly represent the *latent* objective function.
- We need a new Acq. Func. for duels!



Pure Exploration Acquisition Function (PBO-PE)

- The common pure explorative acq. func., *i.e.* $\mathbb{V}[y]$, does not work.
- Propose a pure explorative acq. func. as the variance (uncertainty) of the "winning" probability of a duel:

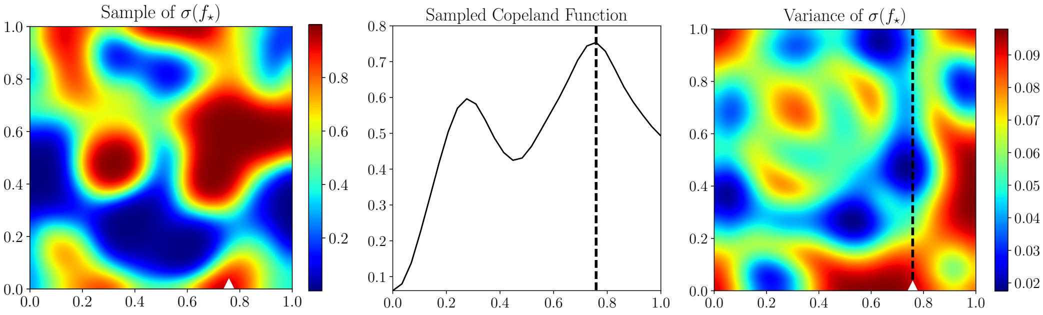
$$\mathbb{V}[\sigma(f_\star)] = \int \left(\sigma(f_\star) - \mathbb{E}[\sigma(f_\star)]\right)^2 p(f_\star|\mathcal{D}, [\mathbf{x}, \mathbf{x}'])df_\star$$



Acquisition Function: PBO-DTS

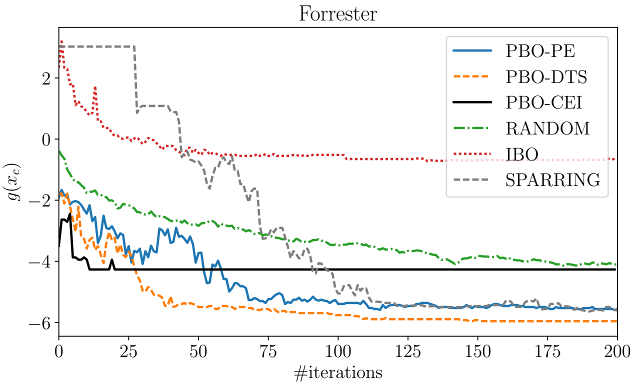
To select the next duel $[\mathbf{x}_{next}, \mathbf{x}'_{next}]$:

1. Draw a sample from surrogate model
2. Take the maximum of *soft-Copeland* score as $\mathbf{x}_{next}$.
3. Take $\mathbf{x}'_{next}$ that gives the maximum in PBO-PE



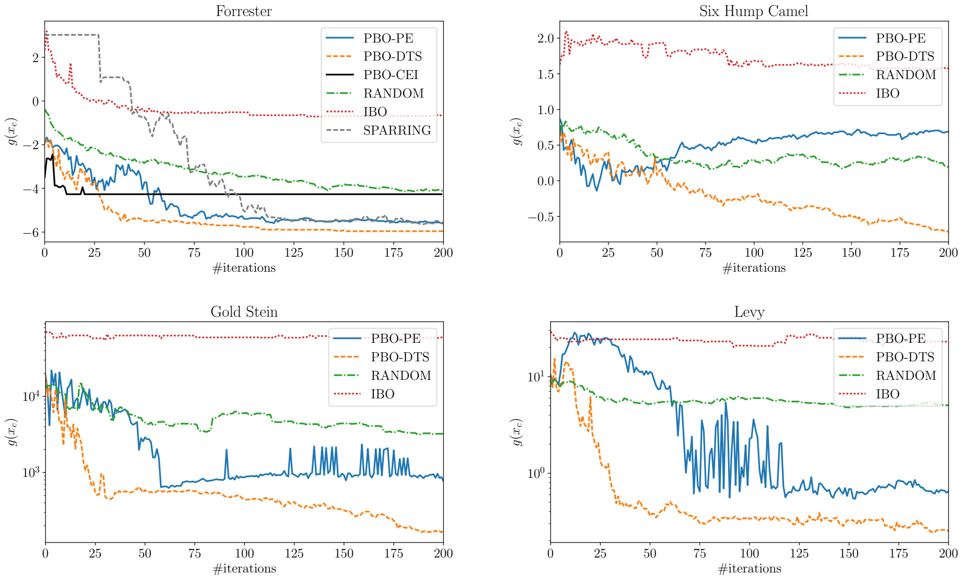
Experiment: Forrester Function

- Synthetic 1D function: Forrester
- Observations drawn with a probability: $\frac{1}{1+e^{g(\mathbf{x})-g(\mathbf{x}')}}$
- $g(x_c)$ shows the value at the location that algorithms *believe* is the minimum.
- The curve is the average of 20 trials.



IBO: [Brochu, 2010]
SPARRING: [Ailon et al., 2014]

Experiments: More (2D) Functions



Summary

- Address Bayesian optimization with preferential returns.
- Propose to build a surrogate model for the preference function.
- Propose a few efficient acquisition functions.
- Show the performance on synthetic functions.

Questions?

Exploration & Exploitation

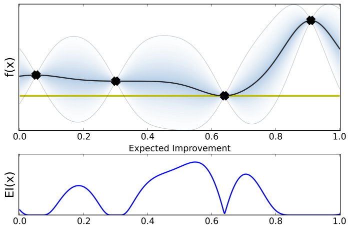
The two ingredients in an acquisition function: Exploration & Exploitation.

Exploration in PBO

- To understand exploration in PBO by designing a *pure explorative* acq. func.
- Exploration in standard BO can be viewed as the action to reduce uncertainty of a surrogate model.
- A purely explorative acq. func.

$$\mathbb{V}[y_\star] = \int (y_\star - \mathbb{E}[y_\star])^2 \, p(y_\star | \mathcal{D}, \mathbf{x}_\star) \mathrm{d}y_\star$$

- Can we extend this idea to PBO?

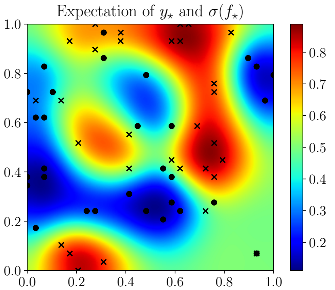


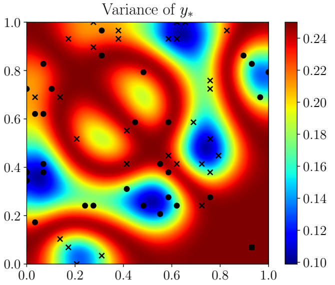
A Straight-Forward Choice

- A straight-forward extension from standard BO:

$$\mathbb{V}[y_\star] = \sum_{y_\star \in \{0,1\}} (y_\star - \mathbb{E}[y_\star])^2 \, p(y_\star | \mathcal{D}, [\mathbf{x}_\star, \mathbf{x}'_\star])$$
$$= \mathbb{E}[y_\star](1 - \mathbb{E}[y_\star])$$

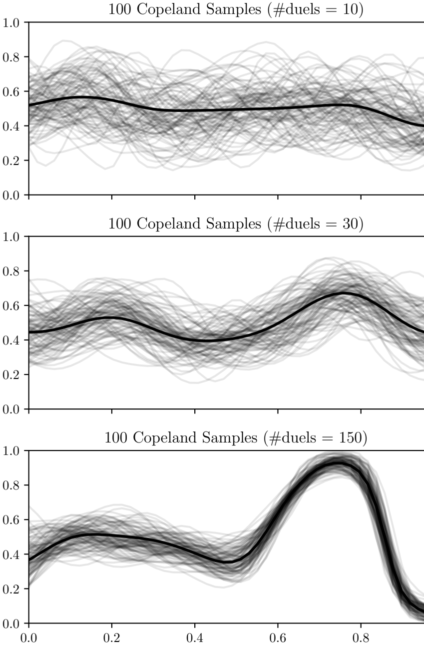
- The maximum variance is always at where $\mathbb{E}[y_\star] = 0.5$!
- The variance may not reduce with observations!





Dueling-Thompson Sampling (DTS)

- To balance exploration & exploitation, we borrow the idea of Thompson sampling by drawing a sample from the surrogate model.
- Compute the *soft-copeland* score on the drawn sample.
- The value $\mathbf{x}_{next}$ that gives the maximum *soft-copeland* score gives a good balance between exploration and exploitation.
- Take it as the *first* value of the next duel.



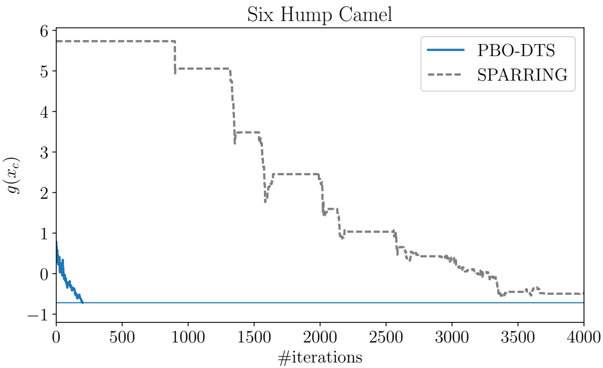
Aleatoric Uncertainty & Epistemic Uncertainty

- The uncertainty of $y_\star$ comes from two sources: the *aleatoric uncertainty* $\sigma(f_\star)$ and the *epistemic uncertainty* $p(f_\star|\mathcal{D}, [\mathbf{x}_\star, \mathbf{x}'_\star], \theta)$

$$p(y_\star = 1|\mathcal{D}, [\mathbf{x}, \mathbf{x}'], \theta) = \int \sigma(f_\star) p(f_\star|\mathcal{D}, [\mathbf{x}_\star, \mathbf{x}'_\star], \theta) df_\star$$

- Aleatoric Uncertainty: the stochasticity of the underlying process
- Epistemic Uncertainty: the uncertainty due to limited observations
- Exploration should focus on *epistemic uncertainty*.

Multi-arm Bandits on 2D



Nir Ailon, Zohar Shay Karnin, and Thorsten Joachims. Reducing dueling bandits to cardinal bandits. In *Proceedings of the 31th International Conference on Machine Learning, ICML 2014, Beijing, China, 21-26 June 2014*, pages 856–864, 2014.

Eric Brochu. *Interactive Bayesian Optimization: Learning Parameters for Graphics and Animation*. PhD thesis, University of British Columbia, Vancouver, Canada, December 2010.

Shengbo Guo, Scott Sanner, and Edwin V Bonilla. Gaussian process preference elicitation. In *Advances in Neural Information Processing Systems 23*, pages 262–270, 2010.

Masrour Zoghi, Zohar S Karnin, Shimon Whiteson, and Maarten de Rijke. Copeland dueling bandits. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 307–315. Curran Associates, Inc., 2015.