Cost-Aware Bayesian Optimization with Adaptive Stopping via Gittins Indices

Qian Xie 谢倩 (Cornell ORIE)

INFORMS Annual Meeting 2025 Job Market Showcase

About Me - Background

- Education:

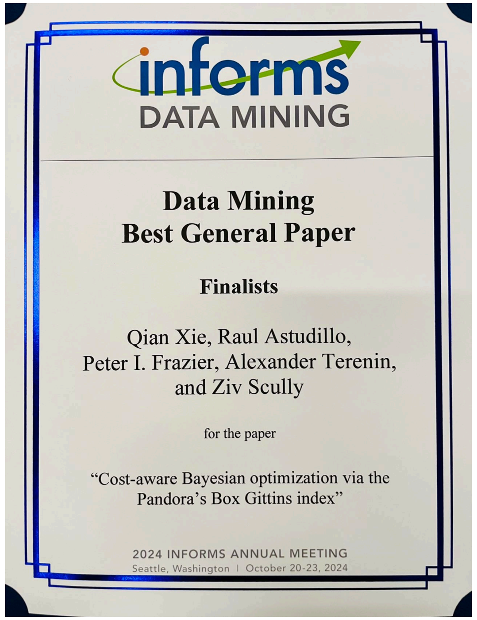
  Tsinghua (Yao Class) ⟶ NYU ⟶ Cornell
  Dissertation working title: *Gittins Indices for Bayesian Optimization*

- Publication:

  Top ML conferences: NeurIPS, ICML, …
  Top journals: OR (major revision), Automatica, IEEE TCNS, ...

- Selected award:

  INFORMS Data Mining Best Paper Competition Finalist



About Me – Research Interests

Theoretical Foundation

- Decision theory
  (incl. Gittins index theory)
- Dynamic programming & MDPs
- Stochastic control
- Bayesian inference

Methodology

- Bayesian optimization
- Active learning
- Reinforcement learning
  (incl. bandits)
- LLM-as-agent

LLM development

- Efficient LLM evaluation (ongoing)
- LLM reasoning (future)

Adaptive experimentation

- Online A/B testing (future)
- Dynamic pricing (future)

Transportation

- Mixed-autonomy traffic control (ongoing)

Scientific discovery

- Drug cocktail discovery (ongoing)
- Fusion reactor design (future)

About Me – PhD Research Projects

- **Data-efficient Black-box Optimization** (Recent)
  - Bayesian optimization via Gittins indices
    [NeurIPS’24 & INFORMS DM Finalist, ICLR’26 (under review), ICML’26 (in prep)]

  - LLM-driven neural architecture search for RL training
    [NeurIPS’25 LAW workshop]

About Me – PhD Research Projects

- Data-efficient Black-box Optimization (Recent)
  - Bayesian optimization via Gittins indices
    [NeurIPS’24 & INFORMS DM Finalist, ICLR’26 (under review), ICML’26 (in prep)]
  - LLM-driven neural architecture search for RL
    [NeurIPS’25 LAW workshop]
- Interactive Online Decision-making (Earlier)
  - Online recommendation (bandits)
    [ICML’23 & OR (major revision)]
  - Online resource allocation (MDP & stochastic game)
    [Automatica (2024)]

About Me – PhD Research Projects

In this presentation

- Data-efficient Black-box Optimization (Recent)
  - Bayesian optimization via Gittins indices

    [NeurIPS'24 & INFORMS DM Finalist, ICLR'26 (under review), ICML'26 (in prep)]

  - LLM-driven neural architecture search for RL

    [NeurIPS'25 LAW workshop]

- Interactive Online Decision-making (Earlier)
  - Online recommendation (bandits)

    [ICML'23 & OR (major revision)]

  - Online resource allocation (MDP & stochastic game)

    [Automatica (2024)]

Black-Box Optimization

Input x $\longrightarrow$  $\longrightarrow$ Performance metric $f(x)$

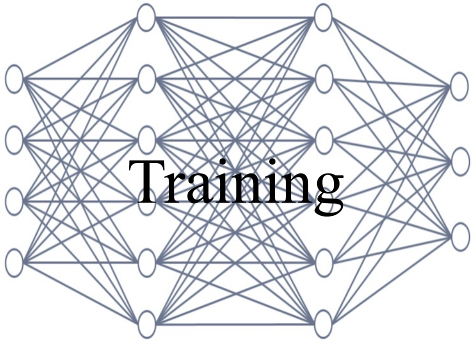
Black-Box Optimization

Input x $\longrightarrow$

$\longrightarrow$ Performance metric $f(x)$

ML model training:

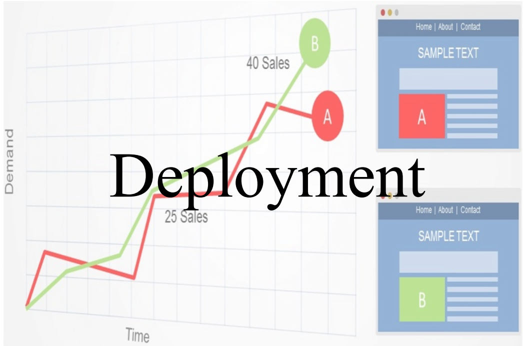
Training hyperparameters $\longrightarrow$



Training $\longrightarrow$ Accuracy

Adaptive experimentation:

Decision/design variables $\longrightarrow$



Deployment $\longrightarrow$ Revenue

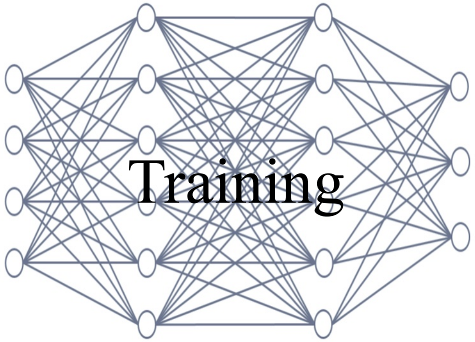
Black-Box Optimization

Input x

Performance metric $f(x)$

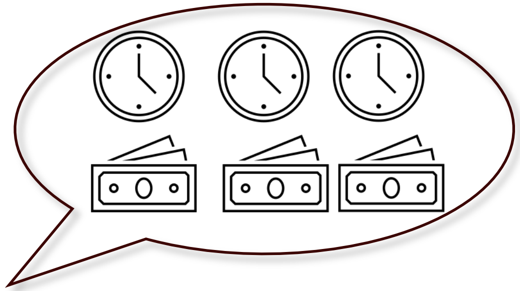
ML model training:

Training hyperparameters



Training

Accuracy

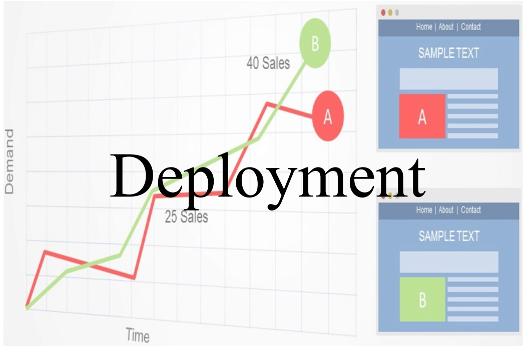


Training time

Compute credits

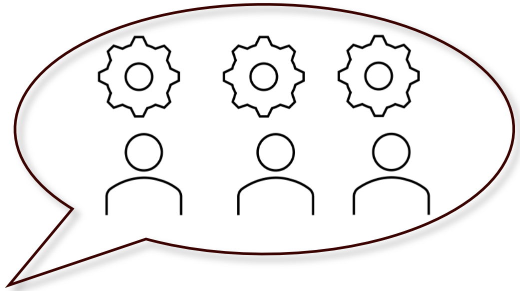
Adaptive experimentation:

Decision/design variables



Deployment

Revenue



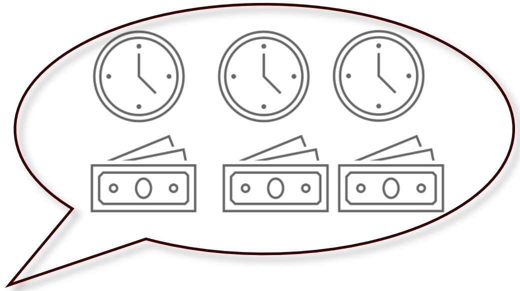
Operational cost

User experience

Black-Box Optimization

expensive-to-evaluate

Input x $\longrightarrow$ $\longrightarrow$ Performance metric $f(x)$

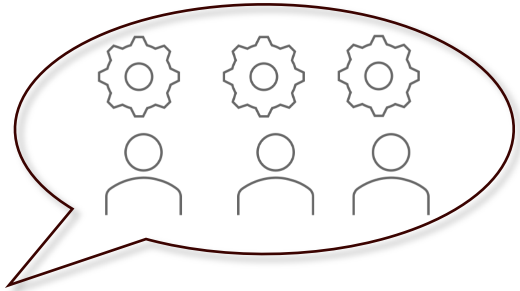


Training time

Compute credits

ML model training:

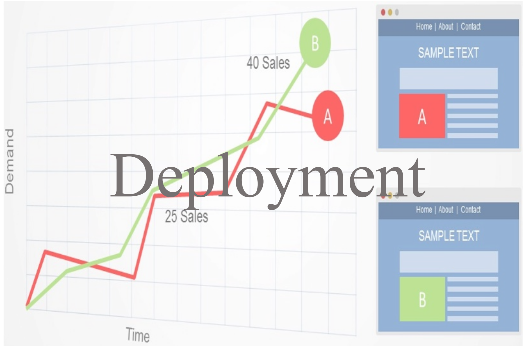
Training hyperparameters $\longrightarrow$ Training $\longrightarrow$ Accuracy



Operational cost

User experience

Adaptive experimentation:



Decision/design variables $\longrightarrow$ Deployment $\longrightarrow$ Revenue

Black-Box Optimization

expensive-to-evaluate

Input x $\longrightarrow$  $\longrightarrow$ Performance metric $f(x)$

High-level goal: Choose $x_1, \ldots, x_T$ to maximize the expected best observed value $\mathbb{E} \max_{t=1,2,\ldots,T} f(x_t)$

Black-Box Optimization

expensive-to-evaluate

Input x $\longrightarrow$  $\longrightarrow$ Performance metric $f(x)$

adaptively

High-level goal: Choose $x_1, \ldots, x_T$ to maximize the expected best observed value $\mathbb{E} \max_{t=1,2,\ldots,T} f(x_t)$

Fewer #evaluations

Black-Box Optimization

expensive-to-evaluate

Input x $\longrightarrow$  $\longrightarrow$ Performance metric $f(x)$

adaptively

High-level goal: Choose $x_1, \ldots, x_T$ to maximize the expected best observed value $\mathbb{E} \max_{t=1,2,\ldots,T} f(x_t)$

Fewer #evaluations

Efficient framework: Bayesian optimization

Bayesian Optimization

$x \longrightarrow$  $\longrightarrow f(x)$

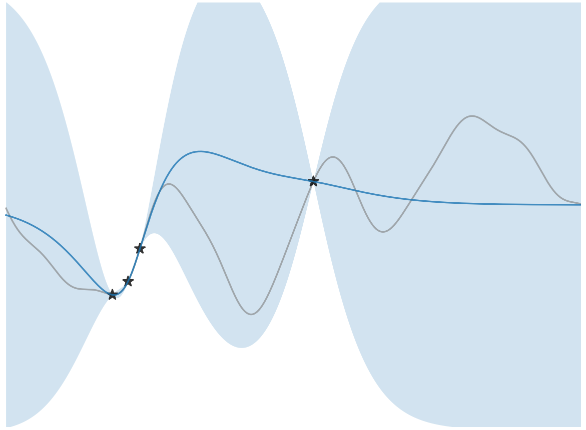
Black-box function

Bayesian Optimization

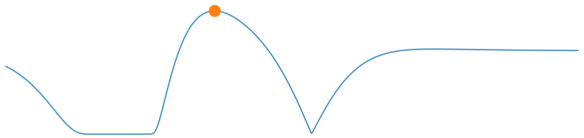
x $\longrightarrow$  $\longrightarrow$ $f(x)$

Black-box function

Update belief
(Bayes' rule)



Probabilistic model
(e.g., Gaussian process)



Decision rule
(e.g., UCB, TS)

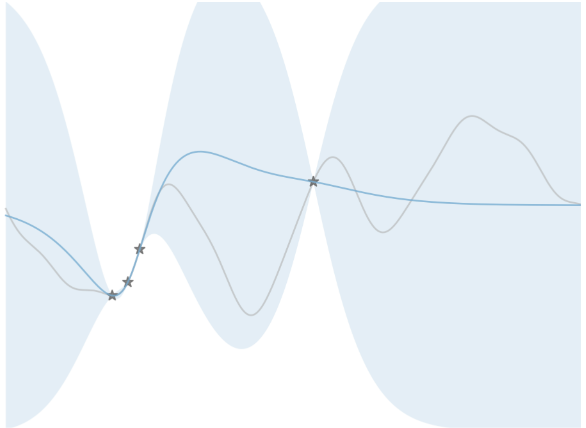
Select next input
(gradient-based
optimization)

Bayesian Optimization

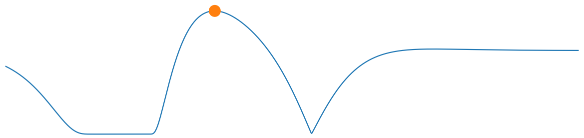
x →  → $f(x)$

Black-box function

Update belief
(Bayes' rule)



Probabilistic model
(e.g., Gaussian process)



Decision rule
(e.g., UCB, TS)

My focus

Select next input
(gradient-based
optimization)

Existing Design Principles

- Improvement-based
- Entropy-based
- Confidence bounds (UCB/LCB)
- Thompson sampling

New Design Principle: Gittins Index

- Improvement-based
- Entropy-based
- Confidence bounds (UCB/LCB)
- Thompson sampling
- Gittins Index

New Design Principle: Gittins Index

- Improvement-based
- Entropy-based
- Confidence bounds (UCB/LCB)
- Thompson sampling
- **Gittins Index**

? Why another principle?

New Design Principle: Gittins Index

- Improvement-based
- Entropy-based
- Confidence bounds (UCB/LCB)
- Thompson sampling
- **Gittins Index**

? Why another principle?

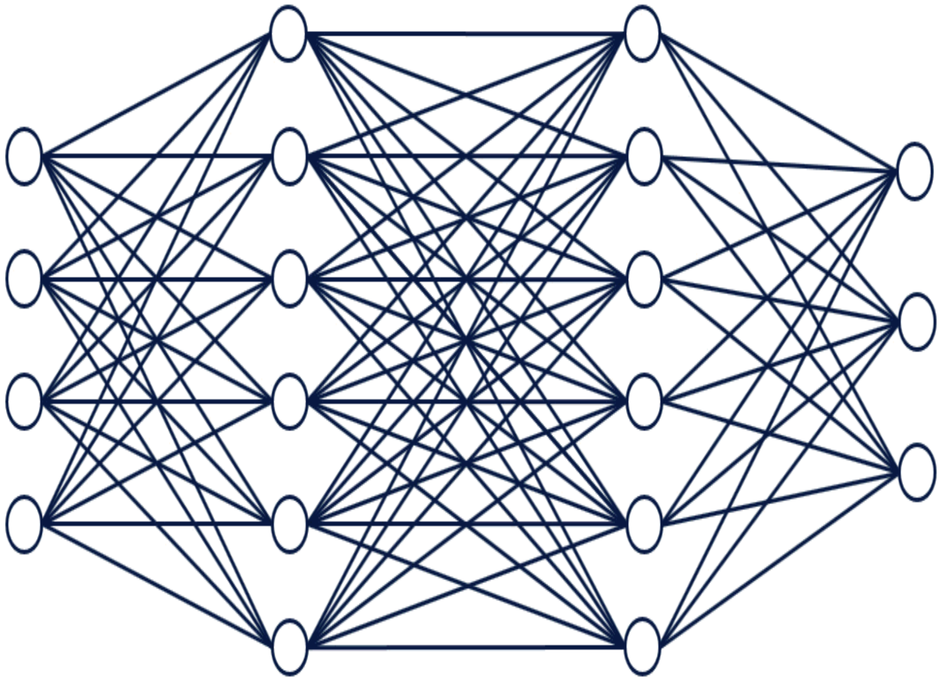
1. Naturally handles practical considerations
2. Performs competitively on benchmarks
3. Comes with theoretical guarantees

Under-explored Practical Considerations

Varying evaluation costs

Smart stopping time

Observable multi-stage feedback



Under-explored Practical Considerations

Varying evaluation costs

Smart stopping time

Observable multi-stage feedback

New design principle:
Gittins index

Why Gittins index?

Varying evaluation costs

Smart stopping time

Observable multi-stage feedback

New design principle:
Gittins index

Why Gittins index?

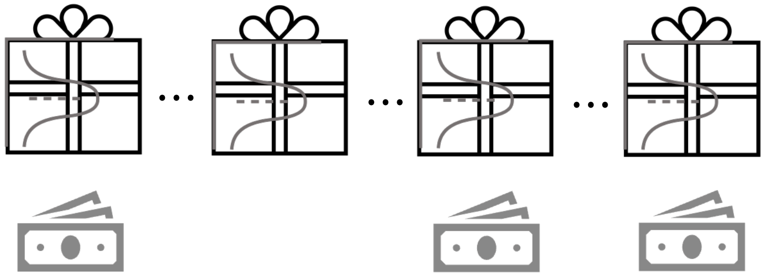
Varying evaluation costs

Smart stopping time

Observable multi-stage feedback

New design principle:
Gittins index

Optimal in related sequential decision problems



Why Gittins index?

Varying evaluation costs

Features in Pandora's box

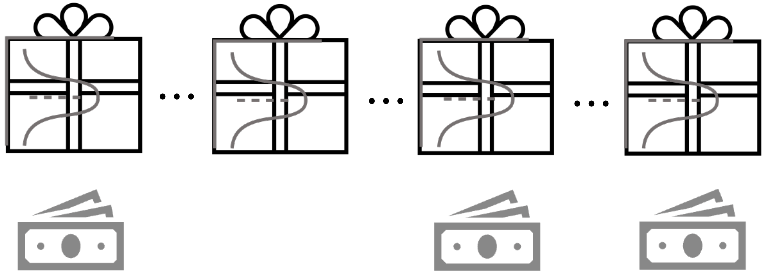
Smart stopping time

Features in Pandora's box

Observable multi-stage feedback

New design principle:
Gittins index

Optimal in related sequential decision problems



Why Gittins index?

Varying evaluation costs

Features in Pandora’s box

Smart stopping time

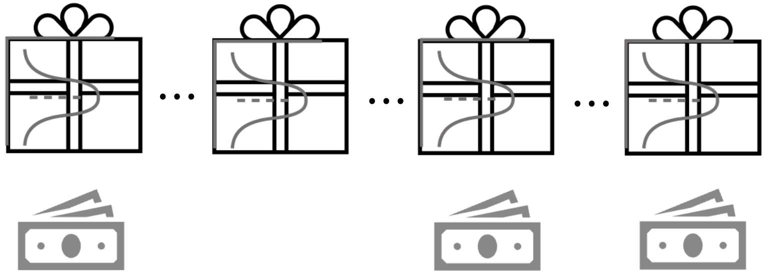
Features in Pandora’s box

Observable multi-stage feedback

Features in Markovian bandits

New design principle:
Gittins index

Optimal in related sequential decision problems



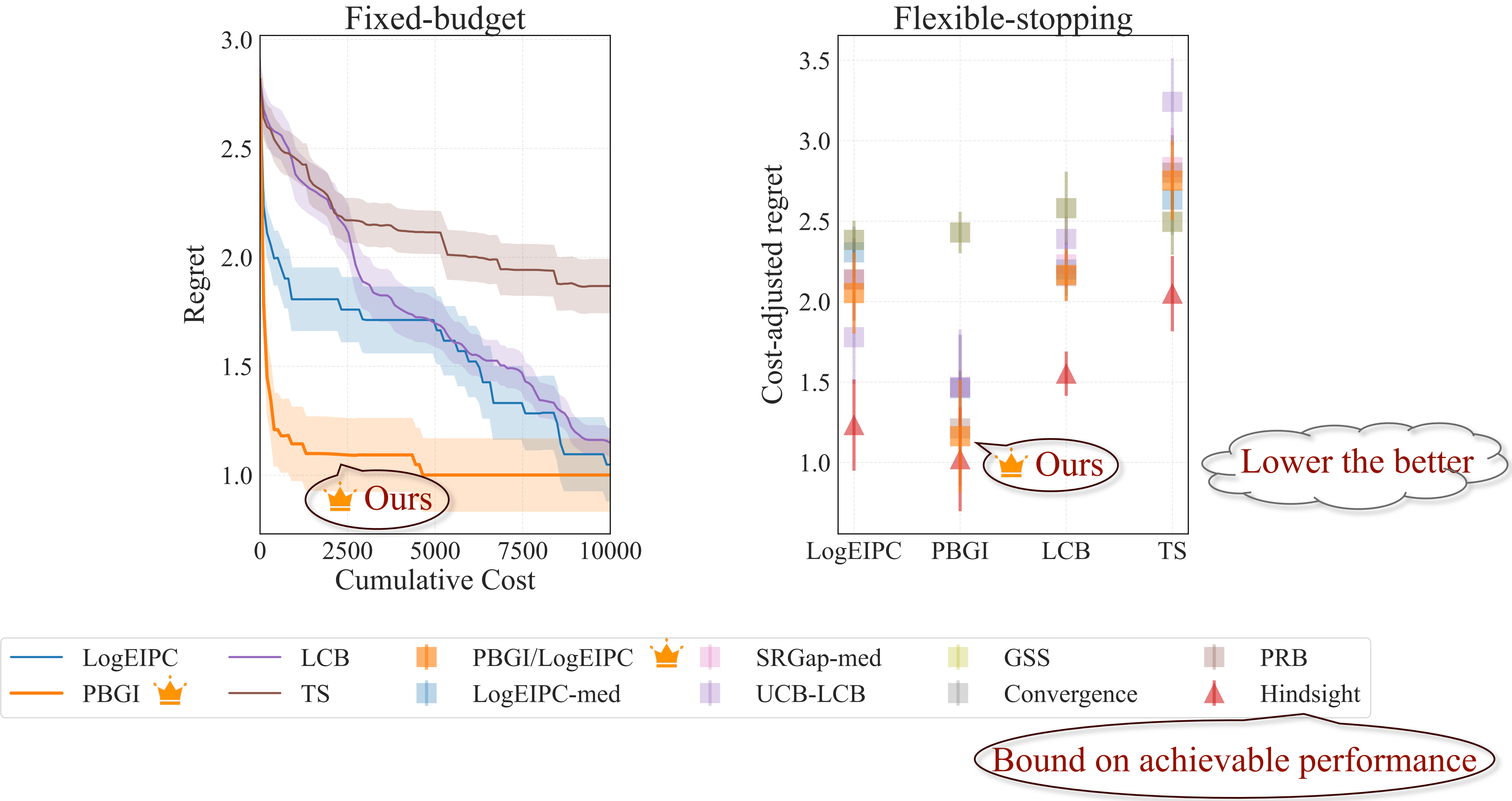
New Design Principle: Gittins Index

- Improvement-based (e.g., LogEIPC)
- Entropy-based
- Confidence bounds (UCB/LCB)
- Thompson sampling (TS)
- Gittins Index (PBGI)

? Why another principle?

1. Naturally handles practical considerations
2. Performs competitively on benchmarks
3. Comes with theoretical guarantees

Gittins Index vs Baselines on AutoML Benchmark



New Design Principle: Gittins Index

- Improvement-based (e.g., LogEIPC)
- Entropy-based
- Confidence bounds
- Thompson sampling
- **Gittins Index**

? Why another principle?

1. Naturally handles practical considerations
2. Performs competitively on benchmarks
3. Comes with theoretical guarantees

Theoretical Guarantee and Empirical Validation

Theorem (No worse than stopping-immediately)

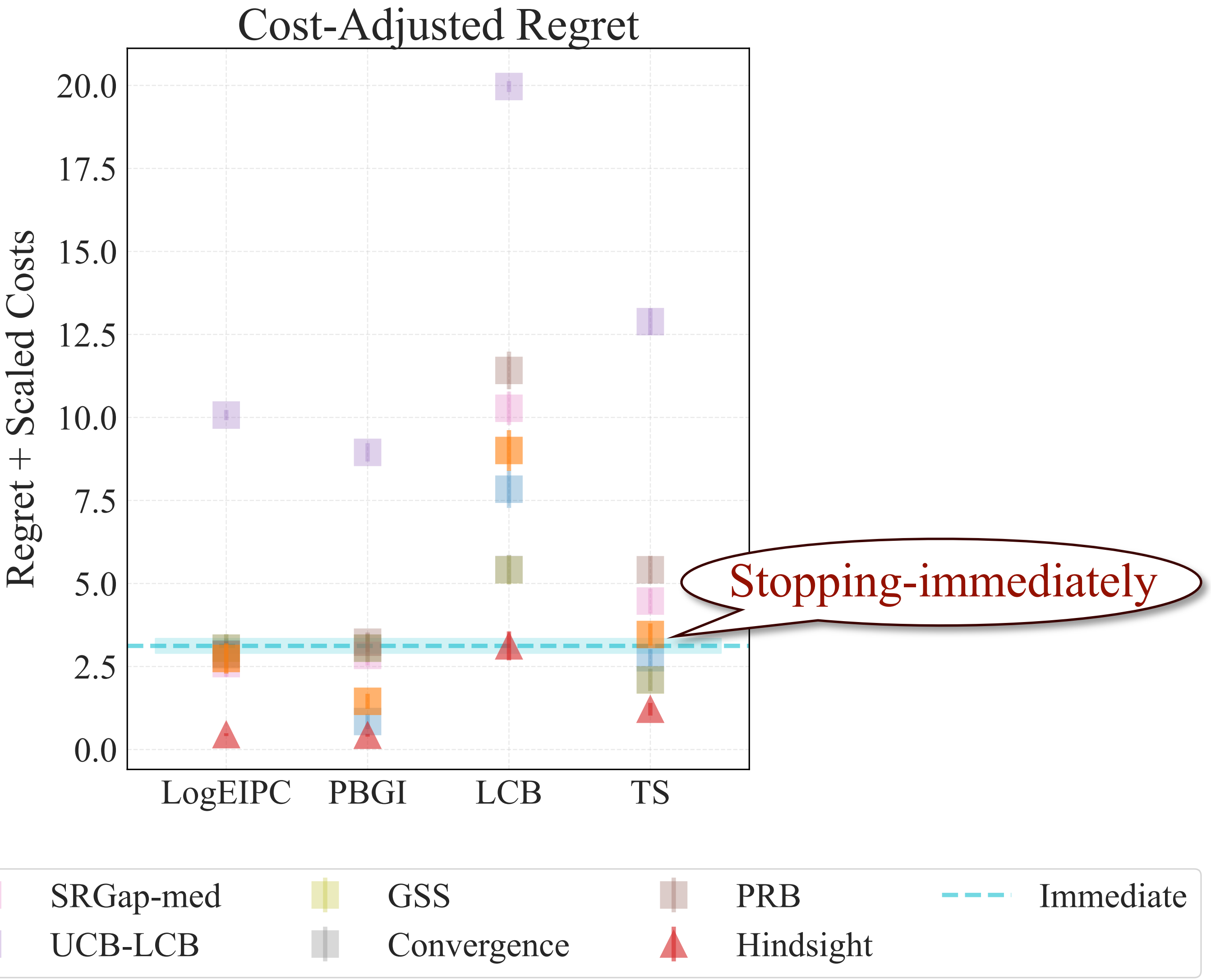
$\mathbb{E}[R(\text{ours}; \text{PBGI})] \leq R[\text{stopping immediately}]$

or LogEIPC   cost-adjusted regret

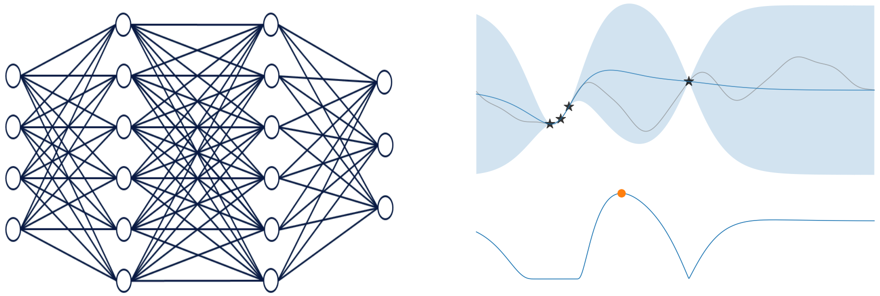
Implication:

- Matches the best achievable performance in the worst case (evaluations are all very costly).

- Avoids over-spending — a property many cost-unaware stopping rules lack.



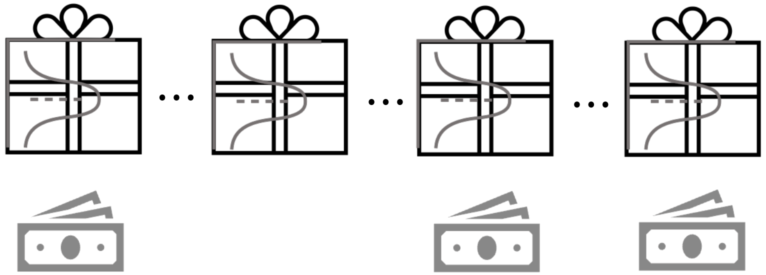
Studied problem



Varying evaluation costs

Adaptive stopping time

Key idea



Link to Pandora's Box problem & Gittins index theory

Impact

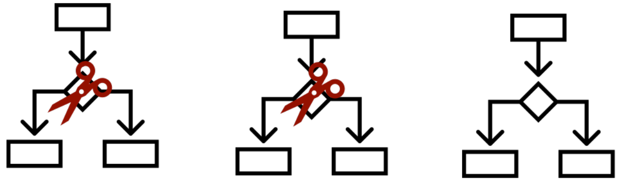
Meta

BoTorch Ax

Competitive empirical performance & interests from practitioners

"Cost-aware Bayesian Optimization via the Pandora's Box Gittins Index." NeurIPS'24.

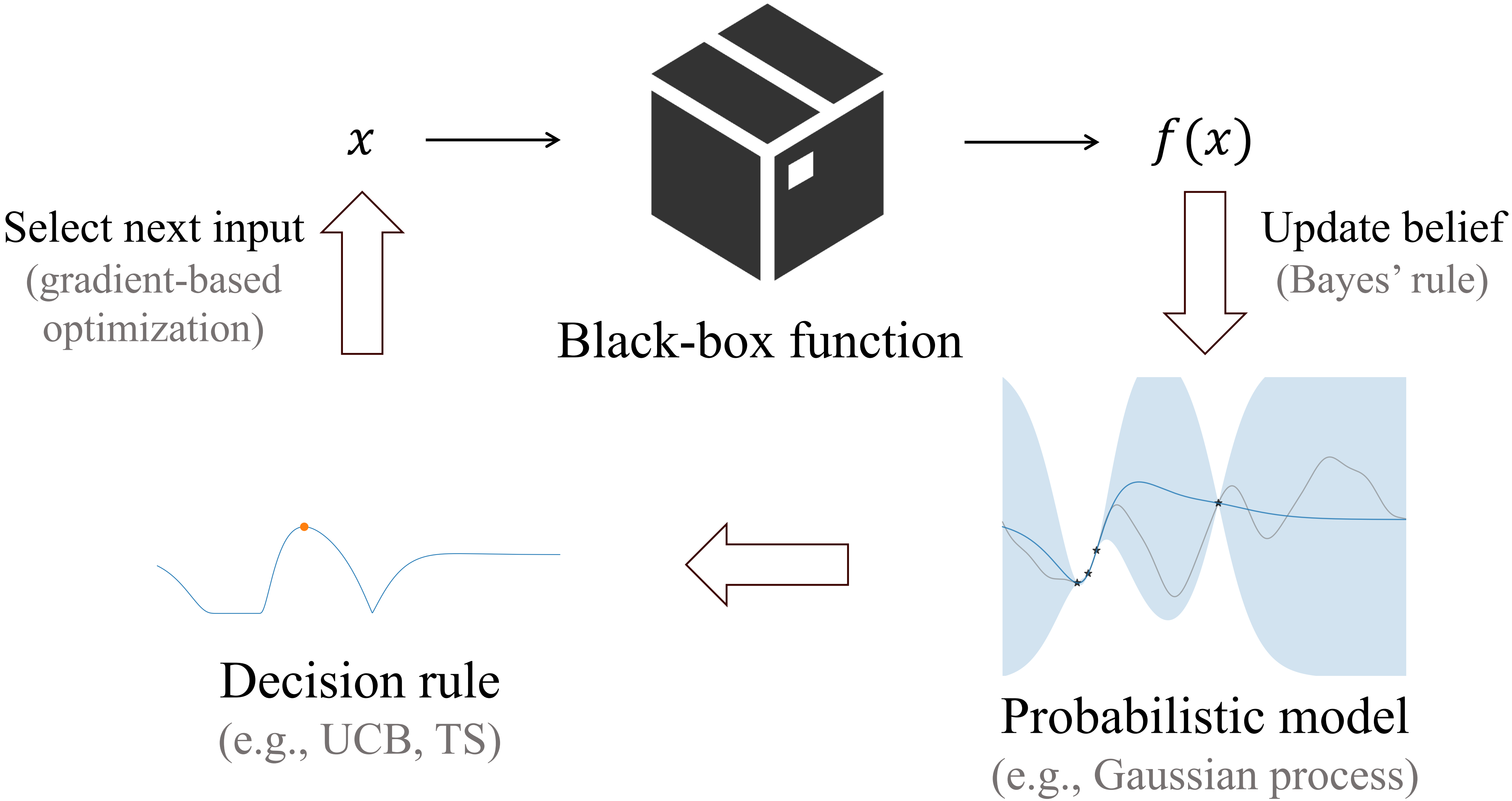
Ongoing work



Sharper theoretical guarantees & black-box optimization w/ multi-stage feedback

"Cost-aware Stopping for Bayesian Optimization." Under review.

Recap: Bayesian Optimization



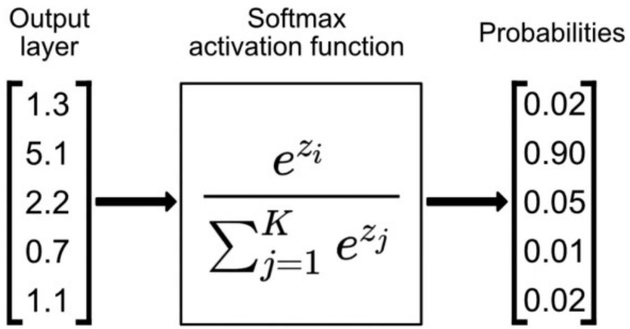
Ongoing: LLM-Driven Black-Box Optimization

$x \longrightarrow$  $\longrightarrow f(x)$

Black-box function

Select next input
(LLM proposal)

Update belief
(feedback-augmented context)



Decision rule
(e.g., Softmax sampling)

ChatGPT Gemini
deepseek Claude

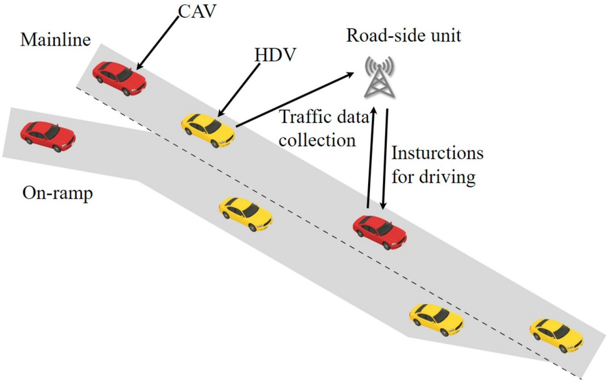
Probabilistic model
(large language model)

Ongoing: LLM-Driven RL Training Optimization

Mixed-autonomy traffic control:

(e.g., Transformer config)

x



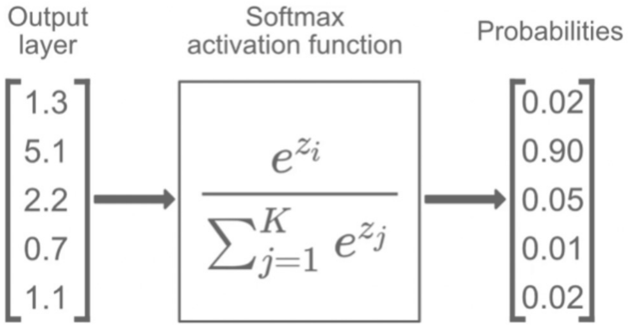
(e.g., average speed)

$f(x)$

Select next input
(LLM proposal)

Update belief
(feedback-augmented context)

Black-box function
(RL training & evaluation)



ChatGPT Gemini
deepseek Claude

Decision rule
(e.g., Softmax sampling)

Probabilistic model
(large language model)