Statement of Purpose

Xingjian Bai *Ph.D. Applicant*

My research interest lies in the convergence of algorithms, mathematics, and deep learning, a domain where elegant theories meet practical models. My journey so far has been a quest to merge these fields: adapting classic algorithms to leverage inaccurate learned predictions [1], and enhancing machine learning models through rigorous theoretical analysis, spanning the fields of robustness [2], generative modeling [4], and reinforcement learning [3]. Integrating the strength of algorithms and deep learning, my ambition is to develop intelligent models with spontaneous and logical abilities, able to perceive, think, and explore the world beyond human-level capabilities.

Algorithms were my idea of fun in high school: Participating in the Chinese Olympiad of Informatics, advancing to the USACO camp, and winning 1st place in the Canadian Computing Olympiad national team selection, I enjoyed learning algorithms with friends worldwide. Later, I delved into algorithms with theoretical foundations; through the Fast Fourier Transform, the Miller-Rabin test, and Euler's iterative method, I glimpsed a dazzling world behind codes, of analytical and algebraic structures I could not understand. To explore the theory underpinning algorithms, I majored in mathematics and computer science at Oxford.

My tipping point occurred when I encountered simulated annealing. Looking back, my realization was remarkably similar to what Prof. Sanjeev Arora described in his blog, Off the Convex Path. It was the first algorithm I encountered that was not provably correct, yet possessed the power to tackle NP-hard problems in natural, non-adversarial settings. Suddenly, all the classic algorithms I learned collapsed into a plane in my brain – a plane of idealized worlds without errors. However, the points outside of the plane are countless and attractive. Real-world problems are intractable, erroneous, and stochastic in nature. What, then, is the suitable paradigm to solve them? Intrigued by this question, I entered the world of deep learning, beginning my journey to combine the most practical models with elegant theories.

Learning-Augmented Sorting My first exploration came about using inaccurate advice to speed up sorting, perhaps the most fundamental algorithmic task. This followed the line of work on learning-augmented algorithms, started by [1], which seeks to design algorithms with theoretical guarantees that can leverage predictions from ML models. Previous works have proposed approximated sorting algorithms that can leverage noisy signals, but no one has invented provably correct sorting algorithms to cope with erroneous advice. Working with Prof. Christian Coester at Oxford, I proposed a setting where a quick-and-dirty comparison function is available, besides the original exact (but slow) comparator. I drew insights from self-balanced data structures and randomized algorithms, and designed a novel algorithm that can precisely sort with a small number of clean comparisons, if given dirty comparisons are mostly correct. Subsequently, I proved the optimality of my algorithm and demonstrated its effectiveness in experiments.

This work [2], appearing at **NeurIPS 2023**, has attracted attention from the theoretical and practical worlds. After hearing about my algorithm at a workshop, one professor told me that he was considering teaching it at his university. Researchers from a drug discovery lab told me that, my new result is potentially helpful to reduce the number of comparative experiments needed to select molecular structures among thousands of candidates.

Wasserstein Distributional Adversarial Robustness of Neural Networks Unlike algorithms with provable guarantees, neural networks (NNs) are vulnerable to worst-case attacks, making them unsafe for real-world applications. Despite the extensive literature on effective adversarial attacks, theoretical tools for analyzing the worst-case behavior of NNs remain underdeveloped. Working with Prof. Jan Obloj at Oxford, we used optimal transport to lift the robustness problem into infinite-dimensional space, then leveraged Distributional Robust Optimization (DRO) sensitivity analysis to evaluate the robustness. As the only team member with an ML background, I worked on translating mathematical tools into tractable algorithms. Specifically, I proposed a new loss function, Rectified DLR, along with an adversarial attack algorithm with certified accuracy bound. Further, to mitigate the intractability of minimax optimization in training, I employed off-policy learning techniques from RL and proposed a novel training algorithm. This work [3] will appear at **NeurIPS 2023**.

Neuro-symbolic control on diffusion models For image generation tasks, diffusion models have all the strengths of "System 1 thinking"[1]: they excel at rendering emotions, styles, and scene diversity. However, they fall short in compositionality, consistency, and generalizability, the hallmarks of a logical system. To address these limitations, I worked with Prof. Jiajun Wu at Stanford to improve the relational compositionality of diffusion models. Diverging from traditional end-to-end diffusion models, I developed a two-stage pipeline using bounding boxes as a bridge and decomposed image generation into the generation of individual objects and relations. This approach achieved state-of-the-art performance in generating multiple objects and multiple relations on synthetic datasets.

Future Directions My aspiration is to develop learning systems that combine the spontaneous strength of ML with the logical precision of algorithms. Such systems would be more interpretable and composable than traditional ML models while surpassing classic algorithms in learning abilities.

To start with, several questions immediately pique my interest:

Xingjian: pick 3 to 4 according to different professors

- **How to invent algorithms to leverage multiple "dirty" predictors?** Countless algorithmic and practical scenarios, like sorting, rely on a potentially costly evaluator (e.g., comparator, attacker, or scoring system). ML models could be trained to predict such evaluations, albeit with errors, and produce a spectrum of "dirty" predictors with varying accuracy and cost. Then, I aim to develop algorithms with provable guarantees that can leverage the spectrum of dirty predictors to yield provable results.

- **How to seamlessly integrate deep learning models with provably correct algorithms?** I believe the key lies in the interface. Previous works in learning-augmented algorithms have modeled predictions with discrete structures, neglecting the structure of networks producing such predictions, for the convenience of complexity analysis. However, this simplification severely limits the collaboration between ML models and algorithms. I aim to design continuous, interactive, bi-directional interfaces tailored between specific types of networks, such as transformers and GNNs, and algorithms. The integrated system can leverage both learning and logical reasoning abilities adaptively.

- **In statistical inference, integer programming, and optimization, what are some problems where ML can be applied to inform traditional algorithms?** Whereas the focus in these fields has predominantly been on worst-case analysis, the advent of ML provides possibilities for a paradigm shift. I aim to identify and tackle problems where the integration of ML can go beyond traditional algorithms and lead to more efficient solutions.

- **How to leverage algorithmic techniques to enhance ML models?** Classic algorithms offer a wealth of powerful tools – recursion, divide-and-conquer, dynamic programming, etc. Integrating them into learning systems can take logical reasoning capabilities to the next level. As an initial attempt, I would try to use implicit neural networks to define networks with algorithmic specifications, thus integrating classic algorithmic behavior into learning systems. The primary challenge involves finding appropriate contexts and robustly training these sophisticated systems.

- **In ML pipelines, which components can be replaced by more transparent, theory-driven systems?** In my thesis[5], I replaced the diffusion denoising network with a fixed point dynamic system. Leveraging results in Differential Equations, the mathematically grounded formulation allows reallocating computation across timesteps. With insights into classic algorithms and mathematics, I am eager to closely examine more ML pipelines, identify more exploitable structures, and replace them with systems with better theoretical properties.

Why Stanford A paragraph matching myself with professors in a specific university.

[1] as described in *Thinking, Fast and Slow*

References

[1] **X. Bai** and C. Coester, “Sorting with Predictions,” in *Conference on Neural Information Processing Systems (NeurIPS)*, 2023, Available: `https://arxiv.org/abs/2311.00749`.

[2] **X. Bai**, G. He, Y. Jiang, and J. Obloj, “Wasserstein Distributional Robustness of Neural Networks,” in *Conference on Neural Information Processing Systems (NeurIPS)*, 2023. Available: `https://arxiv.org/abs/2306.09844`.

[3] J. Karwowski, O. Hayman, **X. Bai**, K. Kiendlhofer, C. Griffin, and J. Skalse, “Goodhart’s Law in Reinforcement Learning,” Under review. Available: `https://arxiv.org/abs/2310.09144`.

[4] **X. Bai** and L. Melas-Kyriazi, “Fixed Point Diffusion Models” Under review.

[5] T. Lykouris and S. Vassilvitskii, “Competitive Caching with Machine Learned Advice,” in *Proc. of the 35th Int. Conf. on Machine Learning (ICML 2018)*, vol. 80, pp. 3302–3311, 2018.