Shortest Edit Path Crossover: A Theory-driven Solution to the Permutation Problem in Evolutionary Neural Architecture Search

Xin Qiu[†] and Risto Miikkulainen[†‡]
[†]Cognizant AI Labs, [‡] The University of Texas at Austin

1. Introduction

Background: Population-based search is a popular method for black-box neural architecture search (NAS). It is usually based on evolutionary algorithms (EAs), which mimic natural evolution by maintaining a population of solutions and evolving them through mutation and crossover. However, most recent evolutionary NAS methods only employ mutation operation, due to the permutation problem in applying crossover to NAS.

The Permutation Problem: This problem is due to isomorphisms in graph space, i.e., functionally identical architectures are mapped to different encodings/representations, making crossover operations disruptive.

Limitations of Existing Solutions:

- Only work on fixed or constrained topologies
- Limited to a particular algorithm or search space
- Only focused on empirical verification without a theoretical analysis of potential solutions

Main Contributions of this Work:

- A new shortest edit path (SEP) crossover operator, which is a generalizable solution that can be applied to arbitrary architectures or search space
- A theoretical analysis of mutation, standard crossover, reinforcement learning (RL), and the proposed SEP crossover in the NAS domain
- Empirical studies on NAS benchmarks, demonstrating the effectiveness of the SEP crossover

2. Method

Main Idea: The proposed SEP crossover consists of two main steps:

- First, given two parent architectures represented by two attributed directed graphs, calculate the shortest edit path that minimizes the graph edit distance (GED) between them
- Second, randomly pick half of the graph edit operations from the shortest edit path, then apply them to one of the parents to obtain the offspring

Motivation: This operator is motivated by a common observation in the literature that the differences in predictive performance between two architectures are positively correlated with their GEDs. This observation suggests that the edits in the SEP encode fundamental differences between two architectures that matter to predictive performance. An offspring that lies in the middle of this SEP can explore the search regions where the parents have fundamental discrepancies. At the same time, the offspring can automatically preserve those common substructures between parents, avoiding unnecessary disruptive behaviors, and thus avoiding the permutation problem. Figure 1 shows how the SEP crossover resolves the permutation problem.

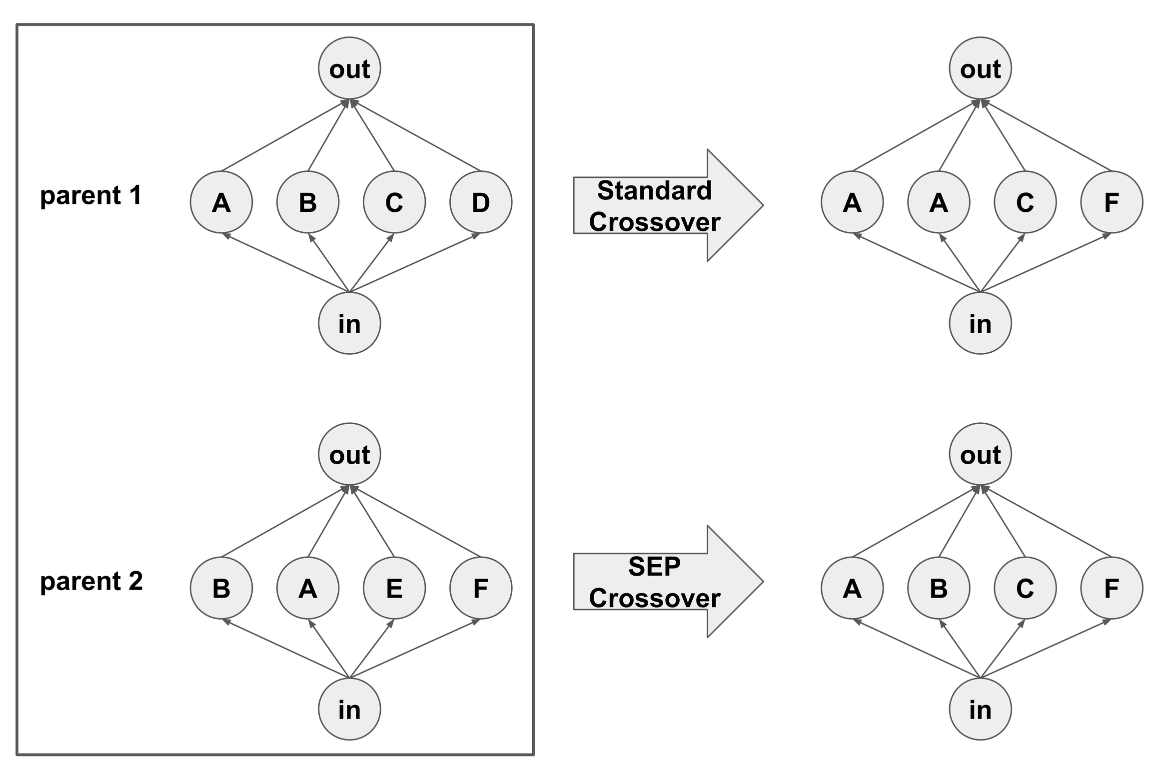

Figure 1: The permutation problem and the SEP crossover solution. The two parent architectures share vertices A and B. Although these two vertices appear in a different order, together they implement the same function, and this function should not be disrupted during crossover. However, standard crossover cannot identify the subgraph isomorphism, and it loses this substructure. In contrast, the shortest edit path calculation recognizes the isomorphism, and as a result, the SEP crossover preserves this substructure. Thus, the SEP crossover only explores the parts that are functionally inconsistent between the two parents.

3. Theoretical Analysis

Brief Summary: In this work, the SEP crossover, standard crossover, mutation, and RL approaches to NAS will be analyzed theoretically, showing that the SEP crossover has an advantage in improving the expected quality of generated graphs. The main points are summarized as follows:

- New interpretations of graph edit distance, crossover and mutation based on *attributed adjacency matrices* are defined. Based on the common assumption that the performance difference between two architectures is positively correlated with their GED, theorems regarding the expected improvement of SEP crossover, standard crossover, and mutation in NAS setup are derived.
- An NAS RL method is interpreted using the newly developed concepts, and two extreme cases whose combinations span the possible states of the RL process are defined, with theorems derived for expected improvement for both.
- A numerical analysis brings these theorems together, showing that the SEP crossover results in more improvement than the other methods in common NAS setups (highlighted results in Figure 2).
- Additional studies verify the robustness of the SEP crossover under inaccurate GED calculations.

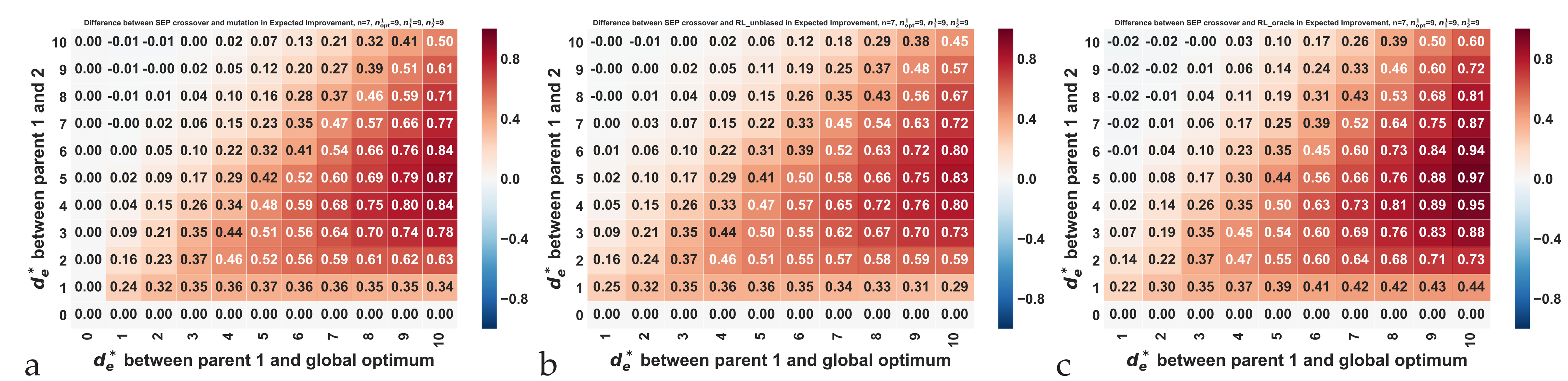

Figure 2: Comparison of expected improvement between SEP crossover, mutation, and RL in NAS-bench-101. (a) Differences between SEP crossover and mutation under different $d^*_{e,\hat{\mathcal{G}}_1,\hat{\mathcal{G}}_2}$ (y-axis) and $d^*_{e,\hat{\mathcal{G}}_{\text{opt}},\hat{\mathcal{G}}_1}$ (x-axis) combinations, where $d^*_{e,\hat{\mathcal{G}}_1,\hat{\mathcal{G}}_2}$ measures the topological differences between the parents, and $d^*_{e,\hat{\mathcal{G}}_{\text{opt}},\hat{\mathcal{G}}_1}$ measures the quality of the parent in terms of its topological difference to the global optimum. (b) Differences between SEP crossover and uniform RL agent (one of the two extreme cases for RL). (c) Differences between SEP crossover and oracle RL agent. The expected improvement of SEP crossover is larger (i.e. more red) than those of mutation and RL almost everywhere. Thus, the SEP crossover has a theoretical advantage over mutation and RL.

4. Empirical Studies

Figure 3 highlights the main experimental results that demonstrate the effectiveness of SEP crossover.

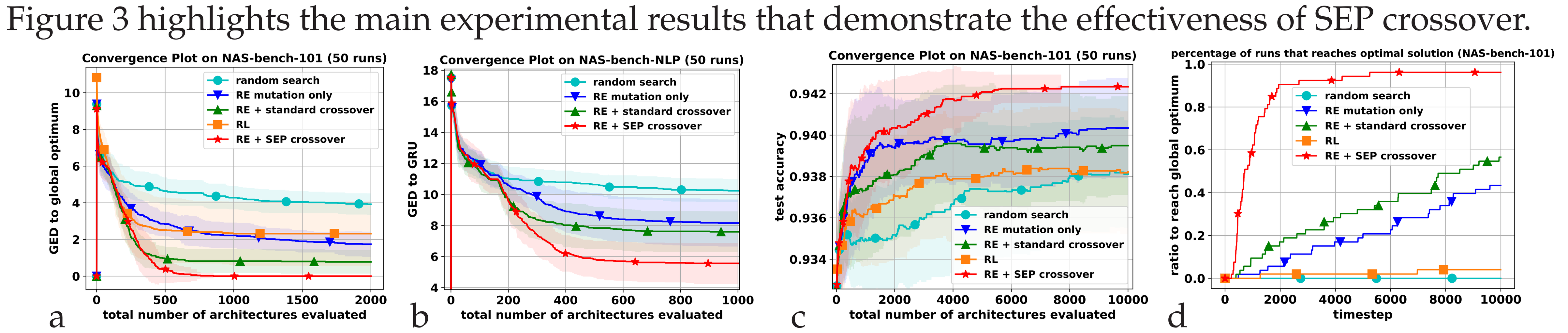

Figure 3: Convergence in noise-free (a,b) and noisy environments (c,d). (a) GED to global optimum in NAS-bench-101. (b) GED to GRU (targeted solution) in NAS-bench-NLP. (c) Average testing accuracy in NAS-bench-101. (d) Percentage of runs that reach the global optimal architecture in NAS-bench-101. In all experiments, the SEP crossover performs consistently better than the other methods in both noise-free and noisy environments. The SEP crossover also reaches the global optimum significantly more efficiently than the other methods in NAS-bench-101. Together the experiments show that the SEP crossover consistently improves evolutionary NAS in practice.

5. Conclusion

The SEP crossover is proposed as a solution to the permutation problem in evolutionary NAS. Its advantage over standard crossover, mutation and RL was first shown theoretically, with a focus on the expected improvement of GED to global optimal. Empirical studies were then performed to verify the applicability of the theoretical results, and demonstrate the superior performance of the SEP crossover in both noise-free and noisy environments. The SEP crossover therefore allows taking full advantage of evolution in NAS, and potentially other similar design problems as well.

6. Source Code & Contact

github.com/cognizant-ai-labs/sepx-paper

Email: qiuxin.nju@gmail.com, risto@cs.utexas.edu