

Alexei Finski - Statement of Research Interests

My interests are in the areas of search and measurement.

1. Local Search

Improving the representation of search solutions arguably provides the most effective way to consistently improve local search, a foundational building block for solving many hard problems. Towards this goal, I developed new approaches to representing and evaluating search solutions.

1.1. Locality in Search

The generative locality approach defines the locality of search solutions at each step of a local search with i) a generative model and ii) a generation randomization process that are used to reach outside an interpolant of previous search solutions in an automated manner.

The approach was applied to address a theoretically grounded limitation of studies conducted by researchers at the University of Washington, Microsoft Research, and Moderna.

1.2. Search and Explainability

The notion of explainability could be formally defined as search according to a human-centric objective function in a search space provided by a data representation. Causal or potentially causal information about rare events is a commonly desired explainability objective.

Causal information can often be contained in sequences of discrete tokens, such as nucleotides or (quantized) values across a composition of functions, associated with a discrete or continuous value indicating an occurrence or property of a rare event.

I developed an approach that enables the enrichment of sequences of discrete tokens associated with rare events in a concise logical representation, and facilitates the identification and interpretable representation of causal information.

2. Software Engineering and Learning

The commonly used libraries for implementing, training, and evaluating learning algorithms often improve usability at the expense of research flexibility.

However, achieving a higher relational abstraction level does not require a trade-off between usability and flexibility, as demonstrated by functional programming approaches.

I developed an approach that provides a higher-order representation of function composition in neural networks and enables code simplicity and research flexibility. This approach could potentially be generalized with the goal of providing extensibility and programmability, comparable to a language such as LISP, without relying on a third-party framework.

3. Distributed Infrastructure

Algorithm analysis and therefore design tools in the emerging area of distributed infrastructure remain limited, but are necessary to ensure the scalability of distributed computation, including learning.

While algorithm parallelism is a meaningful objective in a multi-node/pod infrastructure setting, differences in transfer latencies associated with a computation present an additional consideration. Assuming the same bandwidth and a fixed-sized transfer block of bits, a better algorithm uses more bits of a transferred block (spatial locality) and reuses the transferred bits to a greater extent (temporal locality) as latency increases. A better locality mitigates a higher latency. Amortized cost of access presents a worthwhile objective.

Having briefly explored the area of distributed infrastructure, I am interested in contributing to the development of theoretically grounded approaches in this emerging area.