DECOMPOSABLE PROBABILITY-OF-SUCCESS METRICS IN ALGORITHMIC SEARCH

Tyler Sam, Jake Williams, Abel Tadesse, Huey Sun, George Montañez

{tsam, jlwilliams, gmontanez}@g.hmc.edu
aleulseged20@cmc.edu, hssa2016@pomona.edu

The Walter Bradley Center for NATURAL & ARTIFICIAL INTELLIGENCE

Abstract

To generalize results within the algorithmic search framework for application in subfields of machine learning such as transfer learning, we define a property called decomposability, which describes probability-of-success metrics that can be expressed as a linear operation on a probability distribution. Using decomposable metrics, we generalize certain impossibility results and bounds on success to hold for a variety of ways of measuring success.

Motivation

Previous results within the search framework utilize a success metric that averages the probability of success over all iterations of a given algorithm. This metric is not applicable to cases where one specific iteration of an algorithm is required such as **transfer learning**, where the probability of success at the final step of the algorithm is more relevant than the average probability of success.

Search Framework

ML-as-search

To analyze a variety of search and learning algorithms, Montañez describes a framework that formalizes the search problem, and demonstrates how a wide variety of machine learning problems can be cast into the framework [3].

Regression as Search

- $\Omega =$ the set of possible regression functions;
- $T = \{g : g \in \Omega, \Xi(\mathcal{L}, g, V) \leq \epsilon\}$;
- $F = \{D = \{(x_i, y_i)\}, V = \{(x_j, y_j)\}\mathcal{L}_s\}$;
- $F(\omega_i) = \mathcal{L}(\omega_i)$.

where D is the training data set, V is the test data set, and $\mathcal{L}$ is the loss function.

Within the framework, we have an iterative algorithm which searches for elements of target set according to the current probability distribution generated from previous queries, shown in Figure 1. The abstraction of finding the next probability distribution as a black-box algorithm allows the search framework to encompass all types of search problems.

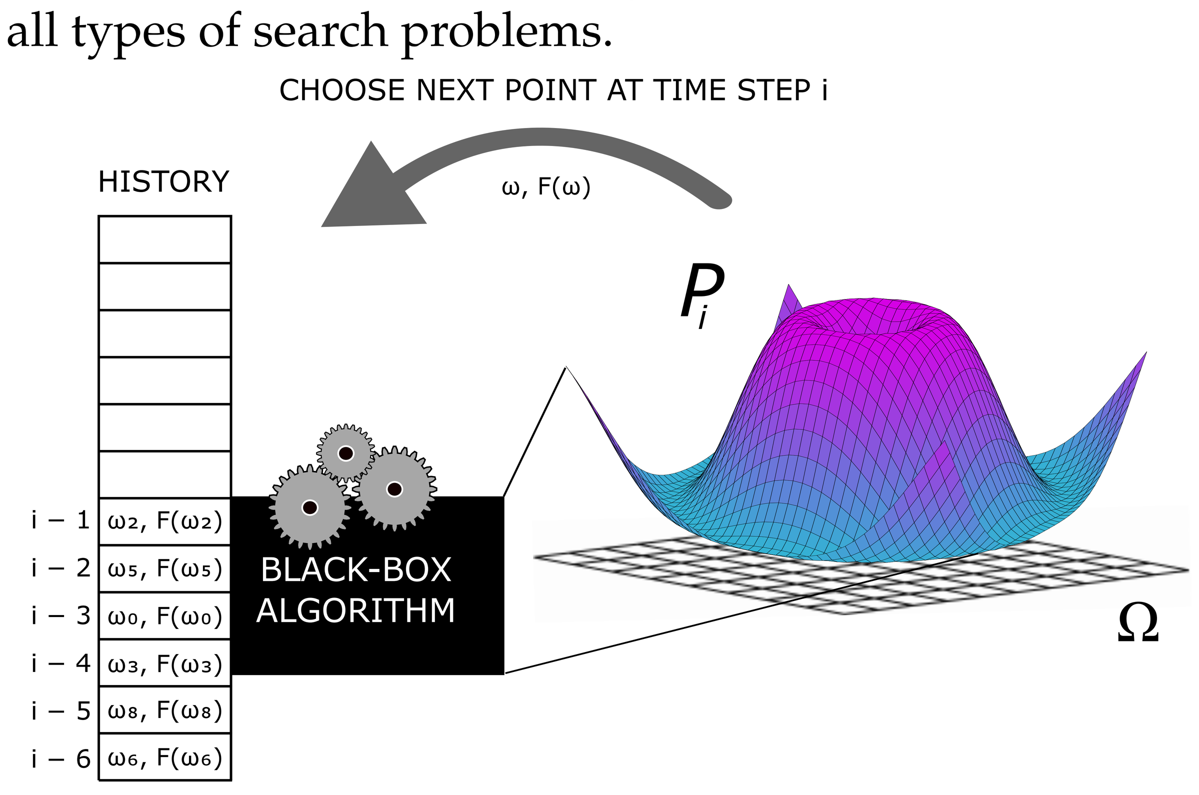

Figure 1: Algorithmic Search Framework.

Expected per-query probability of success

In order to compare search algorithms, Montañez defined the expected per-query probability of success,

$$q(t, f) = \mathbb{E}_{\tilde{P}, H}\left[\frac{1}{|\tilde{P}|}\sum_{i=1}^{|\tilde{P}|} P_i(w \in t)\Big| f\right] = \overline{P}(X \in t|f) \quad (1)$$

where $\tilde{P}$ is the sequence of probability distributions, H is the search history, and t and f are the target set and information resource of the search problem, respectively [1]. This metric of success is useful because $q(t, f) = \mathbf{t}^\top \overline{\mathbf{P}}_f$, where $\overline{\mathbf{P}}_f$ is the average of the vector representation of the probability distribution from the search algorithm at each step.

Results

A probability-of-success metric ϕ is **decomposable** if and only if there exists a $\mathbf{P}_{\phi,f}$ such that

$$\phi(t, f) = \mathbf{t}^\top \mathbf{P}_{\phi,f} = P_\phi(X \in t|f), \quad (2)$$

where $\mathbf{P}_{\phi,f}$ is conditionally independent of t given f. This definition preserves the practical benefits of the expected per-query probability of success, i.e., the ability to express the probability of success as a product of the target set and a probability distribution, and adds additional flexibility, giving us a tool to apply results to a wider breadth of success metrics. For example, we demonstrate that a set of theorems proved by Montañez [2] [3] for the expected per-query probability of success hold for all decomposable metrics.

Famine of Favorable Targets

To demonstrate the utility of decomposable metrics, we prove that the Famine of Favorable Targets, a bound on the proportion of targets on which an algorithm will succeed, applies to all decomposable metrics. For fixed $k \in \mathbb{N}$, fixed information resource f, and decomposable probability-of-success metric ϕ, define

$$\tau = \{T \mid T \subseteq \Omega, |T| = k\}, \text{ and}$$
$$\tau_{q_{\min}} = \{T \mid T \subseteq \Omega, |T| = k, \phi(T, f) \geq q_{\min}\}.$$

Then,

$$\frac{|\tau_{q_{\min}}|}{|\tau|} \leq \frac{p}{q_{\min}} \quad (3)$$

where $p = \frac{k}{|\Omega|}$.
Here, we compare the success of an algorithm against a fixed probability of success $q_{\min}$. This theorem thus upper bounds the number of targets for which the probability of success of the search is greater than $q_{\min}$.

Conclusion

Results in the algorithmic search framework hold for all decomposable probability-of-success metrics. This leads to a number of useful insights:

- Algorithmic performance is conserved with respect to all decomposable probability-of-success metrics.
- Favorable targets are scarce no matter your decomposable probability-of-success metric.
- Without biasing an algorithm towards an expected target set, it will not perform better than uniform random sampling.

Acknowledgment

This work was supported by the Walter Bradley Center for Natural and Artificial Intelligence. We thank Dr. Robert J. Marks II (Baylor University) for providing support and feedback. We also thank Harvey Mudd College's Department of Computer Science for their continued resources and support.

References

[1] G. D. Montañez. The Famine of Forte: Few Search Problems Greatly Favor Your Algorithm. In *Systems, Man, and Cybernetics (SMC), 2017 IEEE International Conference on*, pages 477–482. IEEE, 2017a.

[2] G. D. Montañez. *Why Machine Learning Works*. PhD thesis, Carnegie Mellon University, 2017b.

[3] G. D. Montañez, J. Hayase, J. Lauw, D. Macias, A. Trikha, and J. Vendemiatti. The Futility of Bias-Free Learning and Search. In *32nd Australasian Joint Conference on Artificial Intelligence*, pages 277–288. Springer, 2019.