Gradient-Free Optimization via Associated Memory-Assisted Genetic Algorithms

Abdullah Naeem (anaeem@uno.edu), Md Wasi Ul Kabir (mkabir3@uno.edu), Md Tamjidul Hoque (thoque@uno.edu)
Department of Computer Science, University of New Orleans, New Orleans, LA, USA

Introduction

- Optimization in high-dimensional, non-convex landscapes remains a critical challenge in machine learning, particularly when gradient information is unavailable or unreliable.
- We present a novel gradient-free optimization framework: the Associated Memory-Assisted Local Genetic Algorithm (AMLGA). AMLGA integrates adaptive local search operators with a dynamic associated-memory cache.
- Unlike traditional evolutionary strategies that rely solely on stochastic recombination, AMLGA learns and caches high-quality substructures during the evolutionary process.
- The framework employs a two-point associated-memory crossover operator that recombines cached high-quality segments with current candidates. This approach helps escape local minima and accelerates convergence.

Motivation

- Gradient-based optimization is infeasible for non-differentiable fitness landscapes, such as those encountered in conformational ensemble generators for intrinsically disordered proteins (CEG-IDP).
- We explore AMLGA as a versatile gradient-free alternative, demonstrating its applicability to neural network weight optimization as well as to CEG-IDP workflows using a statistical energy function (IDPEnergy) for fitness evaluation.

Application of AMLGA in CEG-IDP

AMLGA can easily be integrated into the CEG-IDP Workflow as depicted in the Flow Chart in **Fig 1.**

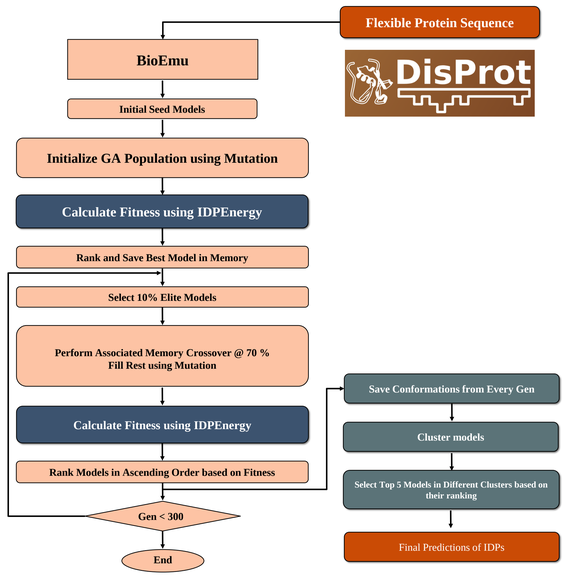


***Figure 1**: AMLGA Integrated Workflow for CEG-IDP.*

Associated Memory Assisted GA

The core innovation is the integration of Associated Memory (AM) into the evolutionary search. Unlike standard GAs that rely solely on random crossover, our approach "learns" and "remembers" the best structural fragments found so far.

Cache Design for up to 2 Point Cross Over

We maintain dynamic memory matrices that store the best-performing structural fragments as depicted in **Fig 2**.
$\boldsymbol{AM_{UT}}$ (Upper Triangular): Stores the best "Head" (N-terminal) fragments.
$\boldsymbol{AM_{LT}}$ (Lower Triangular): Stores the best "Tail" (C-terminal) fragments.
$\boldsymbol{AM_{MID}}$: Stores the best "Middle" segments between variable cut points.

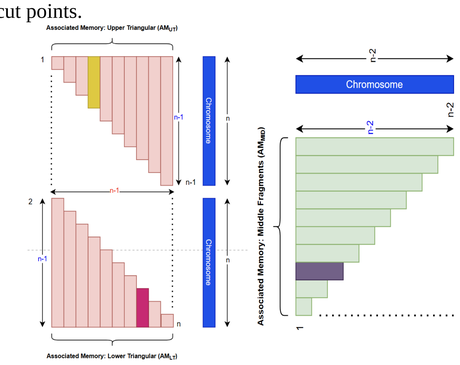


***Figure 2**: Associate Memory Cache for 2-Point Cross Over.*

The Associated Memory Assisted Crossover Operation

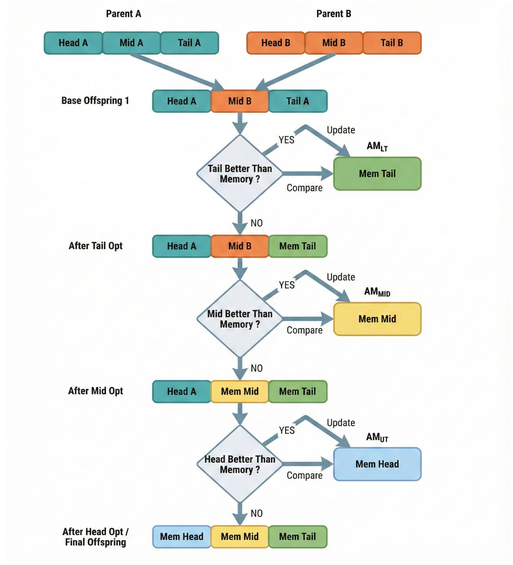


***Figure 3**: Activity Diagram for the AM Assisted Crossover Operation.*

Population Management with Elites

- Elites are preserved.
- Exploited with HGR
- Participate in Crossover to supply strong genetic material to non-elites.
- Tier-2 (non-elite) population is explored and exploited with Mutation, Crossover and Twin Removal as shown in **Fig 4.**

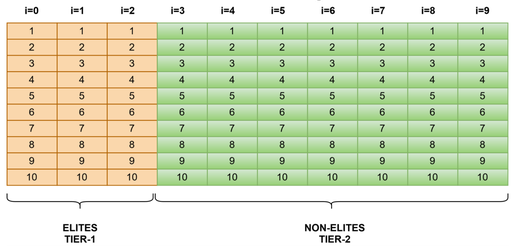
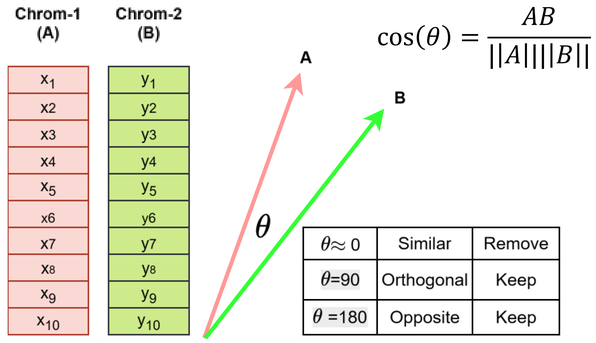


***Figure 4**: Preserved Elites and Non-Elite Population Sorted by Fitness.*

Diversity Management Using Twin Removal



$$\cos(\theta) = \frac{AB}{||A||||B||}$$

***Figure 5**: Twins Identified by Cosine Similarity are Removed and Replaced with a Randomly Generated Candidate to Keep the Population Diverse.*

Homologous Gene Replacement

HGR further improve elites by identifying the best and worst performing gene's loci in the candidate chromosome as in **Fig 6**.

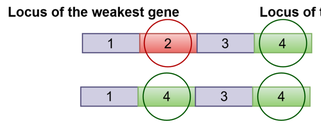


***Figure 6**: HGR Identifies and Replaces Low Performing Genes with Best Performers within the Elite Cohort of the Population.*

Mutation

A randomly selected gene as shown in **Fig 7** is perturbed by adding a value sampled from a Gaussian Distribution with mean 0 and standard deviation, σ.

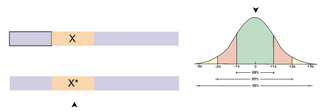


***Figure 7**: Mutation of a Local Gene Through Addition of Gaussian Noise.*

Results

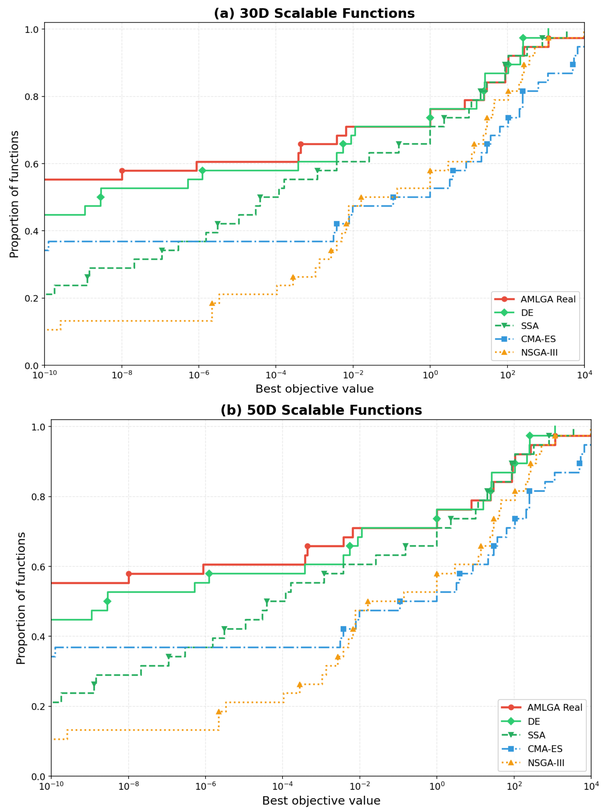


***Figure 8**: Empirical Cumulative Distribution for Best Objective Value.*

- AMLGA (Real) consistently outperforms DE, SSA, CMA-ES, and NSGA-III across benchmark functions. The ECDF analysis in **Fig. 8** clearly shows that AMLGA (Real) dominates these established algorithms.
- It solves 60.5% of functions to high precision (≤1e-6) compared to SSA (36.8%), CMA-ES (36.8%), and NSGA-III (13.2%). The performance gap is consistent across both problem dimensions.

Conclusions

- In this study, we propose a gradient-free optimization of AI models using the Associated Memory Assisted Genetic Algorithm and present it as a potential tool for Disordered Protein Structure Predictions (DPSP).
- On standard benchmarks, AMLGA outperforms state-of-the-art evolutionary and meta-heuristic algorithms.
- We also provide insights on integrating AMLGA into CEG-IDP workflows using a statistical energy function (IDPEnergy) for fitness evaluation.

Acknowledgement

Funding: Research supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number **P20GM103424-21.**