EDAs: theoretical approaches, methodological enhancements, and applications

Josu Ceberio, Unai Garciarena, Jose A. Lozano, Mikel Malagon, Alexander Mendiburu, Roberto Santana, and Jon Vadillo

Intelligent Systems Group (ISG)
University of the Basque Country (UPV/EHU)
http://www.sc.ehu.es/ccwbayes/

In this talk we briefly review a number of relevant EDA research directions participated by researchers of the Intelligent Systems Group in the period 2011-2024. These directions comprise: theoretical approaches, methodological enhacenments and improvements to EDAs, probabilistc modeling in EDAs for permutation-based representations, and challenges in the real-world applications of EDAs.

Our approaches for a theoretical analysis of EDAs include a comprehensive characterization of the behaviors of EDAs based on factorizations [9, 10] and a mathematical analysis of EDAs with distance-based exponential models [38]. Also, the impact of problems such as symmetric problems on EDAs [24] and the limits of effectiveness of EDAs have been investigated [11].

Among the methodological extensions and enhancements of EDAs are the introduction of new sampling strategies [25, 30], the research and design of selection strategies tailored for EDAs [31, 37], the extension to multi-objective problems [19, 20, 44], the use of probabilistic models as surrogates [29], parallel implementations [21], design of hard instances for EDAs [27, 44] and the incorporation of methods for dealing with constrained problems [35]. One recent research direction is the use of deep neural networks to represent and sample high fitness solutions. This includes the use of generative models [14, 8, 18] and the investigation of sampling techniques especifically suited for neural networks [12, 15].

Work on the proposal of new probabilistic models that allow EDAs to deal with permutation-based optimization problems have considered a variety of models including several variants of Mallows models [4, 2, 3, 7, 1], Plackett-Luce [5], doubly stochastic matrix models [33] and other alternative representations [17]. The work in this direction includes extensions to multi-objective problems [43, 42], and the proposal of a software library that implements permutation-based EDAs [16].

Application of EDAs to real-world problems in diverse domains have been proposed. They include problems from bioinformatics [28, 26], planning [41], routing [40], optimization and design of quantum-based methods [23, 32], cryptographic applications [22], compiler flag selection [13], the graph partitioning problem [6], space trajectory optimization [34, 36] and, more recently, the problem of delineating site-specific management zones [39].

References

1. E. Arza, A. Perez, E. Irurozki, and J. Ceberio. Kernels of Mallows models under the Hamming distance for solving the quadratic assignment problem. *Swarm and Evolutionary Computation*, 59:100740, 2020.
2. J. Ceberio, E. Irurozki, A. Mendiburu, and J. A. Lozano. A distance-based ranking model estimation of distribution algorithm for the flowshop scheduling problem. *IEEE Transactions on Evolutionary Computation*, 18(2):286–300, 2013.
3. J. Ceberio, E. Irurozki, A. Mendiburu, and J. A. Lozano. A review of distances for the Mallows and Generalized Mallows estimation of distribution algorithms. *Computational Optimization and Applications*, 62(2):545–564, 2015.
4. J. Ceberio, A. Mendiburu, and J. A. Lozano. Introducing the Mallows model on estimation of distribution algorithms. In *Neural Information Processing: 18th International Conference, ICONIP 2011, Shanghai, China, November 13-17, 2011, Proceedings, Part II 18*, pages 461–470. Springer, 2011.
5. J. Ceberio, A. Mendiburu, and J. A. Lozano. The Plackett-Luce ranking model on permutation-based optimization problems. In *Evolutionary Computation (CEC), 2013 IEEE Congress on*, pages 494–501. IEEE, 2013.
6. J. Ceberio, A. Mendiburu, and J. A. Lozano. A square lattice probability model for optimising the graph partitioning problem. In *Proceedings of the 2017 Congress on Evolutionary Computation CEC-2017*, San Sebastian, Spain, 2017. IEEE Press.
7. J. Ceberio, R. Santana, A. Mendiburu, and J. A. Lozano. Mixtures of generalized Mallows models for solving the quadratic assignment problem. In *Proceedings of the IEEE Congress on Evolutionary Computation CEC 2015*, pages 2050–2057, Sendai, Japan, 2015.
8. J. Ceberio and V. Santucci. Model-based gradient search for permutation problems. *ACM Transactions on Evolutionary Learning and Optimization*, 3(4):1–35, 2023.
9. C. Echegoyen, A. Mendiburu, R. Santana, and J. A. Lozano. On the taxonomy of optimization problems under estimation of distribution algorithms. *Evolutionary Computation*, 21(3):471–495, 2013.
10. C. Echegoyen, R. Santana, A. Mendiburu, and J. A. Lozano. Comprehensive characterization of the behaviors of estimation of distribution algorithms. *Theoretical Computer Science*, 598:64–86, 2015.
11. C. Echegoyen, Q. Zhang, A. Mendiburu, R. Santana, and J. A. Lozano. On the limits of effectiveness in estimation of distribution algorithms. In *Proceedings of the 2011 Congress on Evolutionary Computation CEC-2007*, pages 1573–1580. IEEE Press, 2011.
12. U. Garciarena, A. Mendiburu, and R. Santana. Envisioning the benefits of back-drive in evolutionary algorithms. In *2020 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8. IEEE, 2020.
13. U. Garciarena and R. Santana. Evolutionary optimization of compiler flag selection by learning and exploiting flags interactions. In *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion*, GECCO '16 Companion, pages 1159–1166, 2016.
14. U. Garciarena, R. Santana, and A. Mendiburu. Expanding variational autoencoders for learning and exploiting latent representations in search distributions. In *Proceedings of the 2018 on Genetic and Evolutionary Computation Conference*, pages 849–856. ACM, 2018.
15. U. Garciarena, J. Vadillo, A. Mendiburu, and R. Santana. Adversarial perturbations for evolutionary optimization. In *International Conference on Machine*

Learning, Optimization, and Data Science (LOD-2021), volume 13164 of *Lecture Notes in Computer Science*, pages 408–422. Springer, 2021.

16. E. Irurozki, J. Ceberio, J. Santamaria, R. Santana, and A. Mendiburu. Algorithm 989: perm_mateda: A matlab toolbox of estimation of distribution algorithms for permutation-based combinatorial optimization problems. *ACM Transactions on Mathematical Software (TOMS)*, 44(4):47, 2018.
17. M. Malagón, E. Irurozki, and J. Ceberio. Alternative representations for codifying solutions in permutation-based problems. In *2020 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8. IEEE, 2020.
18. M. Malagón, E. Irurozki, and J. Ceberio. A combinatorial optimization framework for probability-based algorithms by means of generative models. *ACM Transactions on Evolutionary Learning*, 2024. Accepted for publication.
19. M. S. Martins, M. R. Delgado, R. Lüders, R. Santana, R. A. Gonçalves, and C. P. d. Almeida. Hybrid multi-objective Bayesian estimation of distribution algorithm: a comparative analysis for the multi-objective knapsack problem. *Journal of Heuristics*, 24(1):25–47, 2018.
20. M. S. Martins, M. E. Yafrani, M. Delgado, R. Lüders, R. Santana, H. V. Siqueira, H. G. Akcay, and B. Ahiod. Analysis of Bayesian network learning techniques for a hybrid multi-objective Bayesian estimation of distribution algorithm: a case study on MNK landscape. *Journal of Heuristics*, 27(4):549–573, 2021.
21. S. Muelas, A. Mendiburu, A. LaTorre, and J.-M. Peña. Distributed estimation of distribution algorithms for continuous optimization: How does the exchanged information influence their behavior? *Information Sciences*, 268:231–254, 2014.
22. S. Picek, R. Santana, and D. Jakobovic. Maximal nonlinearity in balanced boolean functions with even number of inputs, revisited. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 3222–3229. IEEE, 2016.
23. R. Santana, R. B. McDonald, and H. G. Katzgraber. A probabilistic evolutionary optimization approach to compute quasiparticle braids. In *Proceedings of the 10th International Conference Simulated Evolution and Learning (SEAL-2014)*, pages 13–24. Springer, 2014.
24. R. Santana, R. I. McKay, and J. A. Lozano. Symmetry in evolutionary and estimation of distribution algorithms. In *Proceedings of the 2013 Congress on Evolutionary Computation CEC-2013*, pages 2053–2060, Cancun, Mexico, 2013.
25. R. Santana and A. Mendiburu. Model-based template-recombination in Markov network estimation of distribution algorithms for problems with discrete representation. In *2013 Third World Congress on Information and Communication Technologies (WICT 2013)*, pages 170–175. IEEE, 2013.
26. R. Santana, A. Mendiburu, and J. A. Lozano. An analysis of the use of probabilistic modeling for synaptic connectivity prediction from genomic data. In *Proceedings of the 2012 Congress on Evolutionary Computation CEC-2012*, pages 3221–3228, Brisbane, Australia, 2012. IEEE Press.
27. R. Santana, A. Mendiburu, and J. A. Lozano. Evolving NK-complexity for evolutionary solvers. In *Companion Proceedings of the 2012 Genetic and Evolutionary Computation Conference GECCO-2012*, pages 1473–1474, 2012.
28. R. Santana, A. Mendiburu, and J. A. Lozano. Structural transfer using EDAs: An application to multi-marker tagging SNP selection. In *Proceedings of the 2012 Congress on Evolutionary Computation CEC-2012*, pages 3484–3491. IEEE Press, 2012.
29. R. Santana, A. Mendiburu, and J. A. Lozano. Critical issues in model-based surrogate functions in estimation of distribution algorithms. In *Proceedings of the 4th*

Conference on Swarm, Evolutionary, and Memetic Computing (SEMCCO-2013), Lectures Notes in Computer Science, pages 1–13, Chennai, India, 2013. Springer.
30. R. Santana, A. Mendiburu, and J. A. Lozano. Message passing methods for estimation of distribution algorithms based on Markov networks. In *Proceedings of the 4th Conference on Swarm, Evolutionary, and Memetic Computing (SEMCCO-2013)*, LNCS, pages 419–430, Chennai, India, 2013. Springer.
31. R. Santana, A. Mendiburu, and J. A. Lozano. Customized selection in estimation of distribution algorithms. In *Proceedings of the 10th International Conference Simulated Evolution and Learning (SEAL-2014)*, pages 94–105. Springer, 2014.
32. R. Santana, Z. Zhu, and H. G. Katzgraber. Evolutionary approaches to optimization problems in Chimera topologies. In *Proceedings of the 2016 Conference on Genetic and Evolutionary Computation (GECCO-2016)*, pages 397–404, 2016.
33. V. Santucci and J. Ceberio. Doubly stochastic matrix models for estimation of distribution algorithms. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 367–374, 2023.
34. A. Shirazi, J. Ceberio, and J. A. Lozano. Evolutionary algorithms to optimize low-thrust trajectory design in spacecraft orbital precession mission. In *2017 IEEE Congress on Evolutionary Computation (CEC)*, pages 1779–1786. IEEE, 2017.
35. A. Shirazi, J. Ceberio, and J. A. Lozano. EDA++: Estimation of distribution algorithms with feasibility conserving mechanisms for constrained continuous optimization. *IEEE Transactions on Evolutionary Computation*, 26(5):1144–1156, 2022.
36. A. Shirazi, J. Ceberio, and J. A. Lozano. Trajectory optimization of space vehicle in rendezvous proximity operation with evolutionary feasibility conserving techniques. *Engineering Applications of Artificial Intelligence*, 117:105523, 2023.
37. A. Strickler, O. Castro Jr, A. Pozo, and R. Santana. An investigation of the selection strategies impact on MOEDAs: CMA-ES and UMDA. *Applied Soft Computing*, 62:963–973, 2018.
38. I. Unanue, M. Merino, and J. A. Lozano. A mathematical analysis of EDAs with distance-based exponential models. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 429–430, 2019.
39. J. Velasco, S. Vicencio, J. A. Lozano, and N. M. Cid-Garcia. Delineation of site-specific management zones using estimation of distribution algorithms. *International Transactions in Operational Research*, 30(4):1703–1729, 2023.
40. J. Wang, K. Tang, J. A. Lozano, and X. Yao. Estimation of the distribution algorithm with a stochastic local search for uncertain capacitated arc routing problems. *IEEE Transactions on Evolutionary Computation*, 20(1):96–109, 2015.
41. P. Yang, K. Tang, and J. A. Lozano. Estimation of distribution algorithms based unmanned aerial vehicle path planner using a new coordinate system. In *2014 IEEE Congress on Evolutionary Computation (CEC)*, pages 1469–1476. IEEE, 2014.
42. M. Zangari, A. A. Constantino, and J. Ceberio. A decomposition-based kernel of Mallows models algorithm for bi-and tri-objective permutation flowshop scheduling problem. *Applied Soft Computing*, 71:526–537, 2018.
43. M. Zangari-de Souza, A. Mendiburu, R. Santana, and A. Pozo. Multiobjective decomposition-based Mallows models estimation of distribution algorithm. A case of study for permutation flowshop scheduling problem. *Information Sciences*, 397–398:137–154, 2017.
44. M. Zangari-de Souza, R. Santana, A. Mendiburu, and A. Pozo. On the design of hard mUBQP instances. In *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference*, pages 421–428. ACM, 2016.