Estimation of Distribution Algorithms based on Copula Functions

Rogelio Salinas-Gutiérrez^{*} Department of Computer Science Center for Research in Mathematics (CIMAT) Guanajuato, México rsalinas@cimat.mx Arturo Hernández-Aguirre Department of Computer Science Center for Research in Mathematics (CIMAT) Guanajuato, México artha@cimat.mx Enrique R. Villa-Diharce[‡] Department of Probability and Statistics Center for Research in Mathematics (CIMAT) Guanajuato, México villadi@cimat.mx

ABSTRACT

The main objective of this doctoral research is to study Estimation of Distribution Algorithms (EDAs) based on copula functions. This new class of EDAs has shown that it is possible to incorporate successfully copula functions in EDAs.

Categories and Subject Descriptors

I.2.8 [Artificial Intelligence]: Problem Solving, Control Methods, and Search—Heuristic methods; G.1.6 [Numerical Analysis]: Optimization—Global optimization, Unconstrained optimization

General Terms

Algorithms, Design, Performance

Keywords

EDAs, copula functions, graphical models

1. INTRODUCTION

Estimation of Distribution Algorithms (EDAs) are a well established paradigm in Evolutionary Computation (EC). This class of evolutionary algorithms employs probabilistic models for searching and generating promissory solutions. The goal in EDAs is to take into account the dependence structure in the best individuals and transfer them into the next population. A pseudocode for EDAs is shown in Algorithm 1.

*Doctoral student †Advisor ‡Co-Advisor

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright 2011 ACM 978-1-4503-0690-4/11/07 ...\$10.00.

Algorithm 1 Pseudocode for EDAs

- 1: assign $t \leftarrow 0$ generate the initial population P_0 with N individuals at
- random 2: select a collection of M solutions S_t , with M < N, from
- P_t 3: estimate a probabilistic model \mathcal{M}_t from S_t
- 4: generate the new population by sampling from the dis
 - tribution of S_t . assign $t \leftarrow t+1$
- 5: if stopping criterion is not reached go to step 2

The research on EDAs has been conducted in proposing and enhancing probabilistic models. Most of the probabilistic models used in EDAs for discrete domains are based on graphical models such as Bayesian and Markov Networks [9, 2, 24, 17, 16, 21, 22]. The discrete EDAs have been widely studied and they have been also used in several applications. However, for continuous domains, the Gaussian distribution is the most used assumption over the probabilistic model [12, 13, 14, 6]. For this reason, one of the current challenges for designing new continuous EDAs is finding multivariate models to adequately represent dependencies among the decision variables.

On the other hand, during the last decade, the copula functions have been widely used in many applications where nonlinear dependencies arise among variables. By using copula functions it is possible to separate the effect of dependence from the effect of marginal distributions in a joint distribution. However, they have been barely used in computer science applications where nonlinear dependencies are common and need to be represented.

The following section presents a brief introduction to copula functions and some EDAs based on copula theory.

2. CURRENT RESEARCH

One motivation for this research has been the possibility of proposing new EDAs for continuous domains. The copula based EDAs are able, at the same time, (1) to model dependencies among variables, and (2) to use flexible marginal distributions. The copula theory gives the means for getting such EDAs.

GECCO'11, July 12-16, 2011, Dublin, Ireland.

2.1 Copula Theory

Definition 1. A copula is a joint distribution function of standard uniform random variables. That is,

$$C(u_1,\ldots,u_d) = \Pr[U_1 \le u_1,\ldots,U_d \le u_d] ,$$

where $U_i \sim U(0, 1)$ for i = 1, ..., d.

The following result, known as Sklar's theorem, gives the relevance and practical utility to copula functions.

THEOREM 1 (SKLAR). Let F be a d-dimensional distribution function with marginals F_1, F_2, \ldots, F_d , then there exists a copula C such that for all x in \mathbb{R}^d ,

$$F(x_1, x_2, \dots, x_d) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d))$$

where \mathbb{R} denotes the extended real line $[-\infty,\infty]$. If $F_1(x_1)$, $F_2(x_2), \ldots, F_d(x_d)$ are all continuous, then C is unique. Otherwise, C is uniquely determined on $Ran(F_1) \times Ran(F_2) \times \cdots \times Ran(F_d)$, where Ran stands for the range.

According to Theorem 1 and using the chain rule for differentiating composite functions, any d-dimensional density f can be represented as

$$f(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d)) \cdot \prod_{i=1}^d f_i(x_i) , \quad (1)$$

where c is the density of the copula C, and $f_i(x_i)$ is the marginal density of variable x_i .

Equation (1) shows that the dependence structure is modeled by the copula function. This expression separates any joint density function into the product of the multivariate copula density and the marginal densities. This contrasts with the usual way to model multivariate distributions, which suffers from the restriction that the marginal distributions are usually of the same type. The separation between marginal distributions and a dependence structure explains why the copula functions are suitable tools for modeling dependencies in a joint distribution.

The copula theory was introduced by Sklar [23] and has been developed during the last fifty years. The interested reader of this theory is referred to [11, 15, 25].

2.2 Related works

To the best of our knowledge the theses [3, 1] are the first attempts to incorporate a multivariate Gaussian copula function in EDAs. Since then, other related works have been published. These papers present EDAs based on (1) the Gaussian copula function [30, 28], and (2) Archimedean copula functions with a fixed dependence parameter [29, 26, 27, 7, 31]. Unlike the previous papers that use Archimedean copula functions, the work [10] presents a way of estimating the copula parameter. These works use copula functions to model pairwise relationships among all variables and do not employ graphical models. On the other hand, we have proposed the use of the maximum likelihood method and copula entropies in order to (1) estimate the copula parameters, and (2) build a graphical model which establishes the most important dependencies between variables [19, 20, 18].

2.3 Our proposed EDAs

Our research has been conducted for designing new continuous EDAs. These EDAs are described below.

- In [19] we have presented an EDA based on a chain graphical model and bivariate copula functions. Every chain link in the graphical model represents the dependence between two decision variables. A bivariate copula function is used for modeling such dependence. The relationship between the mutual information and the bivariate copula entropy [8] is used for measuring the dependence between variables. The probabilistic model used in this paper can be seen as a generalization of the well known EDA MIMIC_C^G [12, 13]. In this work, for the first time, the estimation of copula parameters by means of the maximum likelihood function is presented.
- In [20], a regular vine [4, 5] is considered as a graphical model for designing a new EDA. This EDA is based on a particular family of regular vines called *D-vine*. In this work we propose a greedy algorithm for building such graphical model and we also show the theoretical relationship between the entropy of a trivariate copula function and the conditional mutual information.
- The previous works consider the use of a fixed copula function. Based on this observation, the work [18] is a recent contribution for selecting copula functions. A tree graphical model is employed in this paper. A copula function is selected for modeling the dependence between variables in each tree branch. The copula function is selected from a set of six bivariate copula functions and it is selected according to the likelihood function.

Some results obtained by D-vine EDA for three well known test functions in ten dimensions are reported in Table 1. For the Rosenbrock problem, the range of values reached by a D-vine EDA is closer to the global minimum than the range of values of the other EDAs. This would mean that a dependence structure based on a D-vine EDA is more adequate than the chain structure of the MIMIC. More details about these results can be seen in [20].

Figure 1 shows recent results by using a copula selection procedure in EDAs. In this figure, the success rate is reported by two chain and two tree graphical models. These results show a clear advantage for probabilistic models that are incorporated with a copula selection procedure. More results are presented in [18].

A general procedure for estimating the graphical model in our work has been the minimization of the Kullback-Leibler divergence between the true density and the proposed density. This procedure has also used the natural relationship between the entropy of a bivariate copula function and the marginal mutual information of two variables [8].

3. CONCLUSIONS

The use of copula functions in EDAs has opened a new research area. Our approach has been based on copula functions whose parameters are estimated by means of maximum likelihood, and the most important dependencies are represented by a graphical model.

Algorithm	Best	Median	Mean	Worst	Std. deviation
Ackley					
D-vine EDA	2.71E-005	4.11E-005	4.09E-005	5.16E-005	5.82E-006
MIMIC	2.89E-005	3.91E-005	3.92E-005	4.95E-005	5.21E-006
UMDA	2.83E-005	4.08E-005	4.15E-005	5.07 E - 005	4.63E-006
Rosenbrock					
D-vine EDA	0.20	5.75	5.28	9.21	2.70
MIMIC	0.91	6.40	6.42	13.48	2.72
UMDA	7.93	8.02	8.10	10.28	0.42
Sphere					
D-vine EDA	3.37E-007	7.75E-007	7.48E-007	9.99E-007	1.69E-007
MIMIC	4.48E-007	8.47E-007	8.14E-007	9.96E-007	1.52 E-007
UMDA	4.20E-007	8.56E-007	8.11E-007	9.94E-007	1.64E-007

Table 1: Descriptive results for the fitness of three test functions.



Figure 1: Success rate results for (a) a separable function, and (b) a non-separable function. The horizontal axis represents the dimension problem and the vertical axis represents the success rate. The solid line is used for the EDAs based on copula selection and the dashed line for the EDAs based on the Gaussian copula.

The obtained performance results suggest that the incorporation of copula functions is a real option for designing new continuous EDAs. Moreover, our reported experiments have shown that modeling adequately the dependencies between decision variables is a good mechanism for getting better performance results. Future work for our investigation will focus on the design of algorithms with strategies for enhancing their performance.

4. ACKNOWLEDGMENTS

The first author acknowledges support from the National Council of Science and Technology of México (CONACyT) through a scholarship to pursue graduate studies in the Department of Computer Science at the Center for Research in Mathematics.

5. REFERENCES

- R. Arderí-García. Algoritmo con Estimación de Distribuciones con Cópula Gaussiana. Universidad de La Habana, La Habana, Cuba, June 2007. Bachelor's thesis, in Spanish.
- [2] S. Baluja and S. Davies. Using optimal dependency-trees for combinatorial optimization: Learning the structure of the search space. In D. Fisher, editor, *Proceedings of the Fourteenth International Conference on Machine Learning*, pages 30–38. Morgan Kaufmann, 1997.
- [3] S. Barba-Moreno. Una propuesta para EDAs no paramétricos. Master's thesis, Centro de Investigación en Matemáticas, Guanajuato, México, December 2007. In Spanish.
- [4] T. Bedford and R. M. Cooke. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics* and Artificial Intelligence, 32(1):245–268, January 2001.
- [5] T. Bedford and R. M. Cooke. Vines a new graphical model for dependent random variables. *The Annals of Statistics*, 30(4):1031–1068, August 2002.
- [6] P. Bosman. Design and Application of Iterated Density-Estimation Evolutionary Algorithms. PhD thesis, University of Utrecht, Utrecht, The Netherlands, 2003.

- [7] A. Cuesta-Infante, R. Santana, J. Hidalgo, C. Bielza, and P. Larrañaga. Bivariate empirical and n-variate archimedean copulas in estimation of distribution algorithms. In WCCI 2010 IEEE World Congress on Computational Intelligence, pages 1355–1362, July 2010.
- [8] M. Davy and A. Doucet. Copulas: a new insight into positive time-frequency distributions. *Signal Processing Letters*, *IEEE*, 10(7):215–218, 2003.
- [9] J. De Bonet, C. Isbell, and P. Viola. MIMIC: Finding optima by estimating probability densities. In Advances in Neural Information Processing Systems, volume 9, pages 424–430. The MIT Press, 1997.
- [10] Y. Gao. Multivariate estimation of distribution algorithm with laplace transform archimedean copula. In W. Hu and X. Li, editors, 2009 International Conference on Information Engineering and Computer Science, ICIECS 2009, Wuhan, China, December 2009.
- [11] H. Joe. Multivariate models and dependence concepts. Chapman & Hall, London, 1997.
- [12] P. Larrañaga, R. Etxeberria, J. Lozano, and J. Peña. Optimization by learning and simulation of bayesian and gaussian networks. Technical Report KZZA-IK-4-99, Department of Computer Science and Artificial Intelligence, University of the Basque Country, 1999.
- [13] P. Larrañaga, R. Etxeberria, J. Lozano, and J. Peña. Optimization in continuous domains by learning and simulation of gaussian networks. In A. Wu, editor, *Proceedings of the 2000 Genetic and Evolutionary Computation Conference Workshop Program*, pages 201–204, 2000.
- [14] P. Larrañaga, J. Lozano, and E. Bengoetxea. Estimation of distribution algorithm based on multivariate normal and gaussian networks. Technical Report KZZA-IK-1-01, Department of Computer Science and Artificial Intelligence, University of the Basque Country, 2001.
- [15] R. Nelsen. An Introduction to Copulas. Springer Series in Statistics. Springer-Verlag, second edition, 2006.
- [16] M. Pelikan, D. Goldberg, and E. Cantú-Paz. BOA: The bayesian optimization algorithm. In W. Banzhaf, J. Daida, A. Eiben, M. Garzon, V. Honavar, M. Jakiela, and R. Smith, editors, *Proceedings of the Genetic and Evolutionary Computation Conference GECCO-99*, volume 1, pages 525–532. Morgan Kaufmann Publishers, 1999.
- [17] M. Pelikan and H. Mühlenbein. The bivariate marginal distribution algorithm. In R. Roy, T. Furuhashi, and P. Chawdhry, editors, Advances in Soft Computing -Engineering Design and Manufacturing, pages 521–535, London, 1999. Springer-Verlag.
- [18] R. Salinas-Gutiérrez, Hernández-Aguirre, and E. Villa-Diharce. Dependence trees with copula selection for continuous estimation of distribution algorithms. In GECCO '11: Proceedings of the 13th annual conference on Genetic and Evolutionary Computation, 2011. Accepted as full paper.
- [19] R. Salinas-Gutiérrez, A. Hernández-Aguirre, and E. Villa-Diharce. Using copulas in estimation of distribution algorithms. In A. Hernández Aguirre,

R. Monroy Borja, and C. Reyes García, editors, *MICAI 2009: Advances in Artificial Intelligence*, volume 5845 of *Lectures Notes in Artificial Intelligence*, pages 658–668. Springer, 2009.

- [20] R. Salinas-Gutiérrez, A. Hernández-Aguirre, and E. Villa-Diharce. D-vine eda: a new estimation of distribution algorithm based on regular vines. In *GECCO '10: Proceedings of the 12th annual* conference on Genetic and Evolutionary Computation, pages 359–366, New York, NY, USA, 2010. ACM.
- [21] R. Santana-Hermida. Modelación probabilística basada en modelos gráficos no dirigidos en Algoritmos Evolutivos con Estimación de Distribuciones. PhD thesis, Instituto de Cibernética, Matemáticas y Física, La Habana, Cuba, February 2004. In Spanish.
- [22] S. Shakya. DEUM: A framework for an Estimation of Distribution Algorithm based on Markov Random Fields. PhD thesis, The Robert Gordon University, Aberdeen, United Kingdom, April 2006.
- [23] A. Sklar. Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris, 8:229–231, 1959.
- [24] M. Soto, A. Ochoa, S. Acid, and L. de Campos. Introducing the polytree approximation of distribution algorithm. In A. Ochoa, M. Soto, and R. Santana, editors, Second International Symposium on Artificial Intelligence. Adaptive Systems. CIMAF99, pages 360–367, La Habana, 1999. Academia.
- [25] P. Trivedi and D. Zimmer. Copula Modeling: An Introduction for Practitioners, volume 1 of Foundations and Trends[®] in Econometrics. Now Publishers, 2007.
- [26] L. Wang, X. Guo, J. Zeng, and Y. Hong. Using gumbel copula and empirical marginal distribution in estimation of distribution algorithm. In *Third International Workshop on Advanced Computational Intelligence, IWACI 2010*, pages 583–587. IEEE, August 2010.
- [27] L. Wang, Y. Wang, J. Zeng, and Y. Hong. An estimation of distribution algorithm based on clayton copula and empirical margins. In *Life System Modeling and Intelligent Computing*, pages 82–88. Springer, September 2010.
- [28] L. Wang and J. Zeng. Estimation of distribution algorithm based on copula theory. In Y. Chen, editor, *Exploitation of Linkage Learning in Evolutionary Algorithms*, volume 3 of *Adaptation, Learning, and Optimization*, pages 139–162. Springer, 2010.
- [29] L. Wang, J. Zeng, and Y. Hong. Estimation of distribution algorithm based on archimedean copulas. In GEC '09: Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation, pages 993–996, New York, NY, USA, June 2009. ACM.
- [30] L. Wang, J. Zeng, and Y. Hong. Estimation of distribution algorithm based on copula theory. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pages 1057–1063. IEEE Press, May 2009.
- [31] L. Wang, J. Zeng, Y. Hong, and X. Guo. Copula estimation of distribution algorithm sampling from clayton copula. *Journal of Computational Information Systems*, 6(7):2431–2440, 2010.