The Meta-Analytics Innovation: Unifying Metaheuristics and Analytics*

Fred Glover
Meta-Analytics, Inc., 2605 Stanford Ave., Boulder, Colorado 80305
fredwglover@yahoo.com

Carlos Cotta
ETSI Informática, Universidad De Málaga, Malaga, Spain
ccottap@lcc.uma.es

Abstract. Meta-Analytics represents the unification of metaheuristics and analytics, two fields of the foremost interest and practical importance. While metaheuristics provide a modern framework and an arsenal of cutting-edge techniques to handle complex, real-world problems, Analytics embodies the use of prediction and optimization techniques in practical contexts. Thus, their marriage can be regarded as a natural step towards both the creation of effective tools for problems in the Analytics domain and the expansion of the scope of metaheuristic techniques. This introductory chapter describes the advantages obtained by the synergies of the techniques and the avenues for achieving such a unification of methodologies, and discusses some important themes in the field. We also introduce contributions contained in this section, in which these themes are explored in more detail.

*Invited Chapter, *New Ideas in Business and Customer Analytics*, N.J. de Vries and P. Moscato (Eds.), Springer, 2019.

1 Introduction

The Meta-Analytics theme has its origins in a series of seminal developments in optimization and machine learning and their practical applications. The term “Analytics” has gained unprecedented recognition as a referent for analyses that embody prediction and optimization in a broad sense, typically supported by interpretive aids for users. As a result, organizations from a wide range of disciplines have allied themselves with the Analytics area. This includes researchers and practitioners in classical optimization, notably in the fields of engineering, computer science, operations research and management science. As a compelling example, the prestigious *Institute of Management Science and Operations Research*[1] (INFORMS) has adopted the area of Analytics as a primary focus, and has created a new magazine called *Analytics*[2].

Metaheuristics (Glover, 1986; Blum and Roli, 2003; Glover and Kochenberger, 2003; Sörensen et al, 2017), by contrast, emerged from the dawning recognition that many real world problems in business, science and industry are too large or too complex for classical optimization methods to handle effectively. To remedy this problem, recourse was initially made to joining classical optimization with simple heuristic methods, but it soon became clear that more powerful approaches were needed â€“ incorporating various ideas of heuristics, but going beyond them. The methods of metaheuristics were conceived to meet this challenge, with innovative evolutionary and neighborhood search approaches whose first forms appeared in the late 1960s and early 1970s, and which have since undergone substantial refinements and modifications. The name “metaheuristics” itself emerged in the mid-1980s, as the recognition of the essential focus of these new methods became universal.

Meta-Analytics represents the, perhaps inescapable, unification of Metaheuristics and Analytics. While the former provide a modern framework and an arsenal of cutting-edge techniques to handle complex, real-world problems, the latter embodies the use of prediction and optimization techniques in practical contexts. Thus, their marriage can be only regarded as a natural step towards both the creation of effective tools for problems in the Analytics domain and the expansion

[1] https://www.informs.org/
[2] http://analytics-magazine.org/

of the scope of metaheuristic techniques. Indeed, modern versions of these latter techniques, such as evolutionary algorithms (Eiben and Smith, 2003), tabu search (Glover and Laguna, 1997), simulated annealing (Dekkers and Aarts, 1991), swarm intelligence algorithms (Kennedy and Eberhart, 2001), memetic algorithms (Neri et al, 2012; Cotta et al, 2016) and a variety of others which have proved to be highly successful, producing an explosion of publications in international journals and presentations at international conferences. These developments have additionally resulted in the formation of new journals and new societies. The ability to deal with challenging practical problems more effectively, including those from domains that Analytics claims as its focus, lays a foundation for an alliance between Metaheuristics and Analytics. This is notably exemplified by the fact that the Metaheuristics field has made important contributions to predictive and prescriptive analysis, which are prominent concerns of Analytics.

The unification of Metaheuristics and Analytics within Meta-Analytics brings about important advantages that were not fully realized in the past as these two fields evolved largely in isolation from each other. Chief among these are the promise of Metaheuristics to become a source of more effective tools for problems in the Analytics domain, and in turn the promise of Analytics to provide a perspective for expanding the scope of algorithmic methods within Metaheuristics. These potentials are accentuated by the fact that many researchers and practitioners in Analytics have not been exposed to the Metaheuristics field, and are unaware of its power for addressing practical applications, while many of those working within Metaheuristics have incompletely appreciated the value of incorporating elements that have become the purview of Analytics. The establishment of Meta-Analytics creates an opportunity to reach an expanded community of decision makers in industry, science and government who can profit from the union of its component areas.

The scope of this union can be glimpsed by elucidating the primary themes of Meta-Analytics and by looking at some previous approaches that have been paving the way. Subsequently, we shall describe the advantages obtained by the synergies of the techniques and the avenues for achieving such a unification of methodologies. We also introduce contributions contained in this section, in which these themes are explored in more detail.

2 Themes of Meta-Analytics

Six themes broadly constitute the main thrusts of the Meta-Analytics area:

1. Using Metaheuristics as a source for creating enhanced predictive and machine learning methods (as in clustering, discrimination, feature detection, pattern recognition and classification, etc.). While such concerns have long been a part of the metaheuristic domain, a more dedicated emphasis on them through Meta-Analytics lays a foundation for significant new advances. A key source of contributions derives from highlighting such advances for their general relevance to Analytics, and hence to the Metaheuristics/Analytics union.

2. Incorporating predictive and machine learning methods to enhance the performance of metaheuristics. In spite of a variety of proposals for exploiting learning within metaheuristics, e.g., see (Glover and Greenberg, 1989; Kelly et al, 1996; Birattari, 2009), very little has been done to pursue this theme. Predictive and machine learning procedures can be applied with metaheuristics in offline (pre-solution and post-solution) stages as well as during run time execution. An important step forward will be supplied by refining and implementing meritorious ideas which have been inadequately investigated, and by developing new proposals to capitalize on the opportunities opened up by joining analytics and metaheuristics.

3. Creating special mechanisms and interfaces for interpreting outcomes and relationships uncovered by metaheuristic solution processes. This focus has the goal of enabling users to interact with Meta-Analytic procedures to achieve greater insight and yield better decisions. This interaction includes adaptive exploration of model assumptions as well as decision rules for guiding the methods studied (Meignan et al, 2015).

4. Developing integrative methods that capitalize on one or more of the preceding themes to build highly effective algorithms that utilize domain knowledge for solving problem from important classes. Tabu search (Glover and Laguna, 1997) and Memetic algorithms (Neri et al, 2012; Cotta et al, 2016) are good examples of this theme.

5. Creating improved methods for analyzing and explaining the operation of alternative solution approaches, including more effective and comprehensive forms of landscape analysis and quasi-decomposition analysis as embodied in vocabulary building strategies.

6. Establishing a repository of important applications in business, science and government where Meta-Analytics provides advances of singular value, leading to improved insights, operations and policies.

3 Some Important Research Avenues

In this section we will explore some research lines that intersect with the area of meta-analytics and are playing (or can play in the near future) a major role in the field. We will highlight such connections as well as their relevance for meta-analytics.

3.1 Ensemble Learning

An ensemble of learning machines is a set of adaptive entities that deliver partial solutions to a given problem, and then integrate these solutions in some manner to construct a final or complete solution to the original problem. Recent advances in machine learning theory have highlighted the importance of understanding why the collective behavior of such a collection of several learning agents can perform substantially better than individual ones. Whereas empirical studies on classification methods have shown that some classifiers perform best in some domains but not in all application domains –a phenomenon linked to the *"No Free Lunch" theorem* by (Wolpert, 1996)–, ensemble methods can purportedly overcome these limitations to some extent, by combining the output of many independent classifiers. It is perhaps a good analog with hybrid metaheuristics (Raidl, 2006), i.e., integrative or collaborative approaches aiming to combining the virtues of different algorithmic approaches to an optimization problem, so that some sort of "emergent phenomena" comes out of the synergy of the techniques employed by the individual components.

These new learning methods have been called "meta-learning schemes" or "meta-classifiers" or "ensembles". In the machine-learning paradigm, ensemble data mining methods strategically advance the power of committee methods, or combine models to achieve better prediction accuracy than any of the individual models could achieve (Oza, 2006). The basic goal when designing an ensemble is to develop it in such a way that it provides independent models whose combination will produce better performance than the individual models in isolation. This involves several strategic decisions such as:

- Accounting for a diverse set of outputs: the diversity of the outputs of individual classifiers in an ensemble is a key issue to the generalization performance of the group as a whole. Consequently, a strategically combination of diverse classifiers can help to reduce the total error (Polikar, 2006). This can be attained (i) by combining different classifiers, (ii) by having each classifier trained with different subsets of data (i.e., by performing horizontal partitions of the data), (iii) by having each classifier trained using different features of the data (i.e., by performing vertical partitions of the data), or (iv) any combination of these approaches among other possibilities.

- Designing an appropriate combination rule: The combination rule leads to a final classification from all participating single classifier's outcome and can be mainly designed in two major ways: (i) train the classifiers over the entire feature space and use a fusion method to integrate their outputs (e.g., [weighted] majority voting, summation, product, etc.) and (ii) use domain expert classifiers (trained to become an expert in a specific part of the total feature space) and pick one depending on its competence, measured either statically (during training) or dynamically (during prediction).

Ensemble learning is relevant to meta-analytics in several regards. On one hand, obtaining the best ensemble from a collection of classifiers (i.e., solving the design decisions sketched above) is an NP-hard problem (Hernández-Lobato et al, 2006), and hence researchers have commonly applied different types of metaheuristic strategies to this end – see, e.g., (Gaber and Bader-El-Den, 2012; Lertampaiporn et al, 2013; Zhang et al, 2015; Haque et al, 2016). On the other hand, they have been routinely applied with success in application domains related to analytics such as customer churn prediction (Xiao et al, 2016), purchasing and marketing (Govindarajan, 2015), predictive analytics (Wang and Wu, 2017) and business process management (Zhao et al, 2016), jut to cite a few.

3.2 Simulation-based Optimization

The sustained increase of the computational power available for scientific purposes is continuously opening new possibilities for investigating systems that were out of reach not so long ago. Particularly, it allows obtaining increasingly accurate simulations of complex systems and processes from the analytics domain, which can in turn be used to optimize these. The term *simulation-based optimization* precisely refers to these kinds of approaches (April et al, 2004, 2006; Better et al, 2007, 2008; Chen and Lee, 2010; Pasupathy and Ghosh, 2013; Better et al, 2015; Amaran et al, 2016; Thengvall et al, 2016).

Metaheuristics constitute the natural tool to tackle the optimization of a simulation system, since black-box optimization is one of the most distinctive realms in which these techniques excel. This does not mean simulation-optimization is a trivial quest though. Simulation models typically exhibit many features that can pose difficulties to the optimization process:

- Non-deterministic behavior: the simulation may incorporate stochastic elements (if only by the presence of some noise in the output variables), resulting in a different outcome in each run. The optimization algorithm must take this into account in order to handle the resulting uncertainty, e.g. (Mininno and Neri, 2010; Tenne, 2012). Some authors have advocated for the term *simeheuristics* to denote approaches that integrate simulation (in any of its variants) into a metaheuristic-driven framework to solve complex stochastic combinatorial optimization problems, see (Juan et al, 2015).

- Granularity of the simulation: it is possible to attain different tradeoffs between the computational cost of the simulation and the accuracy/uncertainty of the output. Going beyond, it is possible to use machine-learning to create surrogate models to be optimized in lieu of the (more computationally expensive) simulation, e.g., (Queipo et al, 2005; April et al, 2006; Better et al, 2007; Han and Zhang, 2012).

- Chaotic behavior: although profoundly different in nature to the issue of noise, the presence of a chaotic regime in the simulated system –e.g., see (Lal and Swarup, 2011)– may offer similar challenges to the optimization technique in terms of uncertainty.

Some connections between simulation optimization and machine learning methods can be also drawn. For example, in active learning approaches (Settles, 2012) the algorithms are allowed to query an oracle for additional data to infer better statistical models, much like simulation optimization methods may take the choice of sampling at each iteration of the simulation, cf. (Better et al, 2015; Amaran et al, 2016).

3.3 Multi-objective Optimization and Analytics

Many problems in the area of analytics are naturally multi-objective: they exhibit multiple cost/benefit functions in partial conflict with each other. Hence, there is often no single optimal solution but a (potentially huge) collection of *non-dominated* or *efficient* solutions providing different optimal tradeoffs between the objectives. This is an issue that has been thoroughly analyzed from the point of view of metaheuristics: nowadays, we have a huge arsenal of techniques and methods by which standard (i.e., mono-objective) optimization methods can be augmented in order to tackle many-objective problems, e.g., (Coello Coello and Lamont, 2004; Jaszkiewicz et al, 2012).

However, no matter how successful these approaches can be in fulfilling their mission, it must be noted that the latter is often finding a large and broad sample of the Pareto front. This is useful only as long as there is a sensible decision-making policy whereby an appropriate solution can be extracted from this Pareto front. This can be a hard problem in itself, involving many of the themes mentioned before such as model building, preference learning, handling large collections of data, etc. Thus, this is an area in which much cross-fertilization between meta-analytics applications is possible.

4 Introducing the Contents of Chapters in this Section

The chapters in this section make several important contributions to Meta-Analytics as instances of the foregoing themes.

The chapter "From Ensemble Learning to Meta-Analytics: A Review on Trends in Business Applications," by Mohammad Nazmul Haque and Pablo Moscato, provides a broad overview of the area of ensemble learning, with particular emphasis on its application to analytics. The authors outline the guiding principles for the construction of an ensemble of learners and the issues that ought to be considered. Then, they proceed to overview the deployment of these approaches on four successful application areas: purchasing and marketing, predictive analytics, business process management and customer churn prediction (this latter area is tackled more in depth in a subsequent chapter by Haque et al.) Particular attention is paid to the interface of ensemble methods and metaheuristics, identifying common intersection points and popular methodological approaches. These issues will be further re-elaborated in a subsequent chapter by Thomschke et al.

The chapter “A Multi-objective Meta-Analytic Method for Customer Churn Prediction,” authored by Mohammad Nazmul Haque, Natalie Jane de Vries and Pablo Moscato, takes its starting point from the observation that researchers typically use single-objective optimization for the important area of ensemble learning in analytics. The authors go a step farther by introducing a metaheuristic multi-objective evolutionary algorithm, which they apply to the problem of customer churn prediction. The authors additionally investigate a complementary symbolic regression-based approach, noting that the multi-objective approach excels at prediction while the symbolic regression-based approach offers useful tools or business analysts. The results of their study demonstrate how combining their new multi-objective approach with symbolic regression analysis is effective for the paired goals of predicting those customers likely to churn and of providing insight into the types of resources companies can invest in to accurately predict churners and prevent them from churning.

Addressing ensemble learning from a different perspective, the chapter “Metaheuristics and Classifier Ensembles,” by Ringolf Thomschke, Stefan Voss and Stefan Lessmann, fills a major gap resulting from the fact that previous studies have failed to consider more than a small number of strategies for ensemble member selection. The authors introduce a comprehensive set of metaheuristics to create alternative ensemble classifiers, and carry out an empirical study to compare the outcomes of applying these alternative classifiers. Based on this study, they identify a highly promising modeling approach and compare it to other ensemble regimes and prediction models. They find that their metaheuristic-based ensemble approach improves upon the state-of-the-art, and in the process also introduce a new method to approximate an optimality gap for predictive classification, leading to promising avenues for future research.

Finally, the chapter “Hotel Classification using Meta-Analytics: a Case Study with Cohesive Clustering,” by Buyang Cao, Cesar Rego and Fred Glover, provides a tailored meta-analytic procedure by developing a new clustering algorithm to address the challenge of analyzing data sets of hotel ratings. The study concerns itself with the commonly occurring situation in clustering applications where coordinates are unknown and only distances between objects are available. A tabu search metaheuristic is employed to assure clusters that manifest a cohesiveness property, which the authors join with a new form of hierarchical clustering for additional refinement.

Computational experiments demonstrate that the algorithm is both robust and extensible, and is suitable for running on a backend to perform the corresponding tasks with little human intervention. Future research is proposed to exploit the susceptibility of the approach to parallel processing.

Acknowledgements

The authors wish to thank Mohammad Nazmul Haque and Pablo Moscato in connection to the subsection on ensemble learning. Carlos Cotta acknowledges support from Spanish Ministry of Economy and Competitiveness and European Regional Development Fund (FEDER) under project EphemeCH (TIN2014-56494-C4-1-P). This research was also supported in part by the Key Laboratory of International Education Cooperation of Guangdong University of Technology.

References

Amaran S, Sahinidis NV, Sharda B, Bury SJ (2016) Simulation optimization: a review of algorithms and applications. Annals of Operations Research 240(1):351–380

April J, Glover F, Kelly J, Laguna M (2004) The exploding domain of simulation optimization. Newsletter of the INFORMS Computing Society 24(2):1–14

April J, Better M, Glover F, Kelly J, Laguna M (2006) Enhancing business process management with simulation-optimization. In: Perrone LF, Wieland FP, Liu J, Lawson BG, Nicol DM, Fujimoto R (eds) 2006 Winter Simulations Conference, IEEE Press

Better M, Glover F, Laguna M (2007) Advances in analytics: Integrating dynamic data mining with simulation optimization. IBM Journal of Research and Development 51(3/4):477–487

Better M, Glover F, Kochenberger G, Wang H (2008) Simulation optimization: Applications in risk management. International Journal of Information Technology & Decision Making 7(4):571–587

Better M, Glover F, Kochenberger G (2015) Simulation optimization to improve decisions under uncertainty. In: Cox T (ed) Breakthroughs in Decision Science and Risk Analysis, Wiley Publishing, pp 59–62

Birattari M (2009) Tuning Metaheuristics. A Machine Learning Perspective, Studies in Computational Intelligence, vol 197. Springer, Berlin Heidelberg

Blum C, Roli A (2003) Metaheuristics in combinatorial optimization: Overview and conceptual comparison. ACM Computing Surveys 35(3):268–308

Chen CH, Lee LH (2010) Stochastic simulation optimization: An optimal computing budget allocation. System engineering and operations research, World Scientific Publishing Company, Singapore

Coello Coello C, Lamont G (2004) Applications of Multi-Objective Evolutionary Algorithms. World Scientific, New York

Cotta C, Gallardo J, Mathieson L, Moscato P (2016) A contemporary introduction to memetic algorithms. In: Wiley Encyclopedia of Electrical and Electronic Engineering, Wiley, pp 1–15

Dekkers A, Aarts E (1991) Global optimization and simulated annealing. Mathematical Programming 50:367–393

Eiben AE, Smith JE (2003) Introduction to Evolutionary Computation. Natural Computing Series, Springer-Verlag, Berlin Heidelberg

Gaber MM, Bader-El-Den M (2012) Optimisation of Ensemble Classifiers using Genetic Algorithm. In: Graña M, Toro C, Posada J, Howlett RJ, Jain LC (eds) Advances in Knowledge-Based and Intelligent Information and Engineering Systems, IOS Press

Glover F (1986) Future paths for integer programming and links to artificial intelligence. Computers and Operations Research 13(5):533–549

Glover F, Greenberg H (1989) New approaches for heuristic search: A bilateral linkage with artificial intelligence. European Journal of Operational Research 39(2):119–130

Glover F, Kochenberger G (eds) (2003) Handbook of Metaheuristics, International Series in Operations Research and Management Science, vol 57, 1st edn. Kluwer Academics Publishers, Boston MA

Glover F, Laguna M (1997) Tabu Search. Kluwer Academic Publishers, Norwell, MA

Govindarajan M (2015) Comparative study of ensemble classifiers for direct marketing. Intelligent Decision Technologies 9(2):141–152

Han ZH, Zhang KS (2012) Surrogate-based optimization. In: Roeva O (ed) Real-World Applications of Genetic Algorithms, InTech

Haque MN, Noman N, Berretta R, Moscato P (2016) Heterogeneous ensemble combination search using genetic algorithm for class imbalanced data classification. PLoS ONE 11(1):e0146,116, 10.1371/journal.pone.0146116

Hernández-Lobato D, Martínez-Muñoz G, Suárez A (2006) Pruning in ordered regression bagging ensembles. In: The 2006 IEEE International Joint Conference on Neural Network Proceedings, pp 1266–1273

Jaszkiewicz A, Ishibuchi H, Zhang Q (2012) Multiobjective memetic algorithms. In: Neri F, Cotta C, Moscato P (eds) Handbook of Memetic Algorithms, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 201–217

Juan AA, Faulin J, Grasman SE, Rabe M, Figueira G (2015) A review of simheuristics: Extending metaheuristics to deal with stochastic combinatorial optimization problems. Operations Research Perspectives 2:62 – 72

Kelly J, Rangaswamy B, Xu J (1996) A scatter-search-based learning algorithm for neural network training. Journal of Heuristics 2(2)

Kennedy J, Eberhart R (eds) (2001) Swarm Intelligence. Morgan Kaufmann Publishers, San Francisco, CA, USA

Lal DK, Swarup K (2011) Modeling and simulation of chaotic phenomena in electrical power systems. Applied Soft Computing 11(1):103 – 110

Lertampaiporn S, Thammarongtham C, Nukoolkit C, Kaewkamnerdpong B, Ruengjitchatchawalya M (2013) Heterogeneous ensemble approach with discriminative features and modified-SMOTEbagging for pre-miRNA classification. Nucleic acids research 41(1):e21

Meignan D, Knust S, Frayret JM, Pesant G, Gaud N (2015) A review and taxonomy of interactive optimization methods in operations research. ACM Trans Interact Intell Syst 5(3):17:1–17:43

Mininno E, Neri F (2010) A memetic differential evolution approach in noisy optimization. Memetic Computing 2(2):111–135

Neri F, Cotta C, Moscato P (eds) (2012) Handbook of Memetic Algorithms, Studies in Computational Intelligence, vol 379. Springer-Verlag, Berlin Heidelberg

Oza NC (2006) Ensemble data mining methods. In: Wang J (ed) Encyclopedia of Data Warehousing and Mining, Idea Group Reference, vol 1, pp 448–453

Pasupathy R, Ghosh S (2013) Simulation optimization: A concise overview and implementation guide. Tutorials in Operations Research 10:122–150

Polikar R (2006) Ensemble based systems in decision making. Circuits and Systems Magazine, IEEE 6(3):21–45

Queipo NV, Haftka RT, Shyy W, Goel T, Vaidyanathan R, Tucher PK (2005) Surrogate-based analysis and optimization. Progress in Aerospace Sciences 41:1–28

Raidl GR (2006) A unified view on hybrid metaheuristics. In: Almeida F, Aguilera MJB, Blum C, Moreno-Vega JM, Pérez MP, Roli A, Sampels M (eds) Hybrid Metaheuristics – HM 2006, Springer, Lecture Notes in Computer Science, vol 4030, pp 1–12

Settles B (2012) Active Learning. Synthesis Lectures on Artificial Intelligence and Machine Learning, Morgan & Claypool Publishers

Sörensen K, Sevaux M, Glover F (2017) A history of metaheuristics. In: Martí R, Pardalos P, Resende M (eds) Handbook of Heuristics, Springer, (available at arXiv:1704.00853 [cs.AI])

Tenne Y (2012) Memetic algorithms in the presence of uncertainties. In: Neri F, Cotta C, Moscato P (eds) Handbook of Memetic Algorithms, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 219–237

Thengvall B, Glover F, Davino D (2016) Coupling optimization and statistical analysis with simulation models. In: Roeder TMK, Frazier PI, Szechtman R, Zhou E, Huschka T, Chick SE (eds) 2016 Winter Simulation Conference, IEEE Press, pp 545–553

Wang L, Wu C (2017) Business failure prediction based on two-stage selective ensemble with manifold learning algorithm and kernel-based fuzzy self-organizing map. Knowl-Based Syst 121:99–110

Wolpert DH (1996) The lack of a priori distinctions between learning algorithms. Neural computation 8(7):1341–1390

Xiao J, Jiang X, He C, Teng G (2016) Churn prediction in customer relationship management via gmdh-based multiple classifiers ensemble. IEEE Intelligent Systems 31(2):37–44

Zhang L, Wang X, Moon WM (2015) PolSAR images classification through GA-based selective ensemble learning. In: Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International, pp 3770–3773

Zhao W, Liu H, Dai W, Ma J (2016) An entropy-based clustering ensemble method to support resource allocation in business process management. Knowl Inf Syst 48(2):305–330