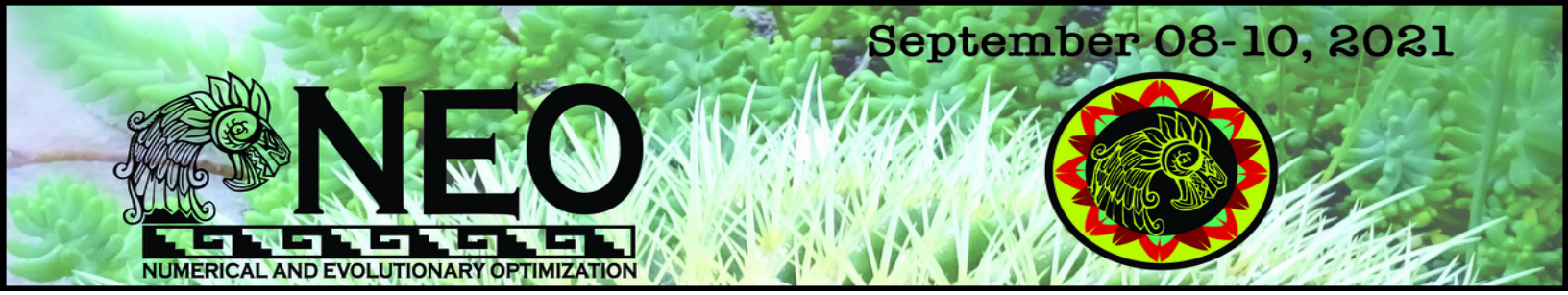
NEO 2021

9th International Workshop on Numerical and Evolutionary Optimization
September 8-10 2021

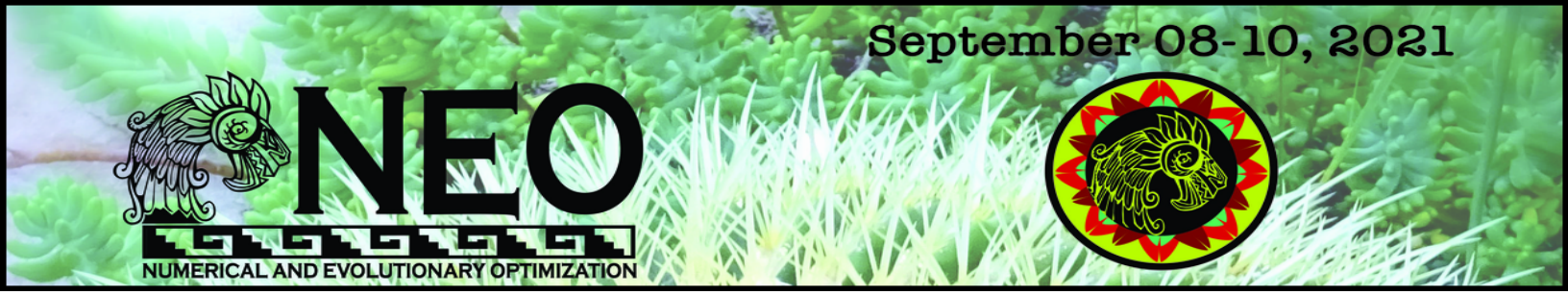
General Chairs:

Luis Gerardo de la Fraga CINVESTAV-IPN

Marcela Quiroz Artificial Intelligence Research Institute, UV



Contents



Foreword

Welcome

Welcome to NEO 2021, the 9th International Workshop on Numerical and Evolutionary Optimization. In this edition, NEO 2021 occurs from September 08 to 10, 2021. As in the last year, the workshop will be held as online only event due to the on-going situation caused by Covid-19.

The goal of the Numerical and Evolutionary Optimization (NEO) workshop series is to bring together people from these and related fields to discuss, compare and merge their complementary perspectives. NEO encourages the development of fast and reliable hybrid methods, that maximize the strengths and minimize the weaknesses of each underlying paradigm; while also applying to a broader class of problems. Moreover, NEO fosters the understanding and adequate treatment of real-world problems, particularly in emerging fields that affect us all such as healthcare, smart cities, big data, among many others.

In this 2021 edition, there will be more than 80 participants coming from all over Mexico. The NEO 2021 will have more than 35 technical presentations addressing different optimization subjects and dealing with a variety of challenging applications. Finally, it is worth to notice that over the years, the NEO community has grown not only in number but also concerning gender diversity. This year we arrange also, as now it is common since 2018, a special session called Women at NEO (W-NEO) as an effort to make more visible the work in optimization performed by our female pairs. We hope you enjoy your participation at the NEO 2021, thank you for your valuable assistance.

Sincerely,

Marcela Quiroz, Universidad Veracruzana

Luis Gerardo de la Fraga, Cinvestav-IPN

NEO 2021 General Chairs

Acknowledgments

We want to thank all participants that helped to make the NEO 2021 such a great success. In particular, we would like to thank Keynote Speakers:
Wolfgang Banzhaf, Michigan State University, USA;
Gustavo Olague, CICESE, Mexico;
Carlos Segura, CIMAT, Mexico; and
Kenneth Sörensen, University of Antwerp, Belgium.

Further, we thank to the institutions
CONACYT,
CINVESTAV-IPN, and
Tecnológico Nacional de México/IT, in Tijuana;

to the academic colleagues
Dr. Daniel E. Hernández Morales (TNM/IT Tijuana)
Oliver Cuate (ESFM-IPN),
Guillermo Morales (Cinvestav-IPN),
Brisbane Ovilla (Cinvestav-IPN);

and to the staff members
Felipa Rosas (Cinvestav-IPN),
Santiago Dominguez (Cinvestav-IPN),
Erika Rios (Cinvestav-IPN),
Jose-Luis Flores (Cinvestav-IPN),
Adriana Martinez (Kreaprom SA de CV and Impakt 45 SA de CV), and
Everett Zhu (MDPI).

NEO 2021 Organizing Committee

Partners

Cinvestav

Universidad Veracruzana

IPN

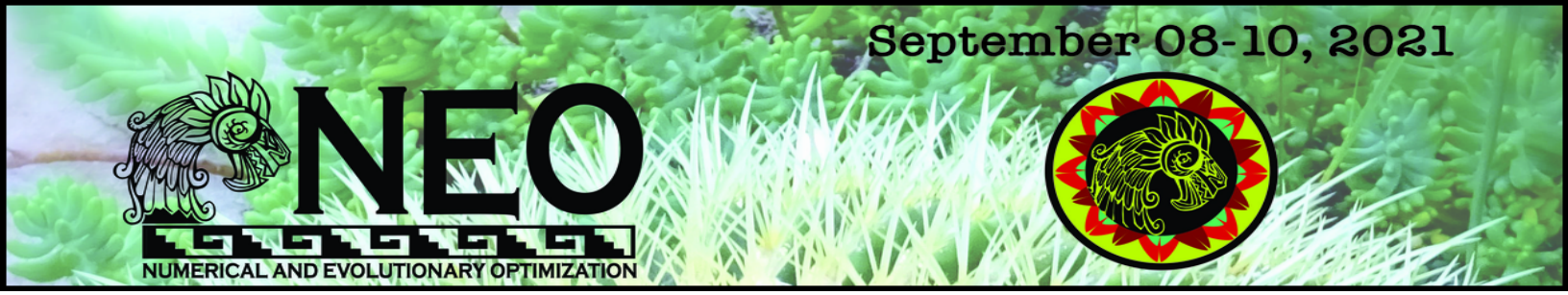
UNIVERSIDAD DE LA SIERRA JUÁREZ
UNSIJ
OAXACA

TIJUANA, BAJA CALIFORNIA

CONACYT
Consejo Nacional de Ciencia y Tecnología

MDPI

TREE LAB



Schedule NEO 2021

Each session ends with a **Session Wrap Up**. Times are 20 minutes per talk, as well as for the wrap ups.

Day 1, September 08, 2021
Research Experience Day

9:15 – 9:30 Welcome to the RED
9:30 – 10:30 Postgrado, la puerta a las oportunidades.
Dr. Carlos Ignacio Hernández, IIMAS-UNAM
10:30 – 11:00 ¿Qué es la optimización y con qué se lleva bien? **Dra. Adriana Lara**, ESFM-IPN
11:00 – 11:30 Una introducción a las heurísticas. **Dra. Marcela Quiroz**, UV
11:30 – 12:00 Ingeniería con programación genética. **Dr. Leonardo Trujillo**, ITT
12:00 – 13:00 Discussion panel.
13:00 – 13:15 Closing
20:00 – 22:00 **Women at NEO**

Day 2, September 09, 2021

09:00 – 09:20 Opening
09:20 – 10:40 Keynote I: **Wolfgang Banzhaf**
Evolution Sometimes Needs Guidance 15
10:40 – 10:50 Coffee break
10:50 – 12:10 Sessions I and II (2 in parallel, 3 talks each)
12:10 – 12:20 Coffee break
12:20 – 13:40 Sessions III and IV
14:40 – 15:00 Lunch break
15:00 – 16:20 Keynote II: **Gustavo Olague**
Brain Programming Vs. Neuroevolution Visiting the Old Fight Between Symbolic and Sub-symbolic Learning 16
16:20 – 16:30 Coffee break
16:30 – 17:50 Sessions V and VI

Session I: 10:50 – 12:10, Room 1, **Covid-19**
Chair: Nelson Rangel-Valdez
(ID 38, in page 23) **Alejandro Castellanos-Alvarez, Laura Cruz, Hector Joaquin Fraire Huacuja, Claudia Gomez, José Alfredo Brambila-Hernández, Nelson Rangel-**

Valdez and Jesús M. Velásquez-Bermúdez. A multi-objective optimization method to estimate parameters of the COVID SEIMR/R-S model with a case study of Mexico.

(ID 42, in page 26) **Miriam Pescador Rojas, Víctor Adrián Sosa-Hernández, and José L. Maravillas-Montero**. Explainable artificial intelligent model to determine the severity on COVID-19 patients.

(ID 28, in page 27) **Angel Iram García Fernández, Nelson Rangel Valdez, Claudia Guadalupe Gómez Santillán, Héctor Joaquín Fraire Huacuja, Laura Cruz Reyes and María Lucila Morales Rodríguez**. Theoretical analysis of the application of a university timetable model in the optimization of hospital resources.

Session II: 10:50 – 12:10, Room 2, **Multi-objective Optimization**
Chair: Günter Rudolph

(ID 45, in page 28) **Fernando Moreno Gómez and Adriana Lara López**. On the seek of an affordable solution to enhance Mexico City's air quality by optimized tree seeding.

(ID 19, in page 29) **Cristian Sandoval, Oliver Cuate, Luis C. González, Leonardo Trujillo and Oliver Schütze**. Simple regression models for the rapid approximation of the hypervolume indicator for multiobjective optimization problems.

(ID 1, in page 30) **Oliver Schütze and Carlos Hernández**. Archiving Strategies for multi-objective evolutionary algorithms.

Session III: 12:20 – 13:40, Room 1, **Circuits and ML I**
Chair: Gerardo de la Fraga

(ID 27, in page 31) **Luis Gerardo De La Fraga, Esteban Tlelo-Cuautle and Martín Alejandro Valencia-Ponce**. Chaotic Oscillators Implemented in FP16 Numbers.

(ID 18, in page 32) **Anibal Sosa, Jhoan Delgado and Maria de Los Angeles González Lima**. Subsample algorithm for Support Vector Regression.

(ID 30, in page 33) **Luis Gerardo De La Fraga**. Differential Evolution with Mixed Precision Numbers.

Session IV: 12:20 – 13:40, Room 2, Optimization in Industry 1
Chair: Oliver Cuate

(ID 20, in page 34) **Alfonso Navarro-Espinoza, Oscar Roberto López-Bonilla, Enrique Efrén García-Guerrero, Esteban Tlelo-Cuautle, Didier López-Mancilla, Carlos Hernández-Mejía and Everardo Inzunza-Gonzáleza**. Traffic flow prediction for smart traffic lights.

(ID 21, in page 36) **Edgar Rene Ramos-Acosta, Enrique Efrén García-Guerrero, Oscar Roberto López-Bonilla, Esteban Tlelo-Cuautle, Didier López-Mancilla, Carlos Hernández-Mejía and Everardo Inzunza-González**. Assembly components detection system based on convolutional neural networks: a case study.

(ID 15, in page 38) **Javier Carmona, Josué Enríquez-Zárate, Armando Cárdenas, Daniel Hernández and Leonardo Trujillo**. Detection of wind turbine blade damage caused by insect erosion using AutoML methods.

Session V: 16:30 – 17:50, Room 1, Circuits and ML 2
Chair: Esteban Tlelo-Cuautle

(ID 43, in page 39) **Gerardo de Jesús Martínez Neri, Luis Fortino Cisneros Sinen-**

cio, Juan Frausto Solis and Javier González Barbosa. Computational methods for the design of transconductance operational amplifiers.

(ID 17, in page 41) **Joel Nation, Leonardo Trujillo, Luis Muñoz and Edgar Galván**. Analysis of transfer learning in a genetic programming system using DeepInsight.

(ID 16, in page 42) **Luis A Cardenas F, Daniel E Hernández, Leonardo Trujillo, Perla Juárez and Jose M. Muñoz**. Implementation of a stack-based genetic programming algorithm in CUDA.

Session VI: 16:30 – 17:50, Room 2, Discrete Optimization 1
Chair: Guadalupe Carmona

(ID 13, in page 44) **Ricardo Pérez-Rodríguez**. A hybrid estimation of distribution algorithm for the quay crane scheduling problem.

(ID 31, in page 45) **Carlos Corpus, José Camacho and Víctor Blanco**. Un algoritmo coevolutivo para el problema binivel de localización del árbol de hubs con fijación de precios.

(ID 35, in page 46) **Diego Soto Monterrubio, Juan Paulo Sánchez-Hernández, Juan Frausto Solis and Javier González Barbosa**. Roulette selection strategies applied to GRSA2-SSP for refinement of the amino acid side chain regions.

Day 3, September 10, 2021

09:00 – 10:20 Keynote III: **Kenneth Sorensen**
A history of Metaheuristics 18
10:20 – 10:30 Coffee break
10:30 – 11:50 Sessions VII and VIII
11:50 – 12:00 Coffee break
12:00 – 13:20 Sessions IX and X
13:20 – 15:00 Lunch break
15:00 – 16:20 Keynote IV: **Carlos Segura**
Design of Population-based Optimizers: Diversity, Data Structures and Algorithms 19
16:20 – 16:30 Coffee break
16:30 – 17:50 Sessions XI and XII
17:50 – 18:00 Closing

Session VII: 10:30 – 11:50, Room 1, Women at NEO
Chair: Miriam Pescador

(ID 40, in page 47) **Dalia Andra Rodriguez Ulloa, Arnoldo Díaz-Ramírez, Rafael Iván Ayala Figueroa, Leonardo Trujillo Reyes and Juan Pablo García Vázquez**. Addressing violence against women and children using IoT, AI, and serious games.

(ID 34, in page 48) **Guadalupe Carmona Arroyo, Marcela Quiroz Castellanos and Efrén Mezura-Montes**. Variable decomposition for large-scale constrained optimization problems using a grouping genetic algorithm.

(ID 41, in page 49) **Miriam Pescador-Rojas, Erika Hernández-Rubio and Amilcar Meneses-Viveros**. Multi-objective evolutionary scheduling in heterogeneous parallel

environments.

Session VIII: 10:30 – 11:50, Room 2, Optimization in Industry 2
Chair: Leonardo Trujillo

(ID 39, in page 50) **Raymundo Juarez, Carmen Borrego, Gonzalo Maldonado and Amalio Ponce**. Crop yield and crop dates forecasting based on biometrical and weather variables: An artificial neural network approach.

(ID 24, in page 55) **Arsenio Lizardi Durán, Guadalupe Castilla Valdez, Juan Frausto Solis and Juan Javier González Barbosa**. A hybrid method for solving Dynamic Job Shop Scheduling Problem.

(ID 26, in page 57) **Oliver Cuate and Oliver Schütze**. The Pareto tracer for degenerated multi-objective optimization problems.

Session IX: 12:00 – 13:20, Room 1, Health
Chair: Esteban Tlelo-Cuautle

(ID 25, in page 58) **Hans Yuan, Timothy Lam and Mauricio de Oliveira**. EMS-Simulator: an Open-Source Discrete-Event Simulation Framework for Medical Emergencies.

(ID 22, in page 59) **Eduardo Enrique Contreras-Luján, Enrique Efrén García-Guerrero, Oscar Roberto López-Bonilla, Esteban Tlelo-Cuautle, Didier López-Mancilla and Everardo Inzunza-González**. Evaluation of machine-learning algorithms for early diagnosis of deep venous thrombosis.

(ID 23, in page 61) **José Jaime Esqueda-Elizondo, J. Reyes Juárez-Ramírez, Oscar Roberto López-Bonilla, Enrique Efrén García-Guerrero, Gilberto Manuel Galindo-Aldana, Laura Jiménez-Beristáin, Alejandra Serrano-Trujillo, Esteban Tlelo-Cuautle and Everardo Inzunza-González**. Attention measurement of an ASD user using EEG signals: A case study.

Session X: 12:00 – 13:20, Room 2, Discrete Optimization 2
Chair: Guadalupe Carmona

(ID 44, in page 63) **Edgar Galvan, Gavin Simpson, Haseena Sreedharan, Marcela Quiroz-Castellanos and Leonardo Trujillo**. Evolving the Upper Confidence Bounds for Trees in Monte Carlo Tree Search.

(ID 29, in page 64) **Stephanie Amador Larrea, Marcela Quiroz Castellanos and Guillermo de Jesús Hoyos Rivera**. An Experimental Study of Grouping Crossover Operators for the Bin Packing Problem.

(ID 36, in page 65) **Jessica Elena Gonzalez San Martin, Laura Cruz-Reyes, Bernabe Dorronsoro, Marcela Quiroz-Castellanos, Héctor Fraire, Nelson Rangel-Valdez and Claudia Gómez-Santillán**. The dynamic bin packing optimization problem: open challenges.

Session XI: 16:30 – 17:50, Room 1, Circuits and ML 3
Chair: Gerardo de la Fraga

(ID 14, in page 67) **Irving Rodriguez Hernández, Juan Frausto Solis, Javier Gonzalez Barbosa, Luis Fortino Cisneros Sinencio and José Luis Purata Aldaz**. GA-DRL: Portfolio selection and deep reinforcement learning agent for daily stock trading.

(ID 32, in page 69) **Natan Vilchis and Adriana Lara**. A multi-objective approach

to deal with emotions in computer-aided musical composition.

Session XII: 16:30 – 17:50, Room 2, Discrete Optimization 3

Chair: Marcela Quiroz

(ID 33, in page 70) **José-Fernando Camacho-Vallejo, Sahori Polino and Juan G. Villegas**. An algorithm based in the Path Relinking metaheuristic for solving a location bilevel problem; an application to a social responsability project.

(ID 37, in page 71) **Citlalli Morales de la Cruz, Laura Cruz-Reyes, Rafael Alejandro Espín-Andrade, José Fernando Padrón-Tristán and Claudia Guadalupe Gómez-Santillán**. Genetic Knowledge Discovery Algorithm Based on Archimedean-Compensatory Fuzzy Logic.

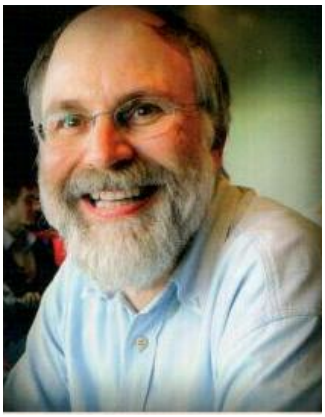
Invited Speakers

Wolfgang Banzhaf

Evolution Sometimes Needs Guidance

Michigan State University, USA

Abstract



Evolutionary computation prides itself with taking inspiration from natural evolution. But often a selective modelling approach has been used in the past that has shut out important natural phenomena from being considered essential to the process. We argue to follow the example of nature more closely, an approach we term 'computational evolution'. Here we discuss one of those issues in more detail: Variations are not totally random in nature. Instead, the process of variation follows pre-described paths. We discuss some of this evidence and its usefulness in applications by creating EC approaches that resemble more closely its natural example.

Website: `https://www.cse.msu.edu/~banzhafw`

Short biography

Wolfgang Banzhaf obtained his doctorate in physics from the Technical University of Karlsruhe (now Karlsruhe Institute of Technology). From 1993 he has taught computer science and engineering in his native Germany and in Canada. Since 2016, he has been a professor of computer science at Michigan State University in East Lansing, where he holds the John R. Koza Chair in Genetic Programming. He is the author of 3 books and more than 200 articles in bio-inspired computing, complex adaptive systems, neural networks, artificial life and evolutionary computing. His best-known works are Genetic Programming (1998), Linear Genetic Programming (2007), and Artificial Chemistries (2015), written together with his students and collaborators.

Gustavo Olague

Brain Programming Vs. Neuroevolution Visiting the Old Fight Between Symbolic and Sub-symbolic Learning

Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Mexico

Abstract

The old fight between Symbolic and Sub-symbolic approaches to artificial intelligence (AI) is not over yet. This friendship/rivalry between researchers of both methods sometimes has turned into a love/hate dilemma. Indeed, the quest for AI captured the imagination of researchers with interest in copying nature to produce intelligent agents. In this talk, we review the history of artificial intelligence regarding symbolic and sub-symbolic paradigms. This setup helps us understand how/why symbolic reasoning was once the masterpiece of AI; it now seems to be void of meaning compared to deep learning. In contrast, neuroevolution attempts to enter the deep-learning revolution still have to prove their value. In this talk, we briefly survey the main points to differentiate the pros and cons of both strategies as research opportunities in this new era of artificial intelligence. Brain programming opens new opportunities to look for solutions to the plethora of challenges that current sub-symbolic research solves in multiple fields. We conclude with a list of research opportunities and reflect on the main task in artificial intelligence.

Short biography

Gustavo Olague received the B.S. and M.S. degrees in industrial and electronics engineering from the Instituto Tecnológico de Chihuahua (ITCH), in 1992 and 1995, respectively, and the Ph.D. degree in computer vision, graphics, and robotics from the Institut Polytechnique de Grenoble (INPG) and the Institut National de Recherche en Informatique et Automatique (INRIA) in France. He is currently a Professor with the Department of Computer Science, Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), México, and also the Director of the EvoVisión Research Team. He is also an Adjunct Professor of engineering with the Universidad Autonóma de Chihuahua (UACH). He has authored over 100 conference proceedings papers and journal articles, co-edited special issues in Pattern Recognition Letters, Evolutionary Computation (MIT Press), and Applied Optics (OSA). He has authored the book Evolutionary Computer Vision (Springer) in the Natural Computing Series. His main research interests are evolutionary computing and computer vision. He is a member of the Editorial Team of the IEEE Access, Neural Computing and

Applications (Springer), and served as the Co-Chair of the Real-World Applications track at the main international evolutionary computing conference, GECCO (ACM SIGEVO Genetic and Evolutionary Computation Conference), in 2012 and 2013.

He has received numerous distinctions, among them the Talbert Abrams Award–first honorable mention 2003–presented by the American Society for Photogrammetry and Remote Sensing (ASPRS) for authorship and recording of current and historical engineering and scientific developments in photogrammetry; Best Paper Awards at major conferences such as GECCO, EvoIASP (European Workshop on Evolutionary Computation in Image Analysis, Signal Processing, and Pattern Recognition), and EvoHOT (European Workshop on Evolutionary Hardware Optimization); and twice the Bronze Medal at the Humies (GECCO award for Human-Competitive results produced by genetic and evolutionary computation).

Kenneth Sörensen

A history of Metaheuristics

University of Antwerp, Belgium

Abstract

The field of (meta)heuristics is still young, which is surprising given the fact that people have been using heuristics (and even metaheuristics) to solve optimization problems since the dawn of humanity. In this talk, we investigate the history of the field of metaheuristics, starting a long time in the past, and passing through several paradigm changes. We end with an outlook at the future of the field of (meta)heuristics.

Short biography

Kenneth Sörensen is full professor at the Faculty of Business and Economics of the University of Antwerp, where he chairs the ANT/OR - University of Antwerp Operations Research Group. His research focuses on innovative applications of operations research and on the study of heuristic optimization algorithms.

Carlos Segura

Design of Population-based Optimizers: Diversity, Data Structures and Algorithms

Centro de Investigaciones en Matemáticas (CIMAT), Guanajuato, México

Abstract

In this talk I will describe some key topics in the proper design of population-based optimizers, including issues that are problem-dependent, as well as problem-independent features. The importance of mastering Data Structures and Algorithms will be shown by describing some of the optimizers that have been developed in recent years at CIMAT and that have won some international competitions, such as the extended round of Google Hash Code 2020. Several algorithms that hold currently the best-known solutions for various optimization problems are discussed.

Short biography

Carlos Segura was born in Santa Cruz de Tenerife, Spain. He received the M.S. degree in computer science from the Universidad de La Laguna, in 2009 and the Ph.D. degree in computer science from the Universidad de La Laguna, in 2012. He has authored and co-authored over 75 technical papers and book chapters, including more than 25 journal papers. His publications currently report over 900 citations in Google Scholar. He serves in the editorial board of several international conferences and was the chair of the Genetic Algorithms Track in the Genetic and Evolutionary Computation Conference 2020 and 2021. His main research interests are: design of evolutionary algorithms, diversity management and problem solving paradigms. He has been part of the winning team of several optimization-oriented competitions, such as the extended version of the Google Hash Code 2020 and the Parameterized Algorithms and Computational Experiments Challenge 2018. He is currently an Associate Researcher of the Computer Science area at the Center for Research in Mathematics (CIMAT).

Women at NEO

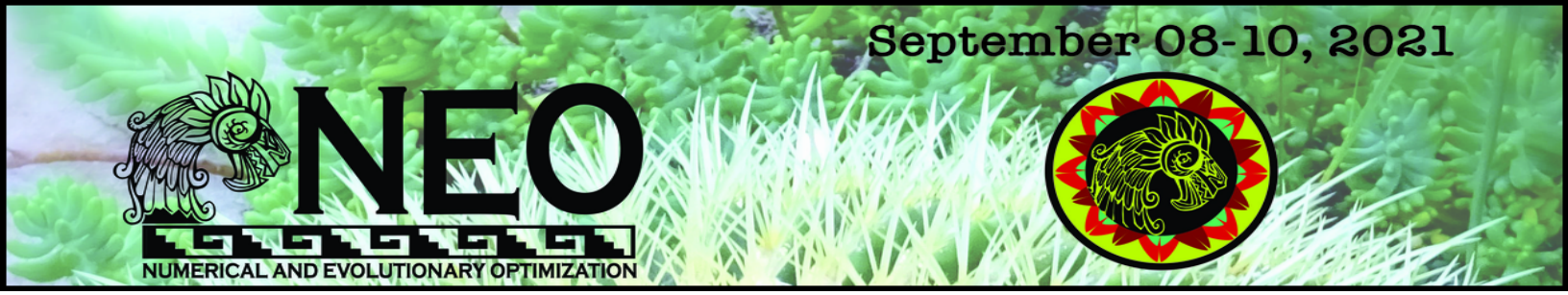
Aim of the session

Women at NEO is a space to promote and intensify research, discussion, and collaboration among the female NEO community. We take time at NEO also for networking and meeting consolidated and young researchers, as well as involved students. One of the goals of W-NEO is to inspire, engage and advice students who are currently working—or planning to work—on optimization subjects lead by the female professors. This section consists of some talks and a short meeting. We will encourage the setting up of specific woman networks with common interests.

It is worth to notice that male researchers/students are also welcome in every part of this session!

Adriana Lara

Marcela Quiroz



NEO 2021 Contributed Talks

Here are all the abstracts of NEO 2021 Workshop. Please, check the schedule (in page 9) in this same notebook to have a list of all the abstracts.

A Multi-Objective Optimization Method to Estimate Parameters of the COVID SEIMR/R-S Model with a Case Study of Mexico

Alejandro Castellanos-Alvarez [a], **Laura Cruz-Reyes** [a], **Héctor Fraire-Huacuja** [a], **José Alfredo Brambila-Hernández** [a], **Nelson Rangel-Valdez** [b], **Claudia Gómez-Santillán** [a] and **Jesús M. Velásquez-Bermúdez** [c]

[a] Graduate Program Division
Tecnológico Nacional de México, Instituto Tecnológico de Ciudad Madero
Cd. Madero 89440, Mexico
alex810_castellanos@hotmail.com, lauracruzreyes@itcm.edu.mx, automatas2002@yahoo.com.mx
claudia.gomez@itcm.edu.mx and alfredo.brambila@outlook.com

[b] CONACyT Research Fellow at Graduate Program Division
Tecnológico Nacional de México
Instituto Tecnológico de Ciudad Madero, Cd. Madero 89440, Mexico
nelson.rangel@itcm.edu.mx

[c] Research Center for Advanced Decision Technologies
FI
USA
jesus.velasquez@decisionware.net

The spread of infectious diseases is of great concern because of their impact on public health, so it is necessary to design strategies to manage these diseases, and it is useful to support the analysis with simulation models of these diseases. Epidemiological models rely on parameters to simulate the behavior of a pandemic. For a pandemic such as COVID-19 these parameters are unknown. For this reason, parameter estimation is an essential area of research to study pandemic behavior and its effects. In this context it is required to develop models and methods to estimate the values of the model parameters.

The SEIMR/R-S model, proposed by (Velásquez-Bermúdez et al, [1]), considers the interregional impacts in a macro-region, and also takes into account the impact of modeling the population divided into socio-demographic segments according to age and economic stratum. Unfortunately, most of the parameters are unknown for a given pandemic, for this reason, parameter estimation is essential to simulate pandemic behavior in such a way that the simulated COVID-19 pandemic data fits the daily historical record data (e.g. historical dead, recovered, exposed, etc.). To our knowledge, this is the first work that addresses parameter estimation for this model.

On the other hand, making manual parameter adjustment to bring the simulation values closer to the actual values is not an easy task, so a method that automatically adjusts parameters from an actual record of a n day period would allow control measures to be implemented to mitigate the effects of the pandemic.

An evolutionary genetic algorithm was used to perform the parameter fitting, in which the chromosome of each individual consisted of a representation of each parameter in the domain of real numbers. We sought to minimize the objective function, which consists of applying the Mean Absolute Percentage Error that existed between the historical series and the series generated from the simulation of the model with the parameters found. The model developed for the adjustment of parameters was applied to the historical dead and recovered data of the Tamaulipas, Mexico conurbation zone: Altamira, Cd. Madero and Tampico.

The experimental results obtained show that the automatic adjustment of the parameters of the epidemiological model achieves a good approximation of the historical series of interest, in such a way that it can be a useful tool to predict the behavior of the spread of COVID-19 can be under control in all the communities considered, if the appropriate restrictions and policies are implemented to control the rates of spread of the virus.

References

[1] Velasquez-Bermudez, Jesus, (in press), SEIMR/R-S. General Epidemic Simulation Model Multi-Infected States - Multi Sociodemographic Segments - Multi-Region Mobility. Theory, European Scientific Journal Special Edition PUBLIC POLICIES IN TIMES OF PANDEMICS

Explainable Artificial Intelligent Model to Determine the Severity on COVID-19 Patients

Miriam Pescador-Rojas[a], **Víctor Andrés Sosa-Hernández**[b,c]
José L. Maravillas-Montero[d]

[a]Escuela Superior de Cómputo
Instituto Politécnico Nacional
Ciudad de México, México
mpescadorr@ipn.mx

[b]Departamento de Computación
Instituto Tecnológico de Monterrey
Estado de México, México
vsosa@tec.mx

[d]Departamento de Biomedicina Molecular
Centro de Investigación y de Estudios Avanzados
Ciudad de México, México
victor.sosah@cinvestav.mx

[c]Red de Apoyo a la Investigación
Universidad Nacional Autónoma de México
e Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán
Ciudad de México, México
maravillas@cic.unam.mx

Data mining is a powerful tool to study the clinical trial databases of COVID-19 patients to discover new knowledge and understand patterns in their medical conditions. Medicine is one field that has successfully integrated machine learning techniques to produce intelligent models that could help improve medical assistance, attention, and diagnosis. However, the machine learning output model should express its decisions in a language that doctors can interpret easily. Among several machine learning techniques, decision trees produce explainable models in a language closer to the doctors. We choose this technique to produce intelligent models to analyze levels of frequency of cells in COVID-19 patient datasets. COVID-19 patients display alterations in different myeloid and lymphoid cells that are associated with several clinical features; among these leucocyte subsets, B cells such as DN subpopulations. Therefore, different levels of cells are associated with the disease severity (mild/moderate, severe, or critical) and the outcome of patients (deceased or surviving). Our preliminary results provide different guidelines such as to understand the immunology response to produce antibodies that combat the disease, and to predict if one patient with moderate or severe severity can change its condition to critical severity.

Theoretical Analysis of the application of a University Timetable Model in the Optimization of Hospital Resources

Angel I. García F. 1[a] ,Nelson R. Valdez 2[a], Claudia G. Gómez S. 3[a]
Laura Cruz R. 4[a] Héctor J. Morales R. 5[a], María L. Morales R 6[a]

[a]Computer Science Department
Tecnológico Nacional de México/Instituto Tecnológico de Cd. Madero
Juventino Rosas S/N, Col. Los Mangos, Cd. Madero, Tamaulipas, C.P. 89440, México.
emails:G20073003@cdmadero.tecnm.mx, nelson.rv@cdmadero.tecnm.mx, claudia.gs@cdmadero.tecnm.mx,
hector.fraire.huacuja@itcm.edu.mx, lauracruzreyes@itcm.edu.mx, lucila.mr@cdmadero.tecnm.mx

Contagious diseases have been a matter of concern throughout the history of mankind. Epidemics such as the plague of Athens, the Black Death, or the Great Plague of London, to name a few extreme cases where they led to the death of up to 40% or 50% of the population of the region, and every year there are different epidemics of viruses worldwide, but only a few reach the level of public concern, Given the recent contingency caused by Covid-19 (SARS-CoV-2) and its exponential spread throughout the world, it has been decided to use various control policies, such policies are included in mathematical models for the prediction of the epidemic, hospital resource management, impact on the economy, among others.

Mathematical modeling allows establishing relationships between phenomena, making approximations, thus providing us with valuable tools to combat epidemics, providing guidelines on how to combat them, and even logistical and economic issues.

The management of hospital resources is a matter of considerable public and political concern, especially in times of contingency, such as the current pandemic of COVID-19, where tools and methods are required to help in decision-making. The vital physical resources for hospitals are fundamental and a necessary part to minimize the impacts due to the pandemic; with this, it should be made clear that for the management of an epidemic, resources must be managed and available, in addition to taking into account the control policies selected by the public health authorities, the coordination of these two aspects is fundamental for the minimization of deaths due to an epidemic.

In this work, the mathematical model for the optimization of hospital resources was formulated with the objective of managing the vital physical resources necessary to minimize the impacts due to the pandemic. This model is based on the review of resource allocation works to make the selection of hospital resources to be allocated. That is, the design of a hospital resource model is proposed based on a model of university schedules of the CB-CTT (Curriculum-Based Course Timetabling) problem with the analysis and analogy to the hospital resource problem defining each of the elements that would be involved, proposing a mathematical programming model based on the university scheduling problem[1].

References

[1] Rangel et al., Practical relaxation of a special case of the Curriculum-Based Course Timetabling problem, Progress in Artificial Intelligence 2014, `https://link.springer.com/article/10.1007/s13748-014-0055-4`

Seeking an Affordable Solution to Enhance Mexico City's Air Quality by Optimized Tree Seeding

Fernando Moreno, Dra. Adriana Lara

Escuela Superior de Física y Matemáticas del
Instituto Politécnico Nacional
Av. Instituto Politécnico Nacional s/n,
Edificio 9 Unidad Profesional "Adolfo López Mateos"
fmorenog0901@alumno.ipn.mx, alara@ipn.mx

High levels of air pollutants represent a problem in many cities all over the world [1]. Reforesting urban areas is a straightforward but challenging solution since different factors like pollen release and optimal location should be considered [3]. In this work, we tackle improving the air quality in Mexico City by urban reforestation. Our analysis includes the first intervention and also the maintenance costs for three years to come. The goal was to provide an affordable solution to the problem, assuming a limited number of available seeding trees. We built a prediction model for the solution using open-access data about air quality [2] and also considering some meteorological aspects [4] particular to this city. We apply a Multi-objective Evolutionary Algorithm to find the numerical solution and determine the number and distribution of the trees all over the counties. Based on publicly available data, this work shows the potential of these numerical tools to suggest solutions to improve the citizen's quality of life.

References

[1] Bingheng Chen and Haidong Kan. Air pollution and population health: a global challenge. *Environmental health and preventive medicine*, 13(2):94–101, 2008.

[2] Dirección de Monitoreo Atmosférico. Base de datos de la red rama y retmet de la ciudad de méxico, Jan 2021.

[3] Neratzia Julia Georgi and K Zafiriadis. The impact of park trees on microclimate in urban areas. *Urban Ecosystems*, 9(3):195–209, 2006.

[4] Roberto Rojano, Jhonny Perez, and Edesnel Freyle. Effect of relative humidity in determining pm10 using a dataram 4 in coastal region of colombia. *Revista Técnica Ingeniería Universidad del Zulia*, 35(2):1–9, 2012.

Simple regression models for the rapid approximation of the hypervolume indicator for multiobjective optimization problems

Cristian Sandoval[a], Oliver Cuate[b], Luis C. González[c], Leonardo Trujillo[a,*] Oliver Schütze[b]

[a]Tecnológico Nacional de México/IT de Tijuana, BC, México
sr_cristian@outlook.es, leonardo.trujillo@tectijuana.edu.mx

[b]Computer Science Department, Cinvestav-IPN, México City, México
ocuate@computacion.cs.cinvestav.mx, schuetze@cs.cinvestav.mx

[c]Autonomous University of Chihuahua, School of Engineering, Chihuahua city, Chih., México
lcgonzalez@uach.mx

Hypervolume (HV, [1, 2]) has become one of the most popular indicators to assess the quality of Pareto fronts. However, the standard way of computing it has a computational complexity of $O(N^{k+1})$ for N solutions and k objectives. In this study, we propose a regression-based approach to learn new mathematical expressions to approximate the HV value and improve at the same time their computational efficiency. In particular, Genetic Programming and Stepwise Linear Regression are used for modelling techniques. The former is used because it can produce compact and efficient symbolic models, while the latter is used to model the HV in the most difficult case considering a large set of predictor variables and interaction terms. Using the HV indicator, we analyze the performance of some state-of-the-art multi-objective evolutionary algorithms (MOEAs). MOEAs are tested in instances of two well-known reference sets (DTLZ and WFG). The solution sets are subjected to a feature extraction process to facilitate the training process of simple models that approximate the value of HV. To evaluate this approach, we exhaustively measure the deviation of the new models against HV ground truth in the DTLZ and WFG benchmark datasets. We also test the new models by using them as a guiding mechanism within an indicator-based algorithm, namely SMS-EMOA. Results are very consistent and promising since the new models report accuracy levels of 94% and above for problems of 3, 4 and 5 objectives. What is more striking is the execution time achieved by these models, which in a direct comparison against standard HV calculation achieved a speedup of over 1000x, all of this based on their relative simplicity of features and model composition.

References

[1] E. Zitzler, L. Thiele, Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach, IEEE transactions on Evolutionary Computation 3 (4) (1999) 257-271.

[2] E. Zitzler, D. Brockhoff, L. Thiele, The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration, in: International Conference on Evolutionary Multi-Criterion Optimization, Springer, 2007, pp. 862-876.

Archiving Strategies for Multi-objective Evolutionary Algorithms

Oliver Schütze[a], **Carlos Hernández**[b]

[a]Computer Science Department
Cinvestav-IPN
Av. IPN 2508, 07260 Mexico City

[b] Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
Universidad Nacional Autónoma de México
Circuto Escolar 3000, C.U., Coyoacán, 04510 Ciudad de México, CDMX

In this presentation we review archiving strategies developed over the last years by the authors that deal with suitable approximations of the sets of optimal and nearly optimal solutions of multi-objective optimization problems by means of stochastic search algorithms. All presented archivers are analyzed with respect to the approximation qualities of the limit archives that they generate and the upper bounds of the archive sizes. The convergence analysis will be done using a very broad framework that involves all existing stochastic search algorithms and that will only use minimal assumptions on the process to generate new candidate solutions. All of the presented archivers can effortlessly be coupled with any set-based multi-objective search algorithm such as multi-objective evolutionary algorithms, and the resulting hybrid method takes over the convergence properties of the chosen archiver.

Chaotic Oscillators Implemented in FP16 Numbers

Luis Gerardo de la Fraga[a], Esteban Tlelo-Cuautle 2[b], Martín Alejandro Valencia-Ponce

[a]Computer Science Department, Cinvestav
Av. IPN 2508, Mexico City, México
fraga@cs.cinvestav.mx
[b]Department of Electronics, INAOE
Tonantzintla Puebla, 72840 México
fraga@cs.cinvestav.mx

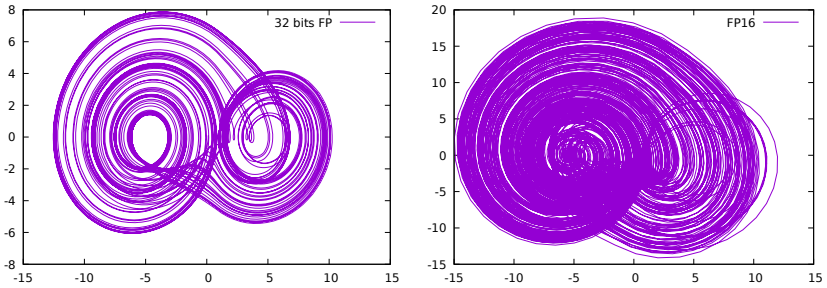
The Lü and the Rossler chaotic oscillators [1] are implemented in FP16 numbers. Both oscillators are integrated using Euler integration method which is the easiest to implement. FP16 numbers [2] are half precision floating point numbers with $|x| \in [6.10 \times 10^{-5}, 6.55 \times 10^{4}]$, which allows a fast calculation time, but because of this reduced range the direct implementation of Euler method with small integration time produce degraded oscillators which enter in periodic instead of chaotic behavior. Differential evolution is used to optimize the oscillators increasing the integration time and oscillator's parameters to improve their Kaplan York dimension.
The Lu oscillator integrated with the Euler method is defined as:

$$\begin{aligned} x_{n+1} &= x_i + hy, \\ ,y_{n+1} &= y_i + hz, \\ z_{n+1} &= z_i + h(-a - b - c + d_1 f(x), \end{aligned}$$

where $f(x) = -5$, if $x < 1$; $f(x) = 5x$, if $|x| \leq 1$; and $f(x) = 5$, if $x > 5$. The values for the coefficients before and after the optimization are shown in the Tab. 1. The used initial state are $[0.1, 0.1, 0.1]$. Ten thousand samples taken at $10h$ and h for the optimized oscillators are shown in the graphs. The optimized h reduce in 10 the calculations and the optimized coefficients improve the oscillator behavior.

Table 1. Coef. values

	Lü	Optimized
a	0.7	2.91
b	0.7	2.3913
c	0.7	1.0533
d_1	0.7	3.1351
h	0.01	0.13



References

[1] E. Tlelo-Cuautle, L.G. de la Fraga, and J.J. Rangel-Magdaleno, Engineering Aplications of FPGAs, Springer, 2016

[2] L.G. de la Fraga, Differential Evolution under Fixed Point Arithmetic and FP16 Numbers, Mathematical and Computational Applications, 26(1), 2021, doi: 10.3390/mca26010013

Subsample algorithm for Support Vector Regression

Jhoan Delgado[a], Anibal Sosa[b] , M.D González Lima[c]

[a]Universidad Icesi
Calle 18 #122-135, Cali, Colombia
jhoan.delgado@u.icesi.edu.co

[b]Depto. Tecnologías de Información y Comunicaciones
Universidad Icesi
Calle 18 #122-135, Cali, Colombia
uasosa@icesi.edu.co

[c]Depto. Matemáticas
Universidad del Norte
Km 5 Via a Pto Colombia, Barranquilla, Colombia
limad@uninorte.edu.co

Support Vector Machines [1] is a well known supervised learning algorithm, for both classification and regression, that has shown impressive results when it comes to classifying or forecasting linear and non-linear data. Despite its impressive results, this algorithm is not suitable for large datasets because its training process may become a computationally expensive task, $\mathcal{O}(n^3)$ time and $\mathcal{O}(n^2)$ memory training complexities, where n is the cardinality of the data set. In this work, we aim to extract a significantly smaller training set for Support Vector Regression (SVR), from which the Support Vectors of the original data set are likely to be determined. To this end we propose an extended and novel implementation of a subsample algorithm inspired by [2], with Bayesian hyperparameter optimization. We study the presence of support vectors from the original data set in each subsample and compare different metrics for performance of the method against a standard SVR implementation [3]. We present results for a variety of benchmark problems with sizes up to $200K$ variables and 13 attributes, where we obtain up to 67.5 times faster convergence than the SVR algorithm with parameter optimization.

References

[1] Vapnik, V., Golowich, S. E., & Smola, A. (1997). Support vector method for function approximation, regression estimation, and signal processing. Advances in neural information processing systems, 281-287.

[2] Camelo, S. A., González-Lima, M. D., & Quiroz, A. J. (2015). Nearest neighbors methods for support vector machines, Annals of Operations Research, 235(1), 85-101.

[3] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. Journal of machine Learning research, 12, 2825-2830.

Differential Evolution with Mixed Precision Numbers

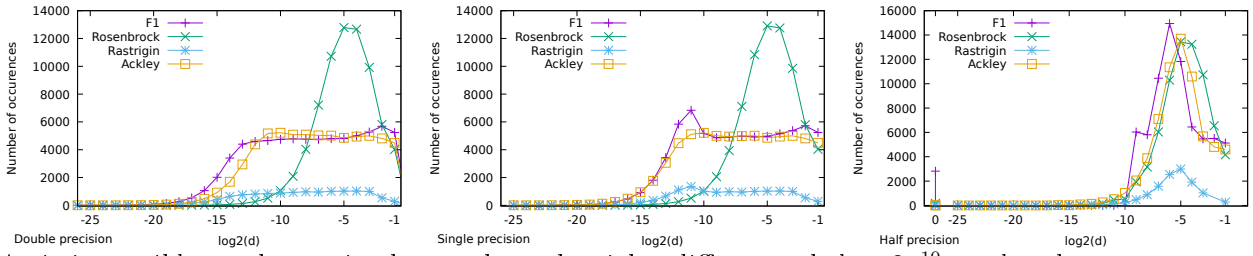
Luis Gerardo de la Fraga

Computer Science Department, Cinvestav
Av. IPN 2508, Mexico City, México
fraga@cs.cinvestav.mx

It is studied the feasibility to use mixed precision numbers in the differential evolution (DE) algorithm. The core of DE is

$$x'_j = \begin{cases} x_{r3,j} + F(x_{r1,j} - x_{r2,j}) & \text{if } U(0,1) < R \text{ or } j = j_{\text{rand}} \\ x_{i,j} & \text{otherwise,} \end{cases}$$

when the new individual x' is generated from three parents x_{r_1}, x_{r_2}, and x_{r_3}, where r_1, r_2, and r_3 are three random numbers within the population size. The expression of the difference $x_{r_1,j} - x_{r_2,j} = d$ for each variable j is crucial in the calculation. Four benchmark functions, one in [1] and three in [2] are used to generate the histograms of those differences shown in the figure below. The average of 10 executions with a different seed for the random number generator are calculated.



As it is possible to observe in the graph at the right, differences below 2^{-10} are lost because cannot be represented with half precision (or fp16) numbers. The option is to maintain the variables in single precision numbers and realize the multiplication with scaled half numbers. Statistics of objective function values and execution time with 30 execution of the used benchmark are the following:

	Double precision		Single precision		Half precision		Mixed precision	
	mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.
F1	-3.224518	< 1e-6	-3.224518	< 1e-6	-3.224609	< 1e-6	-3.224518	< 1e-6
Ros.	0.390024	1.91e-05	0.390024	1.58e-05	0.390389	1.36e-4	0.390031	1.7e-5
Ras.	-33.000000	< 1e-6	32.999823	3.19e-3	-33.000000	< 1e-6	-32.980100	4e-2
Ack.	-6.999990	3.66e-6	-6.999988	1.01e-5	-6.999609	1.19e-3	-6.999985	7.3e-6

Execution times in a Raspberry Pi B which have an ARM processor with half numbers support.

Double	2 min 12.651 sec
Single	2 min 5.893 sec
Half	2 min 0.612 sec
Mixed	1 min 57.262 sec

References

[1] E. Tlelo-Cuautle, L.G. de La Fraga, O. Guillén-Fernández, and A. Silva-Juárez, Optimization of Integer/Fractional Order Chaotic Systems by Metaheuristics and their Electronic Realization, CRC Press, 2021, doi: 10.1201/9781003042181.

[2] L.G. de la Fraga, Differential Evolution under Fixed Point Arithmetic and FP16 Numbers, Mathematical and Computational Applications, 26(1), 2021, doi: 10.3390/mca26010013

Traffic flow prediction for smart traffic lights

Alfonso Navarro-Espinoza[a], Oscar Roberto López-Bonilla[a] , Enrique Efrén García-Guerrero[a], Esteban Tlelo-Cuautle[b], Didier López-Mancilla[c], Carlos Hernández-Mejía[d] and Everardo Inzunza-González[a]

[a]Faculty of Engineering, Architecture and Design
Autonomous University of Baja California
3917 Ensenada-Tijuana highway
Ensenada, Baja California, México, ZIP 22860.
alfonso.navarro@uabc.edu.mx, olopez@uabc.edu.mx, eegarcia@uabc.edu.mx
einzunza@uabc.edu.mx

[b]Department of Electronics
Instituto Nacional de Astrofísica, Óptica y Electrónica
Luis Enrique Erro #1, Santa María Tonanzintla
San Andrés Cholula, Puebla, México, ZIP 72840.
etlelo@inaoep.mx

[c]Centro Universitario de Los Lagos
Universidad de Guadalajara
1144 Enrique Díaz de León, Col. Paseos de la Montaña
Guadalajara, Jalisco, México
didier.lopez@academicos.udg.mx

[d]Doctorado en Ciencias de la Ingeniería
Instituto Tecnológico Superior de Misantla
Km 1.8 carretera Lomas de Cojolite
Misantla, Veracruz, México. ZIP 93821
cmahernandez@gmail.com

Abstract. Nowadays, due to the accelerated growth of the population and, as a consequence, the number of cars in the cities added to the technological limitations of traffic control signs have made vehicular traffic one of the problems of modern life. Many cities worldwide have traffic congestion problems at certain peak hours, producing more pollution, noise, and stress on citizens.
Intelligent Transportation System (ITS) is a significant part of the smart city system. ITS can provide real-time road infrastructure analysis and more efficient road traffic control using big data information and communication technology. With the help of ITS, a safe, mobile, environmentally sustainable, and comfortable road environment can be provided [1]. To achieve ITS, traffic flow prediction is essential because it can provide continuous and precise road status information based on past road conditions. That information can be helpful for the applications involved in ITS, such as traffic congestion control and traffic light control [2]. For example, it can calculate the possibility of congestion on the corresponding road segment and deal with the possible situation in advance [3].
In this paper, we propose using of recurrent neural networks such as gated recurrent units (GRU) and long short-term memory (LSTM) for the traffic flow prediction at an intersection to program the optimal times of the states of the traffic lights. In this manner, it is expected to use it in a smart traffic light controller operating in real-time with sensors to detect the number of vehicles.
A public database [4] was used for the training and testing of the neural network. The database contains

the number of vehicles sampled every five minutes at six intersections for 56 days using different sensors. For this research, four of the six intersections are used to simulate the four lanes of an intersection. The moving average replaces the zeros in the database since there are sections with many of them, probably due to sensor failure. However, applying a mean of all the data causes spikes where they occurred. 75% of the data (35 days) was used for training, and 25% (21 days) for validation and testing. For the preprocessing and design of the neural network, those proposed in [5] are used.
Multiple metrics are used to evaluate the proposed models; both LSTM and GRU achieved an explained variance score of approximately 0.93, a mean absolute error (MAE) near to 10, a root mean squared error (RMSE) near 15, and an approximate R-squared (R^2) of 0.93. GRU requires fewer training parameters and thus uses less memory, executes faster, and trains faster than LSTMs.

Keywords: Deep-Learning; Intelligent Transportation System; Smart City; Recurrent Neural Network; Traffic Flow Prediction;

References

[1] A. Boukerche, J. Wang, A performance modeling and analysis of a novel vehicular traffic flow prediction system using a hybrid machine learning-based model, Ad Hoc Networks Volume 106, 2020.

[2] A. Boukerche, J. Wang, Machine Learning-based traffic prediction models for Intelligent Transportation Systems, Computer Networks, Volume 181, 2020.

[3] A. Boukerche, Y. Tao, P. Sun Artificial intelligence-based vehicular traffic flow prediction methods for supporting intelligent transportation systems, Computer Networks, Volume 182, 2020.

[4] Cristian Axenie & Stefano Bortoli., Road traffic prediction dataset. (Versión v1.0.0) [Data set], Zenodo, doi: `http://doi.org/10.5281/zenodo.3653880`

[5] R Fu, Z Zhang, L Li., Using LSTM and GRU neural network methods for traffic flow prediction, Chinese Association of Automation 2017, pp. 324-328.

Assembly components detection system based on convolutional neural networks: a case study

Edgar Rene Ramos-Acosta[a], Enrique Efrén García-Guerrero[a], Oscar Roberto López-Bonilla[a], Esteban Tlelo-Cuautle[b], Didier López-Mancilla[c], Carlos Hernández-Mejía[d] and Everardo Inzunza-González[a]

[a]Faculty of Engineering, Architecture and Design
Autonomous University of Baja California
3917 Ensenada-Tijuana highway, Ensenada, Baja California, México, ZIP 22860.
edgar.rene.ramos.acosta@uabc.edu.mx, eegarcia@uabc.edu.mx, olopez@uabc.edu.mx
einzunza@uabc.edu.mx

[b]Department of Electronics
Instituto Nacional de Astrofísica, Óptica y Electrónica
Luis Enrique Erro #1, Santa María Tonanzintla
San Andrés Cholula, Puebla, ZIP 72840.
etlelo@inaoep.mx

[c]Centro Universitario de Los Lagos
Universidad de Guadalajara
1144 Enrique Díaz de León, Col. Paseos de la Montaña
Guadalajara, Jalisco, Mexico.
didier.lopez@academicos.udg.mx

[d]Doctorado en Ciencias de la Ingeniería
Instituto Tecnológico Superior de Misantla
Km 1.8 carretera Lomas de Cojolite
Misantla, Veracruz, México. ZIP 93821
cmahernandez@gmail.com

Abstract. The use of artificial intelligence (AI) in the industry has increased in recent years due to the 4th industrial revolution, where AI has been the new strategy to solve the most complex problems related to big data, image recognition, object detection, and among other classification and prediction problems [1, 2, 3]. This article introduces the development and implementation of a zinc component detection system within a manufacturing process, by using convolutional neural networks (CNN). The aim of this paper is the planning, execution, and evaluation of different deep-learning (D-L) algorithms in order to detect five different zinc-based components, under different ambient light conditions and different finishings. The methodology starts by creating a custom dataset, continuing with transformation and image augmentation, designing a CNN architecture known as Convnet, Xception, MobileNetv2, with one output layer, four hidden layers and one last output layer with six neurons. For comparison purposes, other state of the art [4] models with the same dataset are trained by using the transfer learning technique. The training stage is performed with 800 images representing 80% of the dataset, while the validation stage was carried out with 200 images. Seven out of eight models achieved more than 95% accuracy in the testing phase with 200 images per class. The top three test accuracy algorithms were MobileNetV2, Xception and Convnet with 100%, 99%, and 98% of test accuracy respectively. Therefore, these algorithms can be adopted in the manufacturing production line, as good classifier models to be utilized for zinc-based component detection since the confusion matrix showed 100% test accuracy. Furthermore, a high-performance detection smart system is achieved.

Keywords: Deep-Learning; computer vision; Convnet; Xception; MobileNetv2; Industry 4.0.

References

[1] Zamora-Hernández, M.-A., et al., Deep learning-based visual control assistant for assembly in Industry 4.0. Computers in Industry, 2021. 131: p. 103485.

[2] Yun, J.P., et al., Automated defect inspection system for metal surfaces based on deep learning and data augmentation. Journal of Manufacturing Systems, 2020. 55: p. 317-324.

[3] Wei, D., et al., Deep learning assisted vision inspection of resistance spot welds. Journal of Manufacturing Processes, 2021. 62: p. 262-274. ArXiv 2019, `http://arxiv.org/abs/1905.12322`

[4] Sharma, V. and R.N. Mir, A comprehensive and systematic look up into deep learning based object detection techniques: A review. Computer Science Review, 2020. 38: p. 100301.

Detection of wind turbine blade damage caused by insect erosion using AutoML methods

Javier Carmona[a], J. Enríquez-Zárate[b], Armando Cárdenas[a]
Daniel Hernández[a], Leonardo Trujillo[a]

[a]Tecnológico Nacional de México/IT de Tijuana
Tijuana, Baja California, México
javier.carmona@itlp.edu.mx
[b]Tecnológico Nacional de México/IT del Valle de Etla
Santiago Suchilquitongo, Oaxaca, México
jenriquezza@gmail.com

Wind turbines and the increasing advances in their design, as well as the unpredictable environment in which they operate, present challenges in relation to blade leading edges erosion problems. Rain, dust, insects and other natural causes are primarily responsible for the problem. Also, erosion on the surface of wind turbine blades is one of the most critical problems in wind energy development and generation. This paper examines the erosion problem as well as the effect suffered/(experienced) by the blades due to the impact of insects on the leading edges. The Leading-edge Erosion Study (LEES) project database is used as reference [1]. This database was constructed by taking measurements in a low-speed wind tunnel located at Texas AM University operating at a maximum velocity of 90 m/s. These measurements were performed on a vertically mounted NACA63418 airfoil with a chord length of 813 mm spanning the entire height of the tunnel. The model is rotated so that at different angles of attack, different pressure coefficient measurements are obtained. Likewise, classification results are presented using Auto Machine Learning techniques [2] with which it is possible to determine the damage produced by the insects or if there is erosion. The AutoML methods that are analyzed are: AutoSklearn [3], TPOT [4] and H2O DAI [5].

References

[1] David, Maniaci, et al., Owner Reports-Airfoil Performance Degradation due to Roughness and Leading-edge Erosion, data and plots-Raw Data, 2020, Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

[2] Hutter, Frank and Kotthoff, et al., Automatic Machine Learning: Methods, Systems, Challenges, 2019, Springer.

[3] Feurer, Matthias and Klein, et al., Auto-sklearn: efficient and robust automated machine learning, 2019 Springer, Cham, pp. 113-134.

[4] Olson, Randal S and Moore, Jason H, TPOT: A tree-based pipeline optimization tool for automating machine learning, 2016, Workshop on automatic machine learning, pp. 66-74.

[5] Hall, Patrick and Gill, et al., Machine learning interpretability with h2o driverless ai, 2017, H2O. ai, pp. 66-74.

Computational Methods for the Design of Transconductance Operational Amplifiers

Gerardo de Jesús Martínez Neri, Luis Fortino Cisneros, Juan Frausto, Javier González Barbosa

Tecnológico Nacional de México/Instituto Tecnológico de Ciudad Madero

Abstract

Integrated circuit design is a challenging task due to a large number of variables, design parameters, and trade-offs between the electrical characteristics of the circuits. To design an amplifier, it is necessary to determine the size of every transistor in the circuit to maintain the appropriate polarization among the circuit during the operation of its life cycle. As modern applications demand high-speed, low voltage, reliable amplifiers, the number of elements in the circuits increases, thus increasing the complexity of the design process. Paradoxically, while the complexity of the design raises, so the demand for reduced time to market. To find the optimal amplifier design of a high Gain-Bandwith (GBW) product highly linear, operational transconductance amplifier (OTA´s) under a 0.5μm MOSIS CMOS technology. The design will be validated by running Monte Carlo simulations under PVT variations using the SPICE circuit simulator. In this context, the use of Deep Learning is proposed to find the optimal solution for the design of a high gain–bandwidth product (GBW), highly linear, operational transconductance amplifier (OTA´s) under a 0.5μm MOSIS CMOS technology. The design will be validated by running Monte Carlo simulations under PVT variations using the SPICE circuit simulator. There are very good alternative computational algorithms briefly discuss in this presentation, some of them are genetic algorithms, simulated annealing, and so forth.

Key Words: Computational OTA design, Operational Transconductance Amplifier, Heurists, SPICE

Introduction

In the literature, there are several Mathematical models describing circuit behavior for different technologies. Nevertheless, they usually do not consider side effects such as channel length modulation, body effect, and speed saturation, which can affect circuit performance. To take advantage of that issues and using SPICE simulations, can be applied to create a large population of possible subjects. Once this population of possible outcomes is created, computational algorithms as genetic algorithms, Deep Learning, can be applied for determine the best of these outcomes to establish the characteristics of the next batch of possible outcomes. Among the main algorithms are genetic algorithms, simulated annealing, Particle Swarm Optimization, Deep learning, and many others [1] [2]. This process is repeated until to obtain a design that fulfills the design requirements. Unfortunately, a clear description of how to use this methodology is not published.

CONCLUSION

In this work, we discuss the main computational methodologies as Genetic algorithms, Particle Swarm Optimization, and deep learning which are applied to the design Operational

Transconductance amplifiers. The design of integrated circuits consists of determining the physical dimensions of the electronic components that constitute a circuit. Because modern integrated circuits must be increasingly efficient, they result in circuits with a large number of transistors. This results in a large number of computational variables, making the design process difficult. Due to this complexity, the cost in the production and development of integrated circuits reaches levels that affect the profitability of companies. In the final version of the paper, we clearly present how to use the main Computational methods as for obtaining the most adequate OTA designs.

References

[1] K. Maji, R. Kar, D. Mandal, and S. Ghoshal. An evolutionary approach-based design automation of low power CMOS two-stage comparator and folded cascade OTA, AEU-International Journal of Electronics and Communications, vol. 70, no. 4, pp. 398–408, 2016.

[2] A. S. Khade, V. Vyas, and M. Sutaone. Performance enhancement of advanced recycling folded cascode operational transconductance amplifier using an unbalanced biased input stage, Journal of Integration the VLSI, May 2019.

Analysis of Transfer Learning in a Genetic Programming System Using DeepInsight

Joel Nation[a], Leonardo Trujillo[a], Luis Muñoz[a], Edgar Galván[b]

[a]Tecnológico Nacional de México/IT de Tijuana, Tijuana, Mexico
Joelleenation@gmail.com, Leonardo.trujillo@tectijuana.edu.mx, lmunoz@tectijuana.edu.mx

[b]Naturally Inspired Computation Research Group
National University of Ireland Maynooth, Lero, Ireland
edgar.galvan@mu.ie

Abstract

Transfer learning (TL) is a method where what is learned in one problem(source task) by a machine learning (ML) model can be used to simplify the training process in a second problem (target task). TL has become one of the most promising paradigms in ML, since it could allow researchers and practitioners to scale the learning process more efficiently and to maximize the usefulness of pre-trained models. While TL has achieved promising results in Deep Learning, it has been scarcely studied in genetic programming (GP) [1]. Moreover, in general predicting when, or why, TL might succeed, or fail to, is still an open question. This work presents a first attempt to determine when two problems might be compatible for TL to succeed between them. This is done by analyzing the feature space of each problem, as mapped by the DeepInsight method [2]. Previous results on TL for classification tasks with GP are used, such that every possible problem pair is categorized into one of two groups. The first group are TL compatible problems that includes all problem pairs where TL was highly successful, and the second are TL non-compatible problems where TL under-performed baseline methods. Results show that it is possible to distinguish between different groups of problem pairs by aligning their respective feature space representations and computing a distance between them.

References

[1] Muñoz, L., Trujillo, L. & Silva, S. Transfer learning in constructive induction with Genetic Programming, Genet Program Evolvable Mach 21, 529-569 (2020).

[2] Sharma, A., Vans, E., Shigemizu, D. et al. , DeepInsight: A methodology to transform a non-image data to an image for convolution neural network architecture, Sci Rep 9, 11399 (2019).

Implementation of a stack-based genetic programming algorithm in CUDA

Luis A Cárdenas, Daniel E Hernández, Leonardo Trujillo, Perla Juarez, Jose M Muñoz

Posgrado en Ciencias de la Ingeniería
Tecnológico Nacional de México/I.T. de Tijuana
Blvd Industrial S/N,
(luis.cardenas201, daniel.hernandezm, leonardo.trujillo, pjuarez, jose.munoz17)@tectijuana.edu.mx

Genetic programming (GP) is a methodology based on evolutionary algorithms to automatically develop computer programs that perform a user-defined task. It is a machine learning technique used to optimize a population of programs according to a fitness function that assesses the ability of each program to carry out the task in question. In this work, the implementation of a GP system that uses a linear representation and a stack-based interpreter was carried out, for its operation in graphics processing cards (GPU. The implementation was done entirely in CUDA, including specialized kernels for all stages of the evolutionary process, including interpretation and evaluation, selection, mutation, survival and replacement. In particular, a state-of-the-art mutation operator is used for this type of GP, called Uniform Mutation by Addition and Deletion (UMAD). Our system is tested on various symbolic regression problems, to show its performance in terms of quality (error) and efficiency (runtime) on state-of-the-art GPU cards.

References

[1] Muñoz L, Trujillo L, Silva S, Transfer learning in constructive induction with Genetic Programming, Genetic Programming and Evolvable Machine, 2020, vol 21, https://link.springer.com/article/10.1007/s10710-019-09368-y

[2] Muñoz J M, Trujillo L, Hernández D, Castelli M, Tapia J, GSGP-CUDA - a CUDA framework for Geometric Semantic Genetic Programming, Genetic Programming and Evolvable Machines, 2020, 529–569, doi: https://doi.org/10.1007/s10710-019-09368-y

[3] Helmuth T, Mcphee N, Spector L, Program synthesis using uniform mutation by addition and deletion, Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2018, 1127-1134, doi: https://doi.org/10.1145/3205455.3205603

[4] Spector L, Klein J, Keijzer M, The Push3 execution stack and the evolution of control, Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2005, 1689-1696, doi: https://doi.org/10.1145/1068009.1068292

[5] La Cava W, Moore J, Semantic variation operators for multidimensional genetic programming, Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2019, 1056–1064, doi: https://doi.org/10.1145/3321707.3321776

[6] Baeta F, Correia J, Martins T, Machado P, Speed Benchmarking of Genetic Programming Frameworks, Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2021, 768–775, doi: https://doi.org/110.1145/3449639.3459335

[7] Spector L, Robinson A, Genetic Programming and Autoconstructive Evolution with the Push Programming Language, Genetic Programming and Evolvable Machines 3, 2002, 7–40, doi: https://doi.org/10.1023/A:1014538503543

[8] Martínez Y, Trujillo L, Legrand P, Galvan-Lopez E, Prediction of Expected Performance for a Genetic Programming Classifier, Genetic Programming and Evolvable Machines 17, 2016, 409–449, doi: https://doi.org/10.1007/s10710-016-9265-9

[9] Trujillo L, Muñoz L, López U, Hernández D, Untapped Potential of Genetic Programming: Transfer Learning and Outlier Removal, Genetic Programming Theory and Practice XVI. Genetic and Evolutionary Computation. Springer,, 2019, 193-207, doi: https://doi.org/10.1007/978-3-030-04735-1_10

Article

A hybrid estimation of distribution algorithm for the quay crane scheduling problem

Ricardo Pérez-Rodríguez [1,*]

[1] CONACYT – UAQ Autonomous University of Queretaro, Faculty of Engineering
Cerro de las Campanas s/n C.P. 76010, Circuito Universitario, Queretaro, Qro, Mexico;
dr.ricardo.perez.rodriguez@gmail.com
* Correspondence: dr.ricardo.perez.rodriguez@gmail.com

Abstract: This paper considers the quay crane scheduling problem (QCSP). The objective is to determine the best sequence of discharging and loading operations in a ship by each quay crane. This problem is solved with a new hybrid Estimation of Distribution Algorithm (EDA). The proposed approach to tackle the drawbacks of the EDAs, i.e., lack of diversity of the solutions and poor ability of exploitation. The hybridization approach, used in this investigation, is based on a distance-based ranking model and the Moth-flame algorithm. The distance-based ranking model is in charge of modelling the solution space distribution, through an exponential function, that it measures the distance between solutions, meanwhile the heuristic Moth-flame is in charge to determine who would be the offspring, by a spiral function that it identifies the new locations for the new solutions. Based on the results, the proposed scheme, called QCEDA, works to enhance the results obtained from those other EDAs that use complex probability models. The dispersion results of QCEDA scheme is less than other algorithms, used in the comparison section. It means that the solutions found by the QCEDA are more concentrated around the best value, than other algorithms, i.e., the average of solutions of the QCEDA converges better than other approaches to the best-found value. Finally, it is possible to conclude that the hybrid EDAs have a better performance, or equal in effectiveness, than the called pure EDAs.

Keywords: Estimation of distribution algorithm, Mallows model, Moth-flame algorithm, job shop scheduling problem, quay crane scheduling problem

Un algoritmo coevolutivo para el problema binivel de localización del árbol de hubs con fijación de precios

M.C Carlos Eduardo Corpus [a],Dr.Fernando Camacho Vallejo[b],Dr.Víctor Blanco Izquierdo[c]

[a]Centro de Investigación Ciencias Físico Matemáticas
Facultad de Ciencias Físico Matemáticas
Nuevo León, México
carlos.corpuscrd@uanl.edu.mx

[b]Centro de Investigación Ciencias Físico Matemáticas
Facultad de Ciencias Físico Matemáticas
Nuevo León, México
jose.camachovl@uanl.edu.mx

[c]Departamento de Métodos Cuantitativos para la Economía y la Empresa
Facultad de Ciencias Económicas y Empresariales
Granada, España
vblanco@ugr.es

Debido a la globalización que ha sucedido en los últimos años, se han presentado nuevos e interesantes retos. El envío y recepción de bienes o productos es parte fundamental en muchos procesos, por lo que diseñar redes que permitan optimizar el flujo de información o bienes es cada vez más necesario. Para resolver esto, aquellas personas que se encargan de diseñar las redes de distribución utilizan centros especializados llamados hubs que se encargan de recibir los bienes y distribuirlos hacia su destino. Sin embargo, el extravío de bienes o el excesivo pago de cuotas por el uso de estas redes de distribución pueden complicar el proceso. La localización de estos hubs representa una parte fundamental del proceso de la creación de red. En los problemas clásicos de hubs se suelen tener objetivos orientados a la minimización de costos de construcción de la red, lo cual no siempre tiene que ser así. Así mismo, se considera un único tomador de decisiones que actúa en el problema. Sin embargo, en el proceso de la creación y uso de estas redes de distribución se pueden considerar a dos agentes involucrados. El primero es aquel que se encarga de localizar los hubs y crear la red. Por otro lado, está el agente que utiliza la red de la manera más conveniente. Es evidente que los objetivos de estos dos agentes son muy diferentes, mientras que el primero busca maximizar la ganancia obtenida por el uso de esta red, el otro busca minimizar el costo generado por usarla. Ante esta naturaleza jerárquica del problema, se utiliza un área especial de la programación matemática llamada programación binivel.

En esta plática se presentará el problema binivel de localización de hubs en forma de árbol con fijación de precios. Este problema tiene la peculiaridad de que el primer agente (el líder) no solo decide la estructura de la red de hubs en forma de árbol, sino que además decide los precios de utilizarla. Por otro lado, el segundo agente (el seguidor) reacciona a la decisión tomada por el líder y rutea los bienes enviados a través de la red a costo mínimo. Se detalla una reformulación del problema binivel en uno de un solo nivel y se muestran algunos resultados obtenidos, evidenciando las limitantes computacionales de dicha reformulación. Se discutirá brevemente sobre el diseño e implementación de un algoritmo coevolutivo para encontrar buenas soluciones a este problema complejo, además de mostrar posibles lineas de investigación a realizar.

References

[1] Contreras, Ivan & Fernandez, Elena & Marín, Alfredo., The Tree of Hubs Location Problem. European Journal of Operational Research. 202. 390-400. 10.1016/j.ejor.2009.05.044.

[2] Contreras, Ivan. (2015), Hub location problems, In Location science (pp. 311-344). Springer, Cham.

Roulette selection strategies applied to GRSA2-SSP for refinement of the amino acid side chain regions

Diego A. Soto-Monterrubio[1], Juan P. Sánchez-Hernández[2], Juan Frausto-Solís[1] , Juan J. González-Barbosa[1]

[1] Tecnológico Nacional de México/IT Ciudad Madero, México
[2] Universidad Politécnica del Estado de Morelos, México

Abstract: The proteins are macromolecules essentials for living organisms, formed by amino acids. The tertiary structure or Native Structure (NS) of a protein is a quasi-stable conformation in which the protein performs its biological functions. The Protein Folding Problem (PFP) consists of obtaining the NS of a protein knowing its amino acid sequence. In general, in the Literature is published that NS has the lowest Gibbs energy; there are some exceptions for the last statement; however, there are unknown constraints of nature that prevent this situation be fulfilled. One of the best computational methods for PFP in the case of peptides is GRSA2-SSP method, which for these kinds of instances has obtained excellent results. This method uses the golden search heuristic, and also secondary structure solutions for searching a structure close to the NS. Nevertheless, GRSA2-SSP has not been applied for optimizing the lateral amino acids chain; as a consequence its application for larger peptides not always provides good results. In this work, we enhance GRSA2-SSP in several important aspects, and we applied it to a dataset with small, medium, and large peptides. The results show that this method obtains statistically equivalent or better results than the best algorithms in the area. To achieve this we apply two steps: firstly, the GRSA2-SSP algorithm is applied to improve the general protein structure, and secondly, roulette selection strategies are applied to allow a good selection of the amino acid side chains regions for improving to find the structure of the protein with the lowest energy. We evaluate the performance of our method using the RMSD and TM-Score metrics and comparing the initial structure with the refined structure.

Keywords: Protein Folding Problem; Golden Ratio, Simulated Annealing

Addressing Violence Against Women and Children using IoT, AI, and Serious Games

Dalia A. Rodríguez[a], A. Díaz-Ramírez[a], R.I. Ayala Figueroa[a], L. Trujillo[b] , J.P. García-Vázquez[c]

[a]Department of Computer Systems
Tecnológico Nacional de México/IT de Mexicali
Mexicali, BC, México, 21375
{drodriguez, adiaz, rafaelivan}@itmexicali.edu.mx

[b]Department of Electrics and Electronics
Tecnológico Nacional de México/IT de Tijuana
Tijuana, B.C., México, 2414
leonardo.trujillo@tectijuana.edu.mx

[c]Faculty of Engineering
Universidad Autónoma de Baja California
Mexicali, BC, México, 21100
pablo.garcia@uabc.edu.mx

Violence against women and children are major public health problems. Approximately 30% of women worldwide have experienced some kind of violence. Each year one billion children are victims of sexual, physical, or emotional violence or neglect. The increased use of the Internet and social media has exposed women and children to new types of violence such as cyber-bullying and online sexual or emotional harassment. Recent years have seen the use of computer science and related technologies to address violence against women and children. In this talk, we will discuss the results from a systematic review of computer science-based solutions to address violence against women and children. The survey was developed using a literature review of academic documents published from 2010 to 2020. The contributions were categorized into four application domains: online detection, offline detection, safety, and education. These contributions were further categorized based on the computer science approaches and technologies used: artificial intelligence (AI), Internet of Things (IoT), and digital serious games. Additionally, we will discuss the use of machine learning techniques for text mining and natural language processing to detect misogynistic language against women in social networks.

References

[1] Dalia Andrea Rodríguez, et al., A Systematic Review of Computer Science Solutions for Addressing Violence Against Women and Children, IEEE Access 2021, Vol. 9, pp. 1-18, doi: 10.1109/ACCESS.2021.3103459

[2] N I Widiastuti, Convolution Neural Network for Text Mining and Natural Language Processing, OP Conference Series: Materials Science and Engineering, Volume 662, Issue 5, pp. 194-196, 2019, doi: 10.1088/1757-899x/662/5/052010

[3] Shervin Minaee, et al., Deep Learning Based Text Classification: A Comprehensive Review, ACM Computing Surveys, Volume 54, Issue 3, pp. 1-40, doi: 10.1145/3439726

Variable Decomposition for Large-scale Constrained Optimization Problems Using a Grouping Genetic Algorithm

Guadalupe Carmona-Arroyo, Marcela Quiroz-Castellanos, Efrén Mezura-Montes

Artificial Intelligence Research Institute
Universidad Veracruzana
Xalapa, Veracruz, Mexico
gcarmonaarroyo@gmail.com, maquiroz@uv.mx, emezura@uv.mx

Evolutionary Algorithms (EAs) have proven to be a good tool when solving optimization problems. However, EAs perform poorly as the dimension in the problem increases, in other words, when it comes to a Large-scale Optimization Problem. For the above, several strategies have been proposed to deal with large-scale problems. One of them is known as Cooperative Co-evolution (CC) [1], which is inspired by the divide and conquer approach and works basically in three stages: (1) the problem is decomposed into smaller problems, that is, the decision variables of the problem are separated into groups, (2) each of the subproblems is optimized by some EA, and then, (3) the solutions of each subproblem, given by the previous step, cooperate to create a complete solution of the original problem.
There are several variations of CC, in which different decomposition, optimization, and cooperation strategies have been evaluated [2]. In this work, we focus on the first stage: the decomposition of the problem. The challenge at this stage is that a proper decomposition of the problem will make the optimization path easier. Unlike, an inadequate decomposition could cause to obtain local solutions of the problem. Therefore, we are faced with an optimization problem: optimizing the variable grouping.
In this work, we propose a Grouping Genetic Algorithm (GGA) to optimize the group separation of the decision variables since these algorithms have proven to be one of the best strategies to solve combinatorial optimization problems where the optimization of certain groupings is sought [3]. The experiments were performed on 18 benchmark constrained functions in three different dimensions: 100, 500, and 1000 variables [4]. We have compared our proposal against a classical Genetic Algorithm (GA) created for the variable grouping in the same test set. Our proposal obtained better groupings than the GA, that is, the fitness value of the decomposition is lower in most cases. Additively the statistical results show a more stable behavior of our algorithm.
In conclusion, the results obtained indicate that a Grouping Genetic Algorithm is a suitable tool for the decomposition variable in Large-scale Constrained Optimization Problems.

References

[1] Potter, M. A., & De Jong, K. A. (1994, October). A cooperative coevolutionary approach to function optimization. In *International Conference on Parallel Problem Solving from Nature* (pp. 249-257). Springer, Berlin, Heidelberg.

[2] Ma, X., Li, X., Zhang, Q., Tang, K., Liang, Z., Xie, W., & Zhu, Z. (2018). A survey on cooperative co-evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 23(3), 421-441.

[3] Ramos-Figueroa, O., Quiroz-Castellanos, M., Mezura-Montes, E., & Schütze, O. (2020). Metaheuristics to solve grouping problems: A review and a case study. *Swarm and Evolutionary Computation*, 53, 100643.

[4] Sayed, E., Essam, D., Sarker, R., & Elsayed, S. (2015). Decomposition-based evolutionary algorithm for large scale constrained problems. *Information Sciences*, 316, 457-486.

Multi-objective Evolutionary Scheduling in Heterogeneous Parallel Environments

Miriam Pescador-Rojas[a], Erika Hernández-Rubio[b], Amilcar Meneses-Viveros[c]

Escuela Superior de Cómputo
Instituto Politécnico Nacional
Ciudad de México, México
[a]mpescadorr@ipn.mx, [b]ehernandezru@ipn.mx

[c]Departmento de Computación
Centro de Investigación y de Estudios Avanzados
Ciudad de México, México
ameneses@cs.cinvestav.mx

During the last decades, several strategies to solve the job-shop scheduling problem had been designed in diverse areas such as manufacturing, resource allocation, high-performance computing, cloud computing, and heterogeneous parallel process environment. Given n jobs and a heterogeneous computing system, the common goal is to minimize the makespan to dispatch all jobs on different processing units, such as CPUs, GPUs, or FPGA devices. The main idea of these systems is to increase the performance of applications running on them. However, for many problems, there is not a unique solution for the use of heterogeneous processing units, because the performance depends on the problem size. Moreover, a multi-objective scheduling problem introduces different criteria to optimize such as time and energy efficient. To obtain smarts strategies, we couple a multi-objective evolutionary algorithm based on decomposition (MOEA/D) to diverse scheduling policies to identify the set of the best solutions. In this work, we study the behavior of these evolutionary schedulers based on simulations of parallel tasks running in a heterogeneous parallel environment using two scenarios: a multi-core processor and one GPU, and two multi-core processors with two GPUs cards. The final solutions can be used to guide the decision of distributions of jobs depending on scenarios of heterogeneous environments.

Crop yield and Crop Dates Forecasting based on Biometrical and Weather Variables

R. Juárez[a], C. Borrego[a], G. Maldonado[b], A. Ponce[b]

[a]Instituto Politécnico Nacional
Campus Coahuila UPIIC
San Buenaventura, Coahuila. México
rjuarezdt@ipn.mx, cborregos@ipn.mx

[b]Universidad Autónoma de Aguascalientes
Centro de Ciencias Económicas y Administrativas
Aguascalientes, Aguascalientes. México
gmaldona@correo.uaa.mx, aponcem@correo.uaa.mx

1 Introduction

Maximizing yield in the precise dates while minimizing costs is the basic idea in this contribution. Early detection and management of problems limiting plant production can help to increase yields in the precise dates and thus obtain more income for the farm. Statistical models are the most common techniques used for forecasting the yield of crops which enable yield estimation during the growing season. Knowing the predicted crop size in a specific dates can be helpful in making decisions concerning the seasonal planning of cultivation or storage areas, and leads to a human and material resources acquisition or even an anticipated risk management plan. It is also possible to improve the profitability of a farm, as well as to balance the amount of production means used, such as human labor, machinery, material resources and also to react against plagues or unexpected weather conditions. In addition, an early and reliable monitoring of crop yields can help policymakers and agriculture marketing agencies to plan imports and exports. Crop yield and crop dates forecasting increases the profit of processing plants; forecasting tools provide valuable information on the basis of which it is possible to estimate the amount of raw material that these companies will be able to process. Crop prediction for the current growing season is a relatively difficult task due to the complexity of the relationship between the plant growth process and environmental factors such as weather. The application of statistical modeling methods such as MLR is always a basic and initial reference analysis. An essential element in predictive modeling is the accurate assessment of the correctness of the modelâĂŹs functionality. For this purpose, the root mean square error (RMSE) is used.

An artificial neuron is a computational model inspired in the natural neurons. These basically consist of inputs, which are multiplied by weights, and then computed by a mathematical function which determines the activation of the neuron. Another function computes the output of the artificial neuron. ANNs combine artificial neurons in order to process information. An essential feature of ANNs is the ability to learn by means of two variants of learning: supervised and unsupervised learning. The supervised learning process is based on the training set, which includes the learning cases along with the model answers provided to the model. This allows the network answers to be matched with the pattern answers. Training a neural network makes it capable of solving a task similar to the one on which it was trained. Unsupervised learning is based only on providing a series of sample inputs, without considering any information about the expected outputs. The ability to transfer the trained knowledge to new cases is known as generalization. Overfitting in generalization is a risk, which causes excessive fitting to irrelevant learning cases. Overfitting of the network results in poor generalization. One of the widely used comparative methods for ANN-based analyses is MLR. Over the years, ANNs have been successfully used to forecast the yields of agricultural

plants.

2 Variables description

Agriculture is one of many sectors that is vulnerable to ongoing climate change. Changes in precipitation and temperature fluctuations during the growing season may contribute to significant yield losses. As such, environmental factors are commonly used variables in predicting crop yields. The independent variable most often used in ANN models is temperature [78]: minimum temperature (TMN), maximum temperature (TMX), and average temperature. Temperature distribution during the growing season has the strongest impact on plant productivity. Another essential environmental factor determining the proper growth and development of plants is water and its availability in soil. Then, the distribution of precipitation should be considered while conducting research on forecasting models. To be useful as explanatory variables for prediction in regression models, weather variables should be observed at different stages of crop growth prior to harvest. This approach will obviously increase the number of explanatory variables in the prediction model. To estimate the corresponding regression coefficients will require a long series of data, which may not be available in practice. In the research devoted to plant yield forecasting with the use of artificial neural networks, the intensity of solar radiation and wind force have also been used as some of the explanatory variables. The most common predictors in ANN modeling were outlined above. However, the correctness of the prediction model depends not only on the quality of the data but also on the representativeness of the model. Data with outliers, incomplete sets, or erroneous significantly limit the forecasting model capabilities.

The forecasting of the yield of a crop at periodic intervals during the growing season is more difficult than estimating the yield at harvest time. Since the crop yield is a function of biometrical characteristics, an appropriate set of such characteristics having a profound influence on crop yield should be carefully selected for inclusion in the forecasting model. Apart from being useful indicators of the final outcome of the crop yield, these should be such as can be easily measured without much error well in advance of the harvest. To minimize multicollinearity, the biometrical characteristics selected should not have substantial correlation among themselves. Further, a decision in regard to the usefulness of a biometrical characteristic to serve as an indicator of a crop yield should be made on the basis of a study of the relationship between the yield and different biometrical characteristics. Finally, the cost of collection of observations on the biometrical characteristics selected on this basis should also be taken into consideration to ensure that their use is economically feasible on a large scale. The problem of choosing a suitable sampling design for the collection of data on crop yield and different biometrical characteristics has to be viewed in the light of some practical considerations. Frequent observations on biometrical characteristics during the growth period of the crop will be costly and time-consuming then it thus seems necessary to study in depth the effect of sampling design on regression models.

Rancho Medio Kilo RMK, in Aguascalientes Mexico, is a socially and ecologically responsible farm, which produces, processes and distributes vegetables such as Broccoli, cauliflower, carrot, pumpkin, chard, spinach, onion, lettuce, tomato, chili, etc. Conventional broccoli represents 39% of RMK's production for all its products. In harvest time RMK have to hire the necessary personnel, have to get available the machinery and all additional resources for the harvest process. In the RMK, the production surface is divided in Tables which are divided in four sections. See Figure 1. Each Table has around four hectares and each section has around one hectare. 106 and 194 broccoli varieties are randomly grown on may in Table 127. 106 and 194 broccoli varieties are also randomly grown on july in Table 143. Just 194 broccoli variety is grown on may in Table 108. Then 194 broccoli variety is the cold-resistant variety. Maximizing yield while minimizing costs and caring for the environment are the basic goals of agricultural production [1]. In this contribution a Broccoli yield and crop date forecasting is performed in order to make decisions concerning the seasonal planning of cultivation, and to balance the amount of production means used, such as human labor, machinery, resources anticipation, and also to be prepared to react against plagues or

unexpected weather conditions. The yield forecast evaluation is measured respect to variation interval of $[-15\%, 15\%]$. The crop date forecast evaluation is measured respect to a period of thirty days before the start of harvest. The forecasting model is performed by two approaches: MLR and ANN. The vegetative

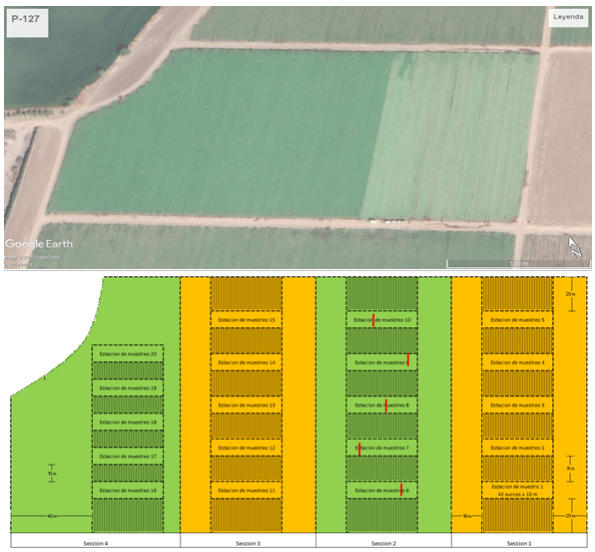


Figure 1: Broccoli table in Aguascalientes, Mexico. Satellite image and table drawing with sections and stations.

cycle of broccoli ranges from 58 to 120 days depending on the genetic characteristics of the varieties, agronomic management, and weather conditions at the time of planting, growing and harvesting. Jaramillo N. J; Diaz D.C. (2006) like Maroto, B.J.V .; Pomares F. and Baixauli C. (2007) say that from the germination of the seed to the seedling there are approximately 30 days. The seedling is characterized by leaves and roots formation, three or four well-formed leaves with 10-12cm height and it is ready for transplantation to the field. Since the seedlings have been transplanted to 40 days after the floral primordium has rised and is perfectly visualized. See Figura 2. The broccoli plant now has 70 days. The height, stem diameter, biomass, number of leaves and foliar area show a logarithmic increase there is also a proliferation of leaves. The canopy closure occurs around 35 days after transplantation and shows an accelerated development of the leaves for radiation capturing (Jaramillo N. J; DÃŋaz D.C. 2006). The floral head appears from 40 to 45 days after transplantation and when the plant has from 18 to 20 leaves. From this moment a linear growth begins for the plant and specially the floral head. It is confirmed by the leaf emission rate decrease, the leaf surface evolution rate decrease and the stem growth rate decrease. (Jaramillo N. J; DÃŋaz D.C. 2006) Hegarty (1979) cited by Maroto, B.J.V. (1989.), studied the influence of the physical conditions of the soil and humidity, on flower emergence. The inflorescence is presented when the flowers are not opened yet and it last from 20 to 25 days. The inflorescence presents an exponential growth in diameter and biomass. Around 55 days after transplantation a period of slow growth begins. Then, from 60 to 65 days after transplantation, a faster period appears until harvest. The translocation of photoassimilates occurs in this stage. Then the stem's diameter increases and height presents a second growth peak due to head's size increasing. (Jaramillo N. J; DÃŋaz D.C. 2006) Crop cycle depends on varieties, edaphoclimatic and water conditions, and even cultural and fertilization practices.

Figure 2: Broccoli life cycle.

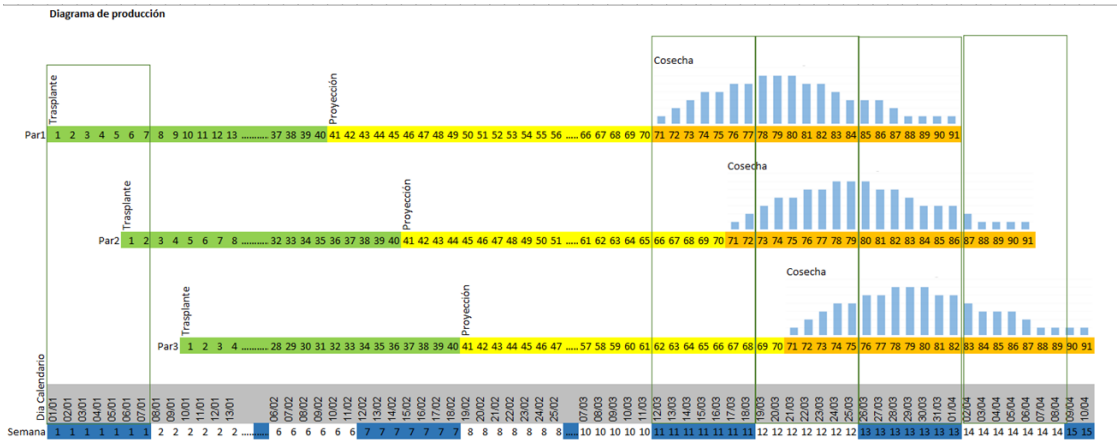


Figure 3: Crop yield duration.

References

[1] Patryk Hara, Magdalena Piekutowska, and Gniewko Niedbała. "Selection of Independent Variables for Crop Yield Prediction Using Artificial Neural Network Models with Remote Sensing Data". In: *Land* 10.6 (2021), p. 609.

A hybrid method for solving Dynamic Job Shop Scheduling Problem

Arsenio de Jesús Lizardi-Durán, Guadalupe Castilla Valdez,
Juan Frausto Solis, J. Javier González Barbosa

TecNM/IT Cd Madero, 1° de Mayo s/n Col. Los Mangos C.P.89440, Cd. Madero Tamaulipas, México

Abstract

The Job Shop Scheduling Problem (JSSP) is usually modeled as an optimization problem with linear and no-linear constraints. JSSP can be classified into static and dynamic problems. In the static version, the configuration and data remain the same for all the periods of the scheduling application; this version is called static JSSP or simply JSSP. The second version, known as DJSSP (Dynamic JSSP) is related to many factors is that can change in real problems. This paper presents the new algorithm or DJSSP named SAL-DJSSP that is based on Simulated Annealing that overtakes many of those of the literature.

Key Words: Dynamic Job Shop Scheduling, Simulated Annealing, Heuristics

Introduction

The Job Shop Scheduling Problem (JSSP) is modeled as an optimization problem with linear and no-linear constraints which is classified as an NP-hard problem [1]. JSSP can be classified into static and dynamic problems. In the static version, the configuration and data remain the same for all the periods of the scheduling application; this version is called static JSSP or simply JSSP. The second version, known as DJSSP or Dynamic JSSP, is related to many factors that can change in real problems: failure of machines, changes in cost, demand level for an emergency, and many other situations that may occur in real-life [2-3]. In other words, in DJSSP unexpected events occur as is common in real-world manufacturing systems.

Simulated Annealing (SA) is one of the best algorithms for single optimization [4-5] and was successfully applied to static optimization for single and multi-objective [6-7]. SA was designed for static optimization problems [4]. In this work, we propose a new algorithm named SAL-DJSSP based on SA for DJSSP that overtakes many of those of the literature [2] [3][8].

The paper is organized as follows: After a brief introduction in the first section about Dynamic Job Shop scheduling problems area, we present in section 2 the SAL-DJSSP algorithm and the material and methods on it. Then we present in section three the experimentation and results obtained with SAL-DJSSP. Finally, we present the conclusions and references at the end of the paper.

Materials and Methods

JSSP is a difficult problem for the two versions of JSSP and it is classified as an NP-Hard problem. JSSP has been extensively studied; so far no one has found that an exact method can solve it in polynomial time. Thus, metaheuristics have been developed for solving JSSP, and most of the literature is for static JSSP. There are few works for solving the Dynamic Job Shop Scheduling Problem (DJSSP). In the last problem, unexpected events occur as is common in real-world manufacturing systems.

In this paper, we address DJSSP with three types of dynamic events:

- machine breakdowns,
- new job arrivals, and
- changes in the processing time

We propose the new hybrid algorithm SAL-DJSSP based on simulated annealing which uses a local search strategy. SAL-DJSSP is tuned with the AndyMark method [9-10]. SAL-DJSSP has

well-designed strategies to achieve the best solution such a those related to escape from local optima. The way to design these strategies is presented in the final version of the paper.

Experimentation and Results

We evaluate the proposed algorithm by using a set of benchmark instances; the performance of SAL-DJSSP obtained with the experimentation is compared with the reported for in the literature. The numerical results obtained for this benchmark indicate that the proposed algorithm obtains the best general solutions.

The main reason for the superiority of SAL-DJSSP is that it profits from the advantages of the participant's methods; other elements that help SAL-DJSSP to achieve very good solutions for this dynamic problem are: a) a well mathematical definition of constraints and the objective function, b) the well-selected processes for escaping from local optima and c) the good data structures, and finally, d) the techniques for tuning the algorithm.

Conclusion

In this paper, we propose the SAL-DJSSP algorithm for solving Dynamic Job Shop problems. We also present the main features which help this algorithm overtakes those of the main area. In the final paper, we analyze the main reasons for the superiority of the SAL-DJSSP algorithm.

References

[1] Peter Brucker, Yu N. Sotskov, and F. Werner. The Complexity of shop-scheduling problems with a fixed number of jobs: a survey.Mathematical Methods of Operations Research, 65(3):461–481, 2007.

[2] Nilsen Kundakcı and Osman Kulak. Hybrid genetic algorithms for minimizing makespan in dynamic job shop scheduling problem.Computers & Industrial Engineering, 96:31–51, 2016.

[3] Shyh-Chang Lin, Erik D. Goodman, and William F. Punch III. A genetic algorithm approach to dynamic job shop scheduling problem. InICGA, pages 481–488, 1997.

[4] S. Kirkpatrick, C.Gelatt, M.Vecchi. Optimization by simulated annealing. Science, 671–680, 1983.

[5] Peter J. M. van Laarhoven and Emile H. L. Aarts. Simulated annealing: Mathematics Theory and Its Applications, pages 7–15. Springer Netherlands, 1987

[6] Hernández L., Frausto J., Castilla G., González-Barbosa, J., Terán D. & Morales L. A Hybrid Simulated Annealing for Job Shop Scheduling Problem. International J. of Combinatorial Optimization Problems and Informatics, 10(1), 6-15. https://ijcopi.org/ojs/article/view/111, 2018.

[7] Frausto J., L. Hernández, G. Castilla, J. González-Barbosa, and J, Sánchez. Chaotic Multi-Objective Simulated Annealing and Threshold Accepting for Job Shop Scheduling Problem. Math. Comput. Appl., EISSN 2297-8747, MDPI, 2021.

[8] Liping Zhang, Liang Gao, and Xinyu Li. A hybrid genetic algorithm and tabu search for a multi-objective dynamic job shop scheduling problem.International Journal of Production Research, 51(12):3516–3531, 2013.

[9] Frausto J., Sanvicente H., Imperial F. (2006) AndyMark: An Analytical Method to Establish Dynamically the Length of the Markov Chain in Simulated Annealing for the Satisfiability Problem. Lecture Notes in Computer Science, v. 4247. Springer, https://doi.org/10.1007/11903697_35, 2006.

[10] Frausto J., F. Martínez-Rios, Golden Annealing Method for Job Shop Scheduling Problem, in Foundations of Intelligent Systems, Springer book, Aijun An, Stan M., Zbigniew W. Ras, D.Slezak; ISBN 978-3-540-68123-6, https://www.springer.com/gp/book/9783540681229, 2008

The Pareto Tracer for Degenerated Multi-objective Optimization Problems

Oliver Cuate[a], **Oliver Schütze**[b]

[a] Mathematics Department of the Physics and Mathematics School (ESFM)
Instituto Politecnico Nacional
Avenida Instituto Politécnico Nacional s/n Edificio 9
Col. San Pedro Zacatenco Del, Nueva Industrial Vallejo, Gustavo A. Madero, 07738 Ciudad de México

[b]Computer Science Department
Cinvestav-IPN
Av. IPN 2508, 07260 Mexico City

In many real-world applications, the problem arises that several objectives have to be optimized at the same time leading to so-called multi-objective optimization problems (MOPs). Recently, the Pareto Tracer has been proposed, a continuation method that is capable to perform a movement along the Pareto set/front of a given MOP. In this presentation, we show some extensions of the Pareto Tracer to efficiently deal with degenerated MOPs, i.e., problems where not all of the considered objectives are in conflict with each other.

EMS-Simulator: an Open-Source Discrete-Event Simulation Framework for Medical Emergencies

Hans Yuan[*], Timothy Lam[†], Maurício C. de Oliveira[‡]
University of California, San Diego

Abstract

Computer simulations can be a powerful tool in the planning, evaluation and management of Emergency Medical Services (EMS). In this paper we introduce the open-source tool EMS-Simulator, which is a discrete-event highly customizable and flexible simulator framework built for simulation of medical emergencies. EMS-Simulator was written with object-oriented principles in Python for ease of development, modularity, and wide accessibility and has been released on the Python Package Index (PyPI) and GitHub. In order to illustrate a real-world application of the package we present a case study using data from the Red Cross in Tijuana (RCT), Mexico. In Tijuana, the Red Cross operates a small 13 ambulance fleet serving more than 2 million people. Severe financial limitations prevent the RCT from widely expanding their fleet, motivating the need to develop well-tested heuristics to improve the ambulance dispatch process. The simulations showed that while operations were able to meet a weak majority of cases within 14 minutes, a disaster would immediately deplete the RCT's ability to respond to calls.

Keywords: Emergency Medical Services; Ambulance Dispatch; Simulation Framework; Open-Source; Travel Times; Double Coverage Model;

[*] h3yuan@eng.ucsd.edu
[†] thl024@ucsd.edu
[‡] Mechanical and Aerospace Engineering Department. University of California San Diego. 9500 Gilman Dr, La Jolla, 92093-0411. mauricio@ucsd.edu

Evaluation of machine-learning algorithms for early diagnosis of deep venous thrombosis

Eduardo Enrique Contreras-Luján[a], Enrique Efrén García-Guerrero[a], Oscar Roberto López-Bonilla[a], Esteban Tlelo-Cuautle[c], Didier López-Mancilla and Everardo Inzunza-González[a]

[a]Faculty of Engineering, Architecture and Design
Autonomous University of Baja California
3917 Ensenada-Tijuana highway, Ensenada, Baja California, México, ZIP 22860.
eduardo.enrique.contreras.lujan@uabc.edu.mx, eegarcia@uabc.edu.mx,
olopez@uabc.edu.mx, einzunza@uabc.edu.mx

[b]Department of Electronics
Instituto Nacional de Astrofísica, Óptica y Electrónica
Luis Enrique Erro #1, Santa María Tonanzintla
San Andrés Cholula, Puebla, Mexico. Zip 72840.
etlelo@inaoep.mx

[c]Centro Universitario de Los Lagos
Universidad de Guadalajara
1144 Enrique Díaz de León, Col. Paseos de la Montaña
Guadalajara, Jalisco, Mexico.
didier.lopez@academicos.udg.mx

Abstract. Deep Venous Thrombosis (DVT) is a disease that must be diagnosed on time, otherwise the consequences can be fatal for the patients [1, 2]. Nowadays, we find different approaches in the determination of the condition, either by statistical determination, clinical scoring, D-dimer blood tests, infrared imaging, ultrasonography [3, 4]. Recently, scientists have attracted their attention in the use of machine-learning (M-L) and neural networks for the diagnosis of many diseases [5, 6, 7]. Some countries, such as the USA, Italy, UK, Germany, and Canada have been pioneers in carrying out work with artificial intelligence (AI) within the diagnosis and prediction of DVT, progressively increasing the percentage of accuracy and effectiveness, as time progresses. Suspected DVT patients have no apparent symptoms and failure to diagnose it could be fatal [8, 9]. In this paper, a new machine-learning model is proposed for the diagnosis of DVT, in an efficient, less invasive, and reliable way. This is based on pattern recognition techniques that help in a good timely diagnosis, as well as good-trained M-L models help in decision making and validating whether or not the patient suffers from this disease. The dataset used was developed through a data augmentation algorithm previously published in the state of the art [10]. The dataset is divided into 80% for training and 20% for evaluation by using k fold cross-validation to observe the performance at different times. This study includes the evaluation of several classifiers such as sequential models, Extra Trees Classifier, K-Nearest Neighbors Classifier, Decision Tree Classifier, SVM Classifier, and Random Forest Classifier. Finally, we propose the implementation of these M-L models on an embedded system of high-performance, to develop a smart system for early diagnosis of DVT, which is reliable, portable, open-source, and low-cost.
In this paper, we considered the performance metrics such as Accuracy, F1-Score, Precision, Recall, and Specificity. The performance of different M-L algorithms was evaluated, where the random forest classifier, achieved the highest accuracy of 89.6% and a specificity of 92.01%, compared to the metrics of angiologist physicians who obtain in a traditional way 73.82% and 71.43% respectively, proving to be a better way to diagnose DVT. The proposed smart-system could be useful for early diagnosis of DVT.

Keywords: Machine-learning; deep venous thrombosis; early diagnosis; smart system; embedded system.

References

[1] Nwosisi, C., Cha, S.-H., Yoo Jung An, Tappert, C. C., & Lipsitz, E. Constructing binary decision trees for predicting Deep Venous Thrombosis. 2010, 2nd International Conference on Software Technology and Engineering (ICSTE 2010). https://doi.org/10.1109/icste.2010.5608901

[2] Agharezaei, L., Agharezaei, Z., Nemati, A., Bahaadinbeigy, K., Keynia, F., Baneshi, M., Iranpour, A., & Agharezaei, M. The Prediction of the Risk Level of Pulmonary Embolism and Deep Vein Thrombosis through Artificial Neural Network. 2016, Acta Informatica Medica. https://doi.org/10.5455/aim.2016.24.354.359

[3] Tanno, R., Makropoulos, A., Arslan, S., Oktay, O., Mischkewitz, S., Al-Noor, F., Oppenheimer, J., Mandegaran, R., Kainz, B., & Heinrich, M. P. AutoDVT: Joint Real-Time Classification for Vein Compressibility Analysis in Deep Vein Thrombosis Ultrasound Diagnostics 2018, Springer International Publishing pp. 905–912 https://doi.org/10.1007/978-3-030-00934-2_100

[4] Yokomichi, A. L. Y., Rodrigues, V. da C., Moroz, A., Bertanha, M., Ribeiro, S. J. L., Deffune, E., & Moraes, M. L Detection of factor VIII and D-dimer biomarkers for venous thromboembolism diagnosis using electrochemistry immunosensor 2020, Talanta https://doi.org/10.1016/j.talanta.2020.121241

[5] Litjens G., Kooi T., Bejnordi B.E., Setio A.A.A., Ciompi F., Ghafoorian M., Van Der Laak JAWM, Van Ginneken B., Sánchez C.I. A survey on deep learning in medical image analysis. 2017, Med Image Anal.; pp. 42:60-88. doi: https://doi.org/10.1016/j.media.2017.07.005

[6] Nafee, T., Gibson, C. M., Travis, R., Yee, M. K., Kerneis, M., Chi, G., AlKhalfan, F., Hernandez, A. F., Hull, R. D., Cohen, A. T., Harrington, R. A., & Goldhaber, S. Z. Machine learning to predict venous thrombosis in acutely ill medical patients. 2020, Research and Practice in Thrombosis and Haemostasis, pp. 230–237 https://doi.org/10.1002/rth2.12292

[7] Willan, J., Katz, H., & Keeling, D. The use of artificial neural network analysis can improve the risk-stratification of patients presenting with suspected deep vein thrombosis. 2019, Journal of Haematology pp. 289–296 https://doi.org/10.1111/bjh.15780

[8] Liu, K., Chen, J., Zhang, K., Wang, S., & Li, X, A Diagnostic Prediction Model of Acute Symptomatic Portal Vein Thrombosis. 2019, Annals of Vascular Surgery, pp. 394–399 https://doi.org/10.1016/j.avsg.2019.04.037

[9] Kim, D. J., Byyny, R. L., Rice, C. A., Faragher, J. P., Nordenholz, K. E., Haukoos, J. S., Liao, M. M., & Kendall, J. L. Test Characteristics of Emergency Physician - Performed Limited Compression Ultrasound for Lower - Extremity Deep Vein Thrombosis 2016, The Journal of Emergency Medicine. pp. 684-690 https://doi.org/10.1016/j.jemermed.2016.07.013

[10] Fong-Mata, M. B., García-Guerrero, E. E., Mejía-Medina, D. A., López-Bonilla, O. R., Villarreal-Gómez, L. J., Zamora-Arellano, F., López-Mancilla, D., & Inzunza-González, E. An Artificial Neural Network Approach and a Data Augmentation Algorithm to Systematize the Diagnosis of Deep-Vein Thrombosis by Using Wells' Criteria. 2020, Electronics, https://doi.org/10.3390/electronics9111810

Attention measurement of an ASD user using EEG signals: A case study

José Jaime Esqueda-Elizondo[a,b], J. Reyes Juárez-Ramírez[b], Oscar Roberto López-Bonilla[a], Enrique Efrén García-Guerrero[a], Gilberto Manuel Galindo-Aldana[c], Laura Jiménez-Beristáin[b], Alejandra Serrano-Trujillo[b], Esteban Tlelo-Cuautle[d] and Everardo Inzunza-González[a]

[a]Faculty of Engineering, Architecture and Design
Autonomous University of Baja California
3917 Ensenada-Tijuana highway, Ensenada, Baja California, México, ZIP 22860.
olopez@uabc.edu.mx, eegarcia@uabc.edu.mx, einzunza@uabc.edu.mx

[b]Faculty of Chemical Sciences and Engineering
Autonomous University of Baja California
Technological Road and University, Tijuana, Baja California, México, ZIP 22390.
jjesqueda@uabc.edu.mx, reyesjua@uabc.edu.mx, ljimenezb@uabc.edu.mx, aserrano11@uabc.edu.mx

[c]Faculty of Engineering and Business
Autonomous University of Baja California
State Highway No. 3, Gutiérrez, ZIP 21720 Mexicali, B.C., México
gilberto.galindo.aldana@uabc.edu.mx

[d]Department of Electronics
Instituto Nacional de Astrofísica, Óptica y Electrónica
Luis Enrique Erro #1, Santa María Tonanzintla
San Andrés Cholula, Puebla, C.P. 72840.
etlelo@inaoep.mx

Abstract. Autism Spectrum Disorder (ASD) is a neurodevelopmental life condition characterized by problems with social interaction, low verbal and non-verbal communication skills, and repetitive and restricted behavior [1]. People with ASD usually have variable attention levels because they have hypersensitivity and all the environmental information is a problem for them [2, 3]. Attention is a process that occurs at the cognitive level and allows us to orient ourselves towards relevant stimuli, ignoring those that are not, and act accordingly. This paper presents a methodology based on electroencephalographic (EEG) signals for measuring the attention of a 13 years old boy with ASD. The EEG signals are acquired with an Epoc+ Brain-computer Interface (BCI) [4, 5, 6] via the Emotiv Pro platform while developing several learning activities and processed using Matlab 2019a and EEGLab software for artifact elimination. Then, the band power spectrum density is obtained in order to detect the theta-beta ratio (TBR), which is commonly used in attention detection and neurofeedback, as well as the theta relative power beta and theta/(alpha + beta) (TBAR) [7]. With these features, several machine-learning models will be trained and evaluated. Figure 1 shows the block diagram of the proposed method for attention measurement based on EEG signals. With these findings, better learning scenarios can be developed according to their needs. Also, quantifiable information on the progress of the person with ASD can be obtained to reinforce the perception of the teacher or therapist.

Keywords: Attention; ASD; EEG; Learning Activities; BCI; Features

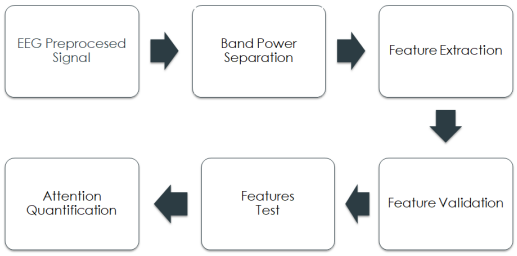


Figure 1: Block diagram for signal processing and feature extraction.

References

[1] American Psychiatric Association and American Psychiatric Association. DSM-5 Task Force Diagnostic and statistical manual of mental disorders : DSM-5, 2013 pp. 947 doi: www.appi.org/Diagnostic_and_Statistical_Manual_of_Mental_Disorders_DSM-5_Fifth_Edition

[2] Amaral C., Mouga S., Simões M., Pereira H.C., Bernardino I., Quental H., Playle R., McNamara R., Oliveira G., Castelo-Branco M. A feasibility clinical trial to improve social attention in Autistic Spectrum Disorder (ASD) using a brain computer interface, Frontiers in Neuroscience, 2018. doi: https://doi.org/10.3389/fnins.2018.00477

[3] Alahbabi Moza, Almazroei Fatima, Almarzoqi Mariam, Almeheri Aysha, Alkabi Mariam, Nuaimi A Al, Cappuccio Massimiliano, Alnajjar Fady Avatar based interaction therapy: A potential therapeutic approach for children with Autism 2017 IEEE International Conference on Mechatronics and Automation (ICMA) pp. 480–484 doi: https://doi.org/10.1109/ICMA.2017.8015864

[4] Lau-Zhu, A. and Lau, M.P.H. and McLoughlin, G. Mobile EEG in research on neurodevelopmental disorders: Opportunities and challenges Developmental Cognitive Neuroscience, 2019 doi: https://doi.org/10.1016/j.dcn.2019.100635

[5] Mehmood F., Ayaz Y., Ali S., Amadeu R De Cassia, Sadia H. Dominance in Visual Space of ASD Children Using Multi-Robot Joint Attention Integrated Distributed Imitation System IEEE Access, 2019 doi: https://doi.org/10.1109/ACCESS.2019.2951366

[6] Wang Hui, Song Quanjun, Ma Tingting, Cao Huibin, Sun Yuxiang Study on Brain-Computer Interface Based on Mental Tasks, 2015 The 5th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems, 2015 pp. 841–845 doi: https://doi.org/10.1109/CYBER.2015.7288053

[7] Fahimi Fatemeh, Guan Cuntai, Goh Wooi Boon, Ang Kai Keng, Lim Choon Guan, Lee Tih Shih Personalized features for attention detection in children with Attention Deficit Hyperactivity Disorder Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2017 pp. 414–417 doi: https://doi.org/10.1109/EMBC.2017.8036850

Evolving the Upper Confidence Bounds for Trees in Monte Carlo Tree Search

Edgar Galván[a], Gavin Simpson[a], Haseena Sreedharan[a]
Marcela Quiroz-Castellanos[b], Leonardo Trujillo[c]

[a]Naturally Inspired Computation Research Group
Department of Computer Science
Maynooth University, Lero, Ireland
edgar.galvan@mu.ie; gavin.simpson.2021@mumail.ie; sreedharan.haseena@gmail.com

[b]Universidad Veracruzana
Artificial Intelligence Research Institute, Xalapa, México
maquiroz@uv.mx

[c]Tecnológico Nacional de México
IT de Tijuana, México
Leonardo.trujillo@tectijuana.edu.mx

Abstract

Monte Carlo Tree Search (MCTS) is a sampling best-first method to search for optimal decisions. MCTS has attained incredible results in the Game of Go in recent years: a game that was thought to be unsolvable by Artificial Intelligence methods. A key element for the success or failure of MCTS is how its statistical tree is built. This tree is the result of the tree policy employed in MCTS. The vast majority of robust MCTS methods use the Upper Confidence Bounds for trees, commonly referred as UCT. In this work, we evolve this expression by means of Evolutionary Algorithms and show how the evolved solutions can be as competitive as the MCTS UCT.

An Experimental Study of Grouping Crossover Operators for the Bin Packing Problem

Stephanie Amador-Larrea [a], Marcela Quiroz-Castellanos [a], Guillermo-de-Jesús Hoyos-Rivera [a]

[a]Artificial Intelligence Research Institute
Universidad Veracruzana
Lomas del Estadio s/n, Xalapa, Veracruz, México.
stephanieamadorlarrea@gmail.com, maquiroz@uv.mx, ghoyos@uv.mx

In the last three decades, the one-dimensional Bin Packing Problem (BPP) has become one of the most studied problems in the area of combinatorial optimization, due to its diverse applications in logistics and industry. Since it is considered an NP-Hard problem, its computational complexity makes traditional solution heuristics non-functional. Among the methods used for its solution are the Grouping Genetic Algorithms, highlighting the performance of the Grouping Genetic Algorithm with Controlled Gene Transmission (GGA-CGT), considered as one of the best in the state-of-the-art for the solution of BPP. GGA-CGT has the particularity of using the variation operators in a controlled way for the induction of the fullest-bin pattern. In this work, we aim to highlight the impact that the crossover operator can have on the final performance of the GGA-CGT algorithm. We present an experimental study for the comparison of the performance of the GGA-CGT algorithm with the grouping crossover operators within the state-of-the- art: Uniform crossover, Exon Shuffling crossover, Greedy Partition crossover and Gene-level crossover. These operators were adapted to the characteristics of the problem and studied to measure their performance, and then compared with the results obtained by the original GGA-CGT algorithm, which has the Gene-level Crossover operator as the predetermined crossover operator. The experiments consisted in implementing the aforementioned GGA-CGT algorithm crossover operators, using different percentages of the population to cross, in order to select the two that solved the highest number of instances in an optimal way. Then, the implementation of the two previously elected operators was performed modifying the GGA-CGT algorithm with respect to the replacement of the solutions, in order to measure their performance and select the best configurations. Finally, the robustness of the operators was studied, with different seeds and the best crossover rates. The experimental results indicate that the Gene-level crossover operator proves to have a greater impact in terms of the number of optimal solutions found, being 1602 out of 1615 benchmark instances, outperforming the other operators for the Hard28 instance class, which has shown the highest degree of difficulty for the BPP algorithms.

The dynamic bin packing optimization problem: open challenges

Jessica González-San Martín[a], Laura Cruz-Reyes[a], Bernabé Dorronsoro[b],

Marcela Quiroz-Castellanos[c], Claudia Gómez-Santillán[a], Nelson Rangel-Valdez[a],

Héctor Fraire[a]

[a] Division of Postgraduate Studies and Research
Instituto Tecnológico de Ciudad Madero
Av. 1o. de Mayo esq. Sor Juana Inés de la Cruz s/n Col.
Los Mangos, Cd. Madero, Tamaulipas, México
jessica_san_martin@hotmail.com
cruzreyeslaura@gmail.com
rangel.nelson1980@gmail.com
cggs71@hotmail.com
hector.fraire2014@gmail.com

[b] Computer Science Engineering Department
University of Cadiz
Avenida de la Universidad, 10
11519 Puerto Real, Cadiz, Spain
bernabe.dorronsoro@uca.es

[c] Artificial Intelligence Research Center
Universidad Veracruzana
Calle Salvador Díaz Mirón 35, Zona Universitaria,
Xalapa-Enríquez, Veracruz, México.
maquiroz@uv.mx

The Bin Packing Problem (BPP) is a classic optimization problem that is known for its applicability and complexity, which belongs to a special class of problems called NP-hard, in which, given a set of items of variable size, we search to accommodate them inside fixed size containers, seeking to optimize the number of containers to be used, that is, using the least number of containers to place the largest number of items possible. For this problem, there are different variants such as one-dimensional (1D-BPP), 2-dimensional (2D-BPP), or 3-dimensional (3D-BPP) bin packing, multiobjective bin packing (MBP), dynamic bin packing (DBP), among others. The BPP has been preserved as a current study problem due to the various applications that it offers; therefore, in the recent state of the art, there are different algorithms, mainly heuristics, for solving the problem, however, currently, there are still open aspects of this problem that have not been studied in-depth, such as changes in the environment (dynamism). To date, there is no heuristic or metaheuristic algorithm capable of finding the optimal solution for all possible instances of a problem of this type despite the efforts of the scientific community. This work presents a review of DBP optimization model and methods to solve it; we identify those methods that provide a better impact on performance and highlight future lines of research on this problem. We also present the GGA-CGT-II algorithm as a case study extended from the Grouping Genetic Algorithm with Controlled Gene Transmission (GGA-CGT) [1] algorithm. In the proposed version, two different strategies were implemented: a) Method to calculate various theoretical lower limits of the optimal solution and select the best one for each instance and b) Instance reduction method to simplify the problem. We conduct extensive experimentation to observe their impact on performance in isolation and then together. The new version of GGA-CGT showed an increment in its performance with benchmark instances for the 1D-BPP.
With this work, it was possible to identify some promising lines of research for the dynamic bin packing [2,3].

References

[1] Quiroz-Castellanos M., Cruz-Reyes L., Torres-Jimenez J., Gómez S. C., F. Huacuja H., & Alvim A.: A grouping genetic algorithm with controlled gene transmission for the bin packing problem. Computational Operation Research pp.52–64 (2015).
[2] Spencer, K. Y., Tsvetkov, P. V., & Jarrell, J. J: A greedy memetic algorithm for a multiobjective dynamic bin packing problem for storing cooling objects Journal of Heuristics, 25(1), 1-45. (2018).
[3] Berndt, Sebastian, Klaus Jansen, and Kim-Manuel Klein: Fully dynamic bin packing revisited. Mathematical Programming 179.1: 109-155. (2020).

GA-DRL: PORTFOLIO SELECTION AND DEEP REINFORCEMENT LEARNING AGENT FOR DAILY STOCK TRADING

Irving O. Rodriguez Hernandez[a], Juan Frausto-Solis[a*] Juan J. González-Barbosa[a] José L. Purata Aldaz[a] Luis F. Cisneros-Sinencio[a]

[a]División de Estudios de Postgrado e Investigación
TECNM/INSTITUTO TECNOLÓGICO DE CIUDAD MADERO
Av. 1o de Mayo esq. Sor Juana Ines de la Cruz s/n Col. Los Mangos C.P.89440
* juan.frausto@gmail.com

Abstract

This work proposes GA-DRL which is applied to New York Stock Exchange (NYSE) sectors. GA-DRL uses Sharpe Ratio (SR) model through a Genetic Algorithm (GA) for portfolio selection followed by a multi-variable powered Deep Reinforcement Learning (DRL) agent for daily portfolio trading. This method builds a diversified and asset number constrained portfolio by optimizing the relationship between expected values and price movements correlations; then GA-DRL trains a DRL agent in a day trading stock market environment for developing a profitable trading strategy. The results show that GA-DRL overtakes, commonly used by investors, buy-and-hold strategies.
Keywords: Investment Portfolio, Deep Reinforcement Learning, Markov Decision Processes

1 Introduction

Portfolio selection and trading are highly time-consuming and complex tasks. GA-DRL integrates a diversified and asset number constrained portfolio by using GenPo-Sharpe [1]; then it trains a DRL agent [2] in a day trading stock market environment intending to develop a profitable trading strategy. The results show GA-DRL overtakes buy-and-hold strategies. Trading in financial markets implies investment strategies that evolve according to market states to alpha generation from a sequence of market decisions (i.e. buy and sell). This work formulates the problem as a Markov Decision Process (MDP) to which an optimal trading policy is pursued.

2 Proposed method

Our MDP is defined as a tuple $M = (S, A, R, \gamma)$ where market features are system states S, trading positions are available actions A to be executed for an agent, R is a reward signal that guides agent's behavior (policy) improvement and γ weighs the effect of immediate and future agent actions. We aim finding an agent's policy that maximizes long-term returns of $\sum_{k=0}^{+\infty} \gamma^k R(s_k, \pi(s_k)) \mid s_0$. Where π is a policy belonging to whole policies set Π, k is a timestep index, s_k represents k timestep state, $\pi(s_k)$ defines an action for s_k while following π; R is a reward function and s_0 is the initial system's state. The First GA-DRL method main step is concerned with portfolio selection by using GenPo-Sharpe [1] which selects a reduced amount of stocks, maximizes portfolio expected return, and minimizes the risk of it as in (1):

$$\max SR = \frac{E(R_{pf}) - R_f}{V(x)}; \sum_{i=1}^{n} x_i = 1; \sum_{i=1}^{n} E(R_{pf}) > R_f; 0 < x_i \leq 1$$

Where $E(R_{pf})$ is the portfolio expected return, $E(R_f)$ is the return of a free-risk investment as treasury bills, $V(x)$ is the variance of the portfolio and x_i represents the capital percentage assigned to i stock.

The GA-DRL method next step solves above mentioned MDP by a DRL agent composed of two Multi-Layer Perceptron (MLP), in an Actor-Critic architecture, one for action execution (actor) and the other for action evaluation (critic); both MLPs shares the same traits in input (MDP state components) and inner layers (fully connected layers) using $Tanh$ activation functions. These MLPs differ on output layer, while actor estimates N function distribution means, that are used to calculate capital assigned to each stock, critic estimates value function $V(S_t)$. We also implement a Proximal Policy Optimization [3] (PPO) algorithm for optimizing the interaction between the agent and environment model.

3 Results

Used datasets consist of daily time series from already mentioned sectors of NYSE. For portfolio selection, we use close prices changes time series from 01-01-2020 to 30-09-2020. For portfolio trading, the stocks time series used are based on previously optimized portfolio and consists of market variables: such as prices, volume, and technical indicators such as simple moving average and relative strength index. We use for training (80% data) and validation (20% data) 01-10-20 to 01-01-2021 timeframe and testing from 02-01-2021 to 01-02-2021. In the carried experimentations we have calculated annualized return (AR) as a performance measure. GA-DRL is competing with GenPo-Sharpe [1], and Equal Weighted (EQ) buy-and-hold investment strategies, having on average 48% higher returns for Technology, 638% higher returns for Health, and 4% higher returns for Energy NYSE sectors

4 Conclusions and future works

This paper presents a GA-DRL method for daily stock trading that showed better AR for Technology, Health, and Energy NYSE sectors from NYSE in comparison to GenPo-Sharpe [1] and EQ strategies. As future work, we see a wide opportunity to implement various Neural Networks architectures and forecasting techniques.

References

[1] Frausto, J., Gonzalez, J., Castilla, G., Purata, J., Soto, D., and Hernandez, L., GenPo-Sharpe: Stock Selection for Investing Portfolio usinga Genetic Algorithm with Sharpe Ratio Applied to Mexican Stock Exchange, NEO 2020, (págs. 92-93).

[2] Sutton, R. S., and Barto, A. G. , Reinforcement Learning: An Introduction, 2018 Cambridge: MIT Press.

[3] Schulman, J. W., Proximal Policy Optimization Algorithms, 2017 Computing Research Repository.

A multi-objective approach to deal with emotions in computer-aided musical composition

Natan Vilchis, Adriana Lara

Escuela Superior de Física y Matemáticas
del Instituto Politécnico Nacional
Av. Luis Enrique Erro S/N, Unidad Profesional Adolfo López Mateos, Zacatenco
Alcaldía Gustavo A. Madero, C.P. 07738, Ciudad de México, México
nvilchist1300@alumno.ipn.mx,alaral@ipn.mx

For computers, designing music is challenging due to the complex use of notes, chords, rhythms, melodies, and other factors to transmit particular human emotions like happiness or calm. We present here a multi-objective optimization model to maximize the musical quality of a music fragment composition. This work combines data mining [1] [2] with contemporary musical theory [3] [4] [5] to extract essential features for describing emotions in melodies. Our analysis was carried out from a data set of songs by different authors, musical genres, rhythms, and ages and later divided into three large groups: those that transmit a negative, positive or neutral emotion [6] [7]. Finally, we present the numerical and acoustic obtained results.

References

[1] Das, Sanchali & Bhattacharyya, Bidyut K. & Debbarma, Swapan *Building a computational model for mood classification of music by integrating an asymptotic approach with the machine learning techniques*, Journal of Ambient Intelligence and Humanized Computing, volume 12, 2021, `https://doi.org/10.1007/s12652-020-02145-1`

[2] Panda, Renato and Malheiro, Ricardo and Paiva, Rui Pedro, Novel audio features for music emotion recognition, IEEE Transactions on Affective Computing volume 11, number 4, pages 614–626, 2018, IEEE.

[3] Mauro de Maria, *Composición integral*, `https://www.maurodemaria.com/`

[4] Mauro de Maria, *18 dimensiones emocionales. Dimensiones rítmicas. Síntesis completa.*, `https://www.maurodemaria.com/`

[5] Mauro de Maria, *18 dimensiones emocionales. Decodificando las emociones de la música desde 18 dimensiones técnicas*, `https://www.maurodemaria.com/`

[6] Scirea, Marco and Togelius, Julian and Eklund, Peter and Risi, Sebastian, *Affective evolutionary music composition with MetaCompose*, Genetic Programming and Evolvable Machines, volume 18, number 4, pages 433–465, 2018

[7] Panda, Renato and Malheiro, Ricardo and Paiva, Rui Pedro, *Novel audio features for music emotion recognition*, IEEE Transactions on Affective Computing, volume 11, number 4, pages 614–626, 2018

An Algorithm Based in the Path Relinking Metaheuristic for Solving a Location Bilevel Problem: an Application to a Social Responsability Project.

José-Fernando Camacho-Vallejo[a], Sahori Polino[a], Juan G. Villegas[b]

[a]Facultad de Ciencias Físico-Matemáticas
Universidad Autónoma de Nuevo León
Ave. Universidad s/n, San Nicolás de los Garza, Nuevo León
jose.camachovl@uanl.edu.mx

[b]Departamento de Ingeniería Industrial
Universidad de Antioquia
Calle 70 No. 52-21, Medellin, Colombia
juan.villegas@udea.edu.co

In this presentation, a capacitated facility location problem is presented, in which the allocation of the users is based on their preferences. The problem arises from a social responsibility project, which involves a situation where the government aims to offer extracurricular workshops for students from Elementary school. Those workshops promote the development of academic, sports, and social skills of the students.

This problem is formulated as a bilevel program. The government is associated with the upper level, which aims to minimize the costs for opening workshops in different schools and the costs of allocating students to each workshop. The lower level is related to the parents association that seeks to optimize the global preferences of the students toward offered workshops.

To solve this problem, an algorithm based on the Path Relinking is proposed. This algorithm handles two *saturated* random solutions. Then, repairing and adjustment phase are applied for both solutions. Once the resulting solutions are obtained after both phases, the trajectory between the two solutions is generated. The main idea is to modify a component of the initial solution to convert it into the guide solution. The adjustment phase will only be applied to the best element found in the trajectory. After that, a local search based on a neighborhood of interchanges is applied. The obtained solution is compared with the incumbent and it is updated -if necessary. Hence, another *saturated* solution is generated and the process is repeated. The new trajectory is traced between the incumbent solution and the new generated one. The algorithm stops when a predetermined number of iterations has been reached.

Computational experimentation is conducted and the results indicate that the proposed algorithm is effective and efficient. The obtained results provide better bounds than the ones obtained by the typical single-level reformulation. Also, the required computational time is significantly reduced (in the small-size instances in which the comparison is possible). Finally, the proposed algorithm is applied to a realistic case-study of Apodaca municipality in Nuevo Leon, Mexico.

Genetic Knowledge Discovery Algorithm Based on Archimedean-Compensatory Fuzzy Logic

Citlalli Morales de la Cruz [a], **Laura Cruz-Reyes** [a]
Rafael Alejandro Espín-Andrade [b], **José Fernando Padrón-Tristán** [a]
Claudia Guadalupe Gómez-Santillán [a]

[a]Division of Postgraduate Studies and Research
Tecnológico Nacional de México, Instituto Tecnológico de Ciudad Madero
Av. 1o. de Mayo esq. Sor Juana Inés de la Cruz s/n
Col. Los Mangos, Cd. Madero, Tamulipas, México
g13071091@itcm.edu.mx, cruzreyeslaura@gmail.com
jfpt76@gmail.com,cggs71@gmail.com

[b]Autonomous University of Coahuila
Blvd. V. Carranza s/n. Col. República Oriente
C.P.25280 Saltillo, Coahuila, México
rafaelalejandroespinandrade@gmail.com

Obtaining knowledge is a problem that aims to find patterns within large volumes of data to create prediction models under limitations such as the imprecision of the information, the representation, and its interpretability.

As described above, some methods rely on techniques based on artificial intelligence such as fuzzy logic, neural networks, and metaheuristic algorithms that have made it possible to address the problem of obtaining knowledge through the use of fuzzy predicates based on a verbal expression made up of linguistic variables translated into the predicate calculus language [1]. However, most of these works are limited since they do not present a formal description of the proposed algorithm and lack a comparative analysis. Within this work, a modification of two genetic algorithms developed by Llorente in 2019[2] is presented, which carries out the first algorithmic effort documented with formality and with experimental support; These algorithms seek to generate knowledge through fuzzy predicates represented by decision trees combining genetic programming.

This work achieves an integration of these two algorithms under a new methodology based on the Compensatory Archimedean Fuzzy Logic instead of the Compensatory Fuzzy Logic, giving rise to the EU-ACFL-GP algorithm.

The first algorithm, called Evolutionary Optimization of Generalized Sigmoidal Function (KD-CFL) [2], takes a set of data and a set of membership functions then create a random population using decision trees by genetic programming. Finally returns the best individuals or the best individual in case of not finding the truth value selected.

The second genetic algorithm called Generalized Membership Function Parameter Optimization (EO-GSF) [2] uses the sigmoid function through a series of families; This function is composed of three parameters that are: β that represents the minimum value before zero, γ defines the diffuse value 0.5, and m that shows the tendency of the function to be a positive, negative sigmoid or convex function; the algorithm discovers these parameters allowing to improve the quality of the predicates to obtain values with high degrees of truth.

This work also foresees the implementation of a new genetic mutation operator that allows the balance between intensification and diversification of the search called constructive mutation, choosing those individuals that remain after selection and crossing operations applying Hill-Climbing local search to them generating individuals until obtaining better individuals than the parents. The advantage of using constructive mutation is that the destructive nature of the mutation operation transfers only the best individuals

than parents to the next generation. Another advantage is that the mutant parents participate after crossing and selection; therefore, individuals who are not good at producing better offspring than their parents mutate [3].
On the other hand, the Method of Increasing Completeness for the Predicate Search uses the concept of the symbolic predicate, which is a form of representation of a fuzzy predicate in which the asterisk (*) the method can use as a wildcard substitution of linguistic variables or unknown fuzzy predicates [4]. This method generates predicates from structures with multiple wildcards concerning predicates that use a single wildcard. This structure has the following advantages: it is more expressive by allowing the user to formulate a symbolic predicate that incorporates knowledge of the discovery problem at different levels and is more comprehensive by allowing search in new areas.
Finally, the reference algorithms [2] use the Compensatory Fuzzy Logic methodology. This work changes to the LDAC, which is the unification of two different types of fuzzy logical systems [5]: Archimedean logic, which represents the logic of non-refutation, and the compensatory logic that figures as the logic of affirmation.
The proposed algorithm uses statistical tests for tests of statistical significance [6]; for the reference algorithm, the proposed algorithm with the implementation of the constructive mutation and Method of Increasing Completeness for the Predicate Search methods obtained a 7% increase in the truth value. Similarly, the impact of the algorithm is evaluated when it includes mutation with and without the use of multiple wildcards. The results of the second option for the first reveal the following: an increase of 2% over the truth value, the execution time is reduced by 56%, the number of variables found in the fuzzy predicates is increased by 30%, and the number of predicate structures is equivalent in both methods; however, this significance depends on the type of operator involved.
In conclusion, the number of predicate structures and several variables is an indicator that was proposed to measure the exhaustiveness of the search in the generation of diverse predicates. With this work, it was possible to identify some promising research and works that could be addressed in the future: 1) use local search strategies and new functions that complement the truth value. 2) it is proposed to develop new classification methods that integrate predicates of different classes with different truth values.

References

[1] Cejas-Montero, J,). La lógica difusa compensatoria/The compensatory fuzzy logic. Ingeniería Industrial, 32(2), 157-161, 2011

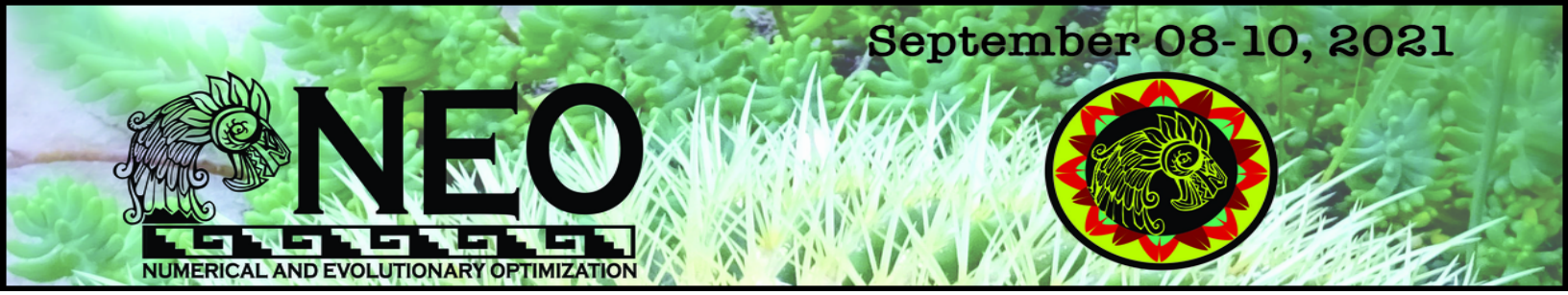
[2] Llorente, C, Evolutionary algorithm for discovering association knowledge using compensatory fuzzy logic. Technological Institute of Ciudad Madero, 2019

[3] Bhardwaj, A, Tiware, A, Varma, M, 'I&' Krishna M, Classification of egg signals using a novel genetic programming approach. Proceedings of the companion publication of the 2014 annual conference on genetic and evolutionary computation, 2011

[4] Eurekas, Eurekas Community, 2021

[5] Espin-Andrade, R, Caballero E, Pedrycz, W,'I&' González, E, Archimedean-Compensatory Fuzzy Logic Systems. International Journal of Computational Intelligence Systems, 2015

[6] Rodríguez-Fdez, I, Canosa A, Mucientes, M, 'I&'Bugarín, A,). STAC: a web platform for the comparison of algorithms using statistical tests. En Proceedings of the 2015 ieee international conference on fuzzy systems (fuzz-ieee), 2015



NEO 2021 Organizing Committee

- **General Chairs**
  Marcela Quiroz, Universidad Veracruzana, México
  Luis Gerardo de la Fraga, Cinvestav-IPN, Mexico City, Mexico

- **Publication Chair**
  Adriana Lara, ESFM - IPN, Mexico City

- **Publicity Chair**
  Marcela Quiroz, Universidad Veracruzana, México

- **Students Chair**
  Juan Gabriel Ruiz Ruiz, Universidad de la Sierra Juárez, Oaxaca, México

- **Series Chairs**
  Oliver Schuetze, Cinvestav-IPN, Mexico
  Leonardo Trujillo, Instituto Tecnológico de Tijuana, México

Index

J

L

M

N

P

Q

R

S

T

V

Y