Final report October 2016

ACOBSEC

Over the last decade, Human-Computer Interaction (HCI) has grown and matured as a field. Gone are the days when only a mouse and keyboard could be used to interact with a computer. The most ambitious of such interfaces are Brain-Computer Interaction (BCI) systems. BCI's goal is to allow a person to interact with an artificial system using brain activity. A common approach towards BCI is to analyze, categorize and interpret Electroencephalography (EEG) signals in such a way that they alter the state of a computer. ACoBSEC's objective is to study the development of computer systems for the automatic analysis and classification of mental states of vigilance; i.e., a person's state of alertness. Such a task is relevant to diverse domains, where a person is required to be in a particular state. This problem is not a trivial one. In fact, EEG signals are known to be noisy, irregular and tend to vary from person to person, making the development of general techniques a very difficult scientific endeavor. Our aim is to develop new search and optimization strategies, based on evolutionary computation (EC) and genetic programming (GP) for the automatic induction of efficient and accurate classifiers. EC and GP are search techniques that can reach good solutions in multi-modal, non-differentiable and discontinuous spaces; and such is the case for the problem addressed here. This project combines the expertise of research partners from five converging fields:
Classification, Neurosciences (University of Bordeaux), Signal Processing, Evolutionary Computation and Parallel Computing in Europe (France Inria, Portugal FCUL, Spain UNEX) and South America (Mexico ITT, CICESE). The exchange program goals and milestones give a comprehensive strategy for the strengthening of current scientific relations amongst partners, as well as for the construction of long-lasting scientific relationships that produce high quality theoretical and applied research.

Introduction:

Our aim was to develop new search and optimization strategies, based on evolutionary computation (EC) and genetic programming (GP) for the automatic induction of efficient and accurate classifiers. EC and GP are search techniques that can reach good solutions in highly multi-modal, non-differentiable and discontinuous spaces; and such is the case for the problem addressed here: the detection of mental states of vigilance. This project combines the expertise of researcher partners from five converging fields: Classification, Neurosciences, Signal Processing, Evolutionary Computation and Parallel Computing, in Europe - France (Inria, University of Bordeaux), Portugal (BioISI) and Spain (UNEX) - and North America - Mexico (ITT and CICESE). The partners complement and enhance their respective disciplines, allowing for a strong multi-disciplinary collaboration and proposal development. The exchange program goals and milestones gave a comprehensive strategy for the strengthening of current scientific relations amongst the partners, and allowed the construction of long-lasting scientific relationships that produced high quality theoretical and applied research.

Description of work packages:

WP1: Classification with Evolutionary Computation

The objective was to develop EC/GP-based classifiers for noisy real world data that exploit the advantages of statistical based approaches while also encouraging the explorative search capabilities of meta-heuristic based search strategies. It is now obvious that in order to solve difficult problems, hybrid approaches such as this could leverage the main strengths of each domain and possibly minimize some of their main weaknesses. Such algorithms have been useful in the particular sense of this project, where biological data and signals are well-known for their high levels of noise, heterogeneous frequency content, chaotic nature and dynamical properties. However, in a broader sense, developing such algorithms facilitated the use of hyper-heuristic based inductive learning in other real-world domain based on the strong cooperation of INRIA, FCUL and ITT.

In this WP, the participants established a common background context and promising lines of research in EC/GP-based classification, the paradigm used to analyze and interpret biological data and signals, the main goals of the proposal. In this sense, it is closely related to WP4 and WP5. In the former case, the conceptual and software-based tools developed in WP1 has been used to solve the classification problems posed in WP4. In the latter case, the EC-based classifiers have been the main subject of study in WP5, where the goal was to derive predictive models of classifier performance, to develop autonomous classification methods that can choose the best search dynamics for the particular problem instance that is addressed in a particular context. Finally, WP1 relates to WP5, since some of the software tools proposed and developed here are actually integrated into the software toolbox that is expected in WP7.

WP2: Acquisition and study of Bio-Signals

From the discovery of the brain electrical activity and the invention of the electroencephalography recording technique (Berger 1929), the oscillatory rhythms of EEG have been correlated to different vigilance states. Those rhythms, defined by their frequency band, are still often used in diagnostics of pathologies. The use of EEG signal is evolving and is more efficient thanks to the joint improvement of the acquisition methods and signal analysis methods. This WP is directly related to the acquisition methods of EEG and to the study of the analysis method of these signals.

First, this WP contributed to a better understanding of the main goal of this project which is depicted in WP4. Secondly, EEG data are hard to record. The recordings are very dependent of the neighborhood (electrical system, lights, etc...). The university of Bordeaux has an EEG recording system (64 electrodes) and was experienced with this kind of recording. The contribution of this WP was to teach to the other partners how to perform a correct recording and how to analyze these signals.

On the other hand, the partners helped the university to develop new analysis tools.

The main objective of this WP was to use the knowledge of an expert of this domain (F. Faïta) to teach to the other partners how the data are obtained and what the usual processing methods are. The goal was also to develop basic tools for EEG signal processing.

WP3: Analysis and feature extraction from EEG recordings of human states of vigilance

In this WP, the goal was to propose, develop and test several different types of feature extraction methods for EEG-based analysis and interpretation of human mental states of vigilance. The proposal was to take a diverse approach, incorporating different types of formal analysis, allowing us to compare results and evaluate the strengths and weaknesses of each method. Moreover, the proposal includes a heuristic approach, automatic feature synthesis through a GP-based search. It is assumed that no single feature can give a global solution to the problem of discriminately describing EEG signals, therefore hybrid features were proposed, composed of the above mentioned features combined in a principled manner using Fuzzy Logic. Finally, the quality of each proposed feature has been validated using a large scale experimental verification exploiting a distributed grid computing model.

The partners collaborating in this WP provided a multidisciplinary view to the problem of feature extraction. First, the problem has been analyzed from the perspective of the signal processing community, by considering wavelet decompositions and regularity based analysis with Hölder exponents. Second, the EEG signals has been analyzed from the perspective of chaos theory. Third, a heuristic approach to automatic feature synthesis has been developed, using EC and GP algorithms. Finally, several features have been combined in order to gain a more general description of EEG data using data fusion techniques from the computational intelligence community, particularly using EC and fuzzy logic.

WP4: Detection and Classification of human states of vigilance through EEG data

Almost all of the partners contributed in this WP, since most of the work developed in other WPs is used and tested here; including EEG recording methods, GP based classifiers and feature extractors.

The general problem addressed in this WP can be posed as follows. The goal was to automatically detect different mental states of vigilance from human test subjects using EEG recordings. It is currently accepted that if one analyses two EEG recording from the same individual it is possible to determine in which case the person was more “relaxed”. However, suppose that only a single recording is available. In this case, the problem becomes increasingly more difficult, since a previous general model would be required. However, most researchers agree that EEG signals are not homogenous across different individuals; i.e., each person can exhibit different signal patterns even when they state that they are experiencing a similar mental experience. Therefore, we were interested in developing new tools to address the problem of feature selection; feature construction and data classification to categorize (classify) the mental state of an (unknown) individual.

The proposed solution incorporates three convergent lines research. First, the EEG data obtained in WP2 is the central object of study. The EEG signals were analyzed using the tools proposed and developed in WP3, providing a set of descriptive features. Then, EEG signals were categorized and classified using the algorithmic tools developed in WP1. This WP also provided the real-world test

case for the analysis and descriptive tools of problem difficulty developed in WP5. Finally, the developed algorithms are being integrated into the software tools developed in WP7.

This WP solves the main goal of the proposal, the detection and classification of mental states of vigilance. In this sense, all other WPs are related with it. The classifiers used were developed in WP1; the subject of study, EEG data, was interpreted and pre-processed in WP2; descriptive features used to discriminate data and extract useful information were developed in WP3; finally, predictors of expected performance for the black-box methods developed in WP5, were validated with the experimental results of this WP.

WP5: Problem analysis and prediction of expected performance

Genetic Programming (GP) deals with the development of evolutionary algorithms (EA's) for automatic program induction [1], However, as for every EA, GP systems are stochastic search process, with many degrees of freedom and heuristic components. Therefore, as of yet, it is not possible to derive, from first principles, weather GP can solve a particular problem or task. A current goal within the GP community is to estimate how hard a problem instance might be for a specific GP. Such a measure could allow researchers to correctly choose and tune a GP search without the need of actually executing the code, which usually is computationally expensive. If we want to measure the difficulty of a problem in GP, we can consider at least two different frames of reference. The first is the problem domain, which is independent of the method used to solve the problem [3]. The second frame of reference is to consider a perspective directly related to the process used to find a solution; in the case of GP this frame of reference corresponds with the search space and fitness landscape.
In this WP, the goal was to use the problem domain as the frame of reference, and characterize the difficulty of a problem based on the expected performance the GP search, a quantity that is derived from domain specific features of each problem instance. This is a pragmatic approach, the evolutionary search is taken as a black-box process and the performance of GP on a set of training problems is used to build predictors of the expected performance on unseen problems, following a machine learning methodology. We refer to such measures of problem difficulty as Predictors of Expected Performance (PEPs).
Landscapes and problem difficulty has been the subject of a good deal of research regarding EA's. For instance, researchers have developed work on landscape correlation, autocorrelation, epistasis, monotonicity, locality and neutrality. However, the main work on this topic relates to evolvability indicators, such as fitness distance correlation and the negative slope coefficient, two empirical tools that measure the underlying difficulty of a given search process; however, such measures are not necessarily correlated with the difficulty of the problem or the expected performance a given solution strategy might achieve. Therefore, this WP followed a more recent line of research, where the goal is to directly predict the performance of an EC algorithm without actually performing the search.

WP6: Modelling Neurological Behaviors

The goal of this WP was to derive descriptive models of the object of study, the dynamics that govern the evolution of mental states of vigilance over a series of prescribed stimuli. Given the detailed analysis of how to characterize and discriminate among different mental states, a detailed description of the dynamics of state transitions induced by external stimuli can be of great relevance for future study.

In particular, the goal was to derive two types of descriptive models. First, computational models derived automatically, using heuristic methods, particularly GP, fuzzy logic and neural networks. Second, models based on differential equations describing the dynamics of mental states at a coarse level. In the former case, the goal was to replicate the behaviors exhibited by human beings, as their mental states develop over time within a given context, thus allowing us to replicate these features within an artificial brain. In the latter case, it is possible to localize the compact invariant sets of a system and deduce sufficient conditions under which the dynamics of the system can be controlled in a specified manner.

WP7: Tools integration in GPLAB

The goal of this WP was to assemble the tools developed in the other WPs into GPLAB. GPLAB is a Genetic Programming toolbox for MATLAB initially created and currently developed by Sara Silva. Most of its functions are used as "plug and play" devices, making it a versatile and easily extendable tool. The main objective of WP7 is to provide to the community the numerical products of our collaboration. A parallel objective is to extend and improve GPLAB with the new developed functionalities.

The GP tools developed in WPs 1, 3, 4 and 5, are merged in the toolbox GPLAB in such a way that (i) they use and are used by the other functionalities already implemented, following the same "plug and play" philosophy, and (ii) together they implement specialized functionalities for EEG classification, feature extraction and difficulty prediction. Despite the fact that WP7 is the concluding part of this project, the integration of the tools will be performed as they are developed, beginning in the second year of the project.

Some of the current features of GPLAB are: 3 modes of tree initialization (Full, Grow, Ramped Half-and-Half) + 3 variations on these, several pre-made functions and terminals for building trees, dynamic limits on tree depth or size (optional) , resource-limited GP (variable size populations) (optional) , dynamic populations (variable size populations) (optional) , 4 genetic operators (crossover, mutation, swap mutation, shrink mutation), configurable automatic adaptation of operator probabilities (optional) , steady-state + generational + batch modes, with fuzzy frontiers between them , 5 sampling methods (Roulette, SUS, Tournament, Lexicographic Parsimony Pressure Tournament, Double Tournament), 3 modes of calculating the expected number of offspring (absolute + 2 ranking methods), 2 methods for reading input files and for calculating fitness (symbolic regression and parity problems + artificial ant problems) , runtime cross-validation of the best individual of the run (optional) , offline cross-validation or prediction of results by any individual (optional) , 4 levels of elitism, configurable stop conditions, saving of results to files (5 frequency modes, optional) , 3 modes of runtime textual output, runtime graphical output (4 plots, optional) , offline graphical output (5 functions, optional) , runtime measurement of population diversity (2 measures, optional) , runtime measurement of average tree level, number of nodes, number of introns, tree fill rate (optional), 4 demonstration functions (symbolic regression, parity, artificial ant, multiplexer) .

Website of the GPLAB toolbox: http://www.gplab.sourceforge.net

1 - Qualitative indicators of progress and success in line with work plan and milestones (description of progress towards the milestones and deliverables)

Work packages and deliverables:

WP1: Classification with Evolutionary Computation

D1.1: Technical report on state of the art GP classification (M06)

Instead of building a technical report for this first deliverable, we decided to publish our first results in international journals, international conferences or book chapters; this has been possible due to the high quality work that has been developed, increasing the dissemination of our work to the research community at large. See **Deliverable1_1.pdf.**

Leonardo Trujillo and Luis Munoz and Enrique Naredo and Yuliana Martinez. NEAT, There's No Bloat. In M. Nicolau et al. editors, 17th European Conference on Genetic Programming, volume 8599, pages 174-185, Granada, Spain, 2014. Springer.
(http://rd.springer.com/chapter/10.1007%2F978-3-662-44303-3_15)

Yuliana Martinez and Leonardo Trujillo and Enrique Naredo and Pierrick Legrand. A Comparison of Fitness-Case Sampling Methods for Symbolic Regression with Genetic Programming. In Alexandru-Adrian Tantar et al. editors, EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation V, volume 288, pages 201-212, Pekin, 2014. Springer.
(http://rd.springer.com/chapter/10.1007%2F978-3-319-07494-8_14)

Emigdio Z-Flores and Leonardo Trujillo and Oliver Schuetze and Pierrick Legrand. Evaluating the Effects of Local Search in Genetic Programming. In Alexandru-Adrian Tantar et al. editors, EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation V, volume 288, pages 213-228, Pekin, 2014. Springer.
(http://rd.springer.com/chapter/10.1007%2F978-3-319-07494-8_15)

Vijay Ingalalli and Sara Silva and Mauro Castelli and Leonardo Vanneschi. A Multi-dimensional Genetic Programming Approach for Multi-class Classification Problems. In M. Nicolau and K. Krawiec and M. I.Heywood and M. Castelli and P. Garci-Sanchez and J.J. Merelo and V. M. R. Santos and K. Sim editors,17th European Conference on Genetic Programming,volume 8599, pages 48-60, Granada, Spain, 2014.Springer.
http://link.springer.com/chapter/10.1007%2F978-3-662-44303-3_5

Classification problems are of profound interest for the machine learning community as well as to an array of application fields. However, multi-class classification problems can be very complex, in particular when the number of classes is high. Although very successful in so many applications, GP was never regarded as a good method to perform multi-class classification. In this work, we present a novel algorithm for tree based GP, that incorporates some ideas on the representation of the solution space in higher dimensions. This idea lays some foundations on addressing multi-class classification problems using GP, which may lead to further research in this direction. We test the new approach on

a large set of benchmark problems from several different sources, and observe its competitiveness against the most successful state-of-the-art classifiers.

Luis Munoz and Sara Silva and Leonardo Trujillo. M3GP: Multiclass Classification with GP. In Penousal Machado and Malcolm I. Heywood and James McDermott and Mauro Castelli and Pablo Garcia-Sanchez and Paolo Burelli and Sebastian Risi and Kevin Sim editors, 18th European Conference on Genetic Programming, volume 9025, pages 78-91, Copenhagen, 2015. Springer.

Data classification is one of the most ubiquitous machine learning tasks in science and engineering. However, Genetic Programming is still not a popular classification methodology, partially due to its poor performance in multiclass problems. The recently proposed M2GP - Multidimensional Multiclass Genetic Programming algorithm achieved promising results in this area, by evolving mappings of the p-dimensional data into a d-dimensional space, and applying a minimum Mahalanobis distance classifier. Despite good performance, M2GP employs a greedy strategy to set the number of dimensions d for the transformed data, and fixes it at the start of the search, an approach that is prone to locally
optimal solutions. This work presents the M3GP algorithm, that stands for M2GP with multidimensional populations. M3GP extends M2GP by allowing the search process to progressively search for the optimal number of new dimensions d that maximize the classification accuracy. Experimental results show that M3GP can automatically determine a good value for d depending on the problem, and achieves excellent performance when compared to state-of-the-art-methods like Random Forests, Random Subspaces and Multilayer Perceptron on several benchmark and real-world problems.

Emigdio Z-Flores, Leonardo Trujillo, Oliver Schütze and Pierrick Legrand. 2014. A Local Search Approach to Genetic Programming for Binary Classification. Proceedings of the seventeenth annual conference on Genetic and evolutionary computation conference (GECCO '15), Ana Esparcia (Ed.). ACM.

In standard genetic programming (GP), a search is performed over a syntax space defined by the set of primitives, looking for the best expressions that minimize a cost function based on a training set. However, most GP systems lack a numerical optimization method to fine tune the implicit parameters of each candidate solution. Instead, GP relies on more exploratory search operators at the syntax level. This work proposes a memetic GP, tailored for binary classification problems. In the proposed method, each node in a GP tree is weighted by a real-valued parameter, which is then numerically optimized using a continuous transfer function and the Trust Region algorithm is used as a local search method. Experimental results show that potential classifiers produced by GP are improved by the local searcher, and hence the overall search is improved achieving significant performance gains, that are competitive with state-of-the-art methods on well-known benchmarks.

Mauro Castelli, Leonardo Trujillo, Leonardo Vanneschi, Sara Silva, Emigdio Z-Flores and Pierrick Legrand. 2014. Geometric Semantic Genetic Programming with Local Search. Proceedings of the seventeenth annual conference on Genetic and evolutionary computation conference (GECCO '15), Ana Esparcia (Ed.). ACM.

D1.2: Seminar on statistical classification methods (M08)

Through the mobility program, we have carried out several seminars between January and June 2014, at the University of Bordeaux, INRIA, University of Extremadura, and Lisbon with our European partners, and at ITT and CICESE on the Mexican side. Participants include: Laurent Vezard, Frederique Faïta Ainseba, Marie Chavent, Yuliana Martinez, Enrique Naredo, Emigdio Z. Flores, Pierrick Legrand, Sara Silva; Francisco Fernandex, Francisco Chavez, Leonardo Trujillo, Daniel Lanza, Gustavo Olague, Eddie Clemente, León Dozal, Daniel Hernández, Victor López, Arturo Sotelo.

As part of the work package a Special Session was held at the EVOLVE 2014 conference, held on July 1 – 4th 2014 in Beijing China; see:
http://www.evolve-conference.org/2011-03-31-04-23-33

D.1.3 Workshop on GP-based classification

The project team members, especially Leonardo Trujillo and Pierrick Legrand, co-organized the Workshop on Numerical and Evolutionary Computation NEO 2015 held from the 23rd to the 25th of September 2015 (see http://neo.cinvestav.mx/NEO2015/) in collaboration with CINVESTAV-IPN, one of the leading research centers in Mexico. At this event, one of the main tracks focused on Genetic Programming, and included the participation of some of the leading researchers in the field, including Dr. Leonardo Vanneschi (winner of the 2015 Award for Outstanding Contributions in Evolutionary Computation) and Dr. Mario Graff. Moreover, all of the student participants were key participants at the event, giving several talks and working as organizers. From this event, a post-proceedings book has being edited, along with collaborators Dr. Oliver Schütze and Dr. Yazmin Maldonado (see http://neo.cinvestav.mx/NEO2015/index.php/post-proceedings), has been published in September 2016 by Springer in the Studies in Computational Intelligence series.

NEO 2015, Results of the Numerical and Evolutionary Optimization Workshop NEO 2015 held at September 23-25 2015 in Tijuana, Mexico. Schütze, O., Trujillo, L., Legrand, P., Maldonado, Y. (Eds.), 17 chapters, series Studies in Computational Intelligence, Springer, 444 pages (2017).

Additionally, we mention the recent theoretical book chapter from our partner from Portugal, related to this WP:

Semantic Learning Machine: A Feedforward Neural Network Construction Algorithm Inspired by Geometric Semantic Genetic Programming, Ivo Gonçalves, Sara Silva, Carlos M. Fonseca, in Progress in Artificial Intelligence: 17th Portuguese Conference on Artificial Intelligence, EPIA 2015, Coimbra, Portugal, September 8-11, 2015. Proceedings. Springer, 811 pages.

Finally, during the last incoming mobility to the university of Bordeaux, Uriel Lopez Islas developed new technics related to this work package. See documents **Report_Uriel.pdf** and **Presentation_Uriel.pdf**.

The main objective of this mobility was to create a Genetic Programming (GP) variant that is robust against the presence of outliers in the training data set for symbolic regression problems.

The first question is, what needs to be done to GP for it to be robust against outliers? with the answer being: first thing is to change the traditional fitness function that GP uses to assess fitness, on individuals, for two new fitness function. The second question would be, why change the fitness function? The reason for this is because the traditional fitness function GP uses to assess fitness is quite sensitive to outliers. Meaning that it has a low breakdown point. To make things clearer, a breakdown point is reached when the model obtained by GP, in this case, does not represent the real model. The new fitness functions are taken from literature on robust linear regression methods and have not been used on any GP system till date, so it is interesting to see what those new fitness function implementations do for GP.

Also, because of the nature that this work addresses sampling methods will be deployed. Sampling methods have shown success on working with only parts of the training data set and not the whole training data set. With this in mind they may also work on finding good models with the presence of outliers. The experimental part is done with benchmark problems. The purpose of this is that the real model of the benchmark function is known and later can be compared with the model that GP finds using symbolic regression. Each benchmark problem will be contaminated artificially with outliers and afterwards, try to obtain the real model with symbolic regression with the presence of outliers. But one must not confuse outliers with noise. So we must contaminate them in an adequate manner. For our problems we are assuming that our benchmark problems are changing in time. So it is needed a proper contamination method to simulate time-series outliers. For this task the Hampel identifier and the Hampel filter has been used.

WP2: Acquisition and study of Bio-Signals

F. Faïta gave a seminar to 3 Mexican PhD students (E. Naredo, Y. Martinez, E. Flores), regarding the state of the art in BCI signal measurements with EEG, covering aspects related to hardware and experimental setups. Moreover, the students participated in the setup and acquisition of new experimental data. Four French master students were also involved in the experimental sessions that were done at the University of Bordeaux. (See attached document **ACOBSEC_WP2_ITT_DOC_EEGREC_R1.0.pdf**).

F. Faïta and P. Legrand gave seminars and teaching session to the student focusing on aspects related to pre-processing of EEG data.

New software for VCN analysis was developed by P. Legrand. He received the support of two undergraduate students (Charlotte Rodriguez and Borjan Geshkovski in June 2015. (**See attached documents Software_EEG_analysis.pdf** and **RapportStage_ExpeVigilance_2015.pdf**).

E. Flores participated to the winter school on BCI (See attached document **ACOBSEC_WP2_ITT_DOC_BCIWS_R1.0.pdf**). And he organized a seminary about advanced correlation Filters (see documents **ACOBSEC_WP2_ITT_DOC_CORRFILT_R1.0.pdf** and **correlation_presentation.pdf**).

Two new rooms for experimental work, and devoted primarily to the project, were built inside the University of Bordeaux.

The following actions are relevant and are part of this project:

- Luis Herrera, during his mobility in France (sept-nov 2015) **developed a new software for the analysis of the biological signals** (see **Interf_Luis.png**). He received courses on mathematics (P. Legrand), signal processing (P. Legrand, E. Grivel) but also on bio-data acquisition (F. Faïta).

- During the last months we started a **new campaign of data Acquisition**. We decided to synchronize the acquisition of simultaneous data (see **Synchro.png**). We successfully recorded EEG, ECG (cardio), EYE (eye tracking), EDA (electro dermal activity) simultaneously.

D2.1 Training report

Please find in the document **list_training.pdf** the list of the Mexican member of the project who had the opportunity to be initiated to the recording methods of EEG signals under the supervision of Frédérique Faïta at the University of Bordeaux, France. The French internships are not mentioned in this document.

WP 3: Analysis and feature extraction from EEG recordings of human states of vigilance

D3.1: Technical report on feature extraction methods for EEG (M6)

Instead of building a technical report, we decided to publish our first results in international journals, international conferences or book chapters (**See Deliverable3_1.pdf)**. These can be found in the following references:

Pierrick Legrand, Laurent Vezard, Marie Chavent, Frédérique Faïta-Ainseba, Leonardo Trujillo, Feature extraction and classification of EEG signals. The use of a genetic algorithm for an application on alertness prediction. In "Guide to Brain-Computer Music Interfacing", Eduardo Miranda, Julien Castet and Benjamin Knapp Eds., Springer.
http://www.springer.com/computer/hci/book/978-1-4471-6583-5

Laurent Vézard, Pierrick Legrand, Marie Chavent, Frédérique Faïta-Ainseba, Julien Clauzel, and Leonardo Trujillo. Classification of EEG signals by evolutionary algorithm. In "Advances in Knowledge Discovery and Management Vol.4", Studies in Computational Intelligence. Volume 527, pp133-153.
http://www.springer.com/engineering/computational+intelligence+and+complexity/book/978-3-319-02998-6

Laurent Vezard, Pierrick Legrand, Marie Chavent, Frédérique Faïta-Aïnseba, Leonardo Trujillo. Detecting mental states of alertness with genetic algorithm variable selection, Applied Soft Computing (2015).

The objective of the present work is to develop a method that is able to automatically determine mental states of vigilance; i.e., a person's state of alertness. Such a task is relevant to diverse domains, where a person is expected or required to be in a particular state of mind. For instance, pilots and medical staff are expected to be in a highly alert state and the proposed method could help to detect possible deviations from this expected state. This work poses a binary classification problem where the goal is to distinguish between a "relaxed" state and a baseline state ("normal") from the

study of electroencephalographic signals (EEG) collected with a small number of electrodes. The EEG of 58 subjects in the two alertness states (116 records) were collected via a cap with 58 electrodes. After a data validation step, 19 subjects were retained for further analysis. A genetic algorithm was used to select a subset of electrodes. Common spatial pattern (CSP) coupled to linear discriminant analysis (LDA) was used to build a decision rule and thus predict the alertness of the subjects. Different subset sizes were investigated and the best compromise between the number of selected electrodes and the quality of the solution was obtained by considering 9 electrodes. Even if the present approach is costly in computation time (GA search), it allows to construct a decision rule that provides an accurate and fast prediction of the alertness state of an unseen individual.

Arturo Sotelo, Enrique D. Guijarro and Leonardo Trujillo. Seizure States Identification in Experimental Epilepsy using Gabor Atom Analysis, Journal of Neuroscience Methods (2015).

Moreover, internal project reports and slides were written by Emigdio Flores, Yuliana Martinez and Enrique Naredo, PhD students from Mexico, during their stay at INRIA and the University of Bordeaux.

D.3.2 Seminar on feature extraction methods for EEG

Emigdio Z-Flores, has given a seminar on this topic at ITT over the last year, for collaborating researchers and especially for the team of graduate students working on the project. The main contents of these talks have been recently accepted for publication in:

- Emigdio Z-Flores, Leonardo Trujillo**,** Arturo Sotelo, Pierrick Legrand and Luis Coria. Regularity and Matching Pursuit Feature Extraction for the Detection of Epileptic Seizures, accepted to Journal of Neuroscience Methods (2016).

In this paper, our main feature extraction methods have been tested on a well-kown EEG dataset, achieving state-of-the-art results. Our feature sets include highly-nonlinear methods based on multi-fractal regularity analysis and the matching pursuit algorithm.

Moreover, Dr. Leonardo Trujillo gave a seminar on this topic at CITEDI-IPN, one of the main research centers in Tijuana Mexico, to a crowd of graduate students and research staff.

Additional work regarding WP3:

An additional French internship started on this topic the 1st of July. Vincent Lenhardt, engineer student in Bordeaux is actually working on the new data recorded (and mentioned in WP2) under the supervision of Eric Grivel and Pierrick Legrand.

Description of the internship topic (in French):
Mots Clefs : traitement du signal biomédical, EEG, ECG, réponse electrodermale (EDA, GSR), pupillométrie, mesure, caractérisation, modélisation, estimation, divergence Kullback Leibler, divergence de Jeffrey, analyse tensorielle, analyse en composantes principales, Matlab.

Des données sont recueillies en parallèle à des fréquences d'échantillonnage différentes et sont de plusieurs types : électroencéphalogramme (EEG), électrocardiogramme (ECG), pupillométrie et réponse galvanique de la peau (galvanic skin response -GSR-).
L'objet du stage est le suivant :
- La première étape consiste à mener une étude bibliographique sur le sujet.

- Parallèlement, il s'agit d'effectuer des pré-traitements sur les signaux reçus de manière à les synchroniser, à disposer des signaux observés aux mêmes instants, etc. Ensuite, on analysera quelle est la combinaison de signaux qui est la plus corrélée aux EEG. On adoptera pour cela des approches de type corrélation.
- Dans un second temps, il s'agit de caractériser ces signaux à partir de modèles appropriés comme les modèles AR multivariés ou MA multivariés ; les propriétés de ces modèles ainsi que les techniques d'estimation des paramètres associés seront tout d'abord étudiées. Ces paramètres estimés pour chaque modèle dans un état cognitif donné seront aussi exploités pour développer une technique de classification ; On analysera ainsi des approches fondées sur la divergence de Jeffrey ou la distance de Rao.
- La méthode développée sera testée sur la base de données et une interface Matlab sera créée.
- Le travail effectué donnera lieu à la rédaction d'un document de synthèse, technique, qui peut être lu par n'importe quel utilisateur.

In October 2016, an additional pool of 9 engineer students was integrated in this work under the supervision of P. Legrand and E. Grivel. See document **Extraction_biomed_eric_grivel.pdf.**

WP4: Detection and classification of human states of vigilance through EEG data

No deliverable was scheduled before M24 for this work package, but before the delivery of the D4.1 our first results were published in the following publications (also related to WP3).

Pierrick Legrand, Laurent Vezard, Marie Chavent, Frédérique Faïta-Ainseba, Leonardo Trujillo, Feature extraction and classification of EEG signals. The use of a genetic algorithm for an application on alertness prediction. In "Guide to Brain-Computer Music Interfacing", Eduardo Miranda, Julien Castet and Benjamin Knapp Eds., Springer.
http://www.springer.com/computer/hci/book/978-1-4471-6583-5

This chapter presents a method to automatically determine the alertness state of humans. Such a task is relevant in diverse domains, where a person is expected or required to be in a particular state of alertness. For instance, pilots, security personnel or medical personnel are expected to be in a highly alert state, and this method could help to confirm this or detect possible problems. In this work, electroencephalographic (EEG) data from 58 subjects in two distinct vigilance states (state of high and low alertness) was collected via a cap with 58 electrodes. Thus, a binary classification problem is considered. To apply the proposed approach in a real-world scenario, it is necessary to build a prediction method that requires only a small number of sensors (electrodes), minimizing the total cost and maintenance of the system while also reducing the time required to properly setup the EEG cap.
The approach presented in this chapter applies a pre-processing method for EEG signals based on the use of a discrete wavelet decomposition to extract the energy of each frequency in the signal. Then, a linear regression is performed on the energies of some of these frequencies and the slope of this regression is retained. A genetic algorithm (GA) is used to optimize the selection of frequencies on which the regression is performed and to select the best recording electrode. Results show that the proposed strategy derives accurate predictive models of alertness.

Laurent Vézard, Pierrick Legrand, Marie Chavent, Frédérique Faïta-Ainseba, Julien Clauzel, and Leonardo Trujillo. Classification of EEG signals by evolutionary algorithm. In "Advances in Knowledge Discovery and Management Vol.4", Studies in Computational Intelligence. Volume 527, pp133-153.

http://www.springer.com/engineering/computational+intelligence+and+complexity/book/978-3-319-02998-6

Laurent Vezard, Pierrick Legrand, Marie Chavent, Frédérique Faïta-Aïnseba, Leonardo Trujillo. Detecting mental states of alertness with genetic algorithm variable selection, Applied Soft Computing (2015).

D 4.1 – D 4.2 Classification of EEG signals

The proposed set of EEG features (discussed in D.3.2), have been successfully applied with impressive results, meriting a recent publication in one of the main journals that cover Neuroscience research and applications of signal processing to related automatic tasks, particularly EEG analysis and classification. The summary of these results are given in the PhD of Laurent Vezard (see document **VEZARD_LAURENT.pdf**) and in the following journal paper.

Emigdio Z-Flores, Leonardo Trujillo**,** Arturo Sotelo, Pierrick Legrand and Luis Coria. Regularity and Matching Pursuit Feature Extraction for the Detection of Epileptic Seizures, Journal of Neuroscience Methods, Volume 266, 15 June 2016, Pages 107–125.
http://dx.doi.org/10.1016/j.jneumeth.2016.03.024

Background

The neurological disorder known as epilepsy is characterized by involuntary recurrent seizures that diminish a patient's quality of life. Automatic seizure detection can help improve a patient's interaction with her/his environment, and while many approaches have been proposed the problem is still not trivially solved.

Methods

In this work, we present a novel methodology for feature extraction on EEG signals that allows us to perform a highly accurate classification of epileptic states. Specifically, Hölderian regularity and the Matching Pursuit algorithm are used as the main feature extraction techniques, and are combined with basic statistical features to construct the final feature sets. These sets are then delivered to a Random Forests classification algorithm to differentiate between epileptic and non-epileptic readings.

Results

Several versions of the basic problem are tested and statistically validated producing perfect accuracy in most problems and 97.6% accuracy on the most difficult case. Comparison with existing methods: A comparison with recent literature, using a well-known database, reveals that our proposal achieves state-of-the-art performance.

Conclusions

The experimental results show that epileptic states can be accurately detected by combining features extracted through regularity analysis, the Matching Pursuit algorithm and simple time-domain

statistical analysis. Therefore, the proposed method should be considered as a promising approach for automatic EEG analysis.

Keywords

- *Hölderian regularity;*
- *Matching Pursuit;*
- *EEG classification*

See **D4142_DETECTION_CLASSIF_JournalOfNeurosciencesMethods_Flores.pdf**.

D 4.3 Performance evaluation of each method and comparative analysis
An evaluation and a comparison of the methods has been realized by Emigdio Z. Flores. Those analyses are part of his PhD thesis (expected defense in June 2017). For more information, please refer to the document **D43_FULL.pdf**.

WP5: Problem analysis and prediction of expected performance

D5.1: Seminar on Problem difficulty in GP (M06)

Seminars/workshops were organized in France (INRIA and University of Bordeaux) and in Spain (University of Extremadura). The main participants in these events were: Yuliana Martinez (ITT), Francisco Chavez (UEX), Francisco Fernandez (UEX), Enrique Naredo (ITT), Emigdio Z. Flores (ITT), Pierrick Legrand (INRIA-UB).

D5.2: Seminar on distributed computing exploiting BOINC and virtualization (M10)

Seminars/workshops were organized in France (INRIA and University of Bordeaux) and in Spain (University of Extremadura). The main participants in these events were: Yuliana Martinez (ITT), Francesco Chavez (UEX), Francesco Fernandez (UEX), Enrique Naredo (ITT), Emigdio Z. Flores (ITT), Luis Muñoz (ITT), Leonardo Trujillo (ITT), Enrique Mediero (UEX), Daniel Lanza (UEX).

Francisco Fernandez and Daniel Lanza gave teaching session to Mexican researchers on tools for big data processing, analysis and management. Based on these workshops, the following papers were submitted and published in relevant scientific forums.

Francisco Fernández de Vega, Leonardo Trujillo, Francisco Chávez, Enrique Mediero and Luis Muñoz. 2014. A Hybrid ECJ+BOINC Tool for Distributed Evolutionary Algorithms, Research in Computing Science, Advances in Computing Science, Control and Communications, Mireya Garcia et al. (Eds.), 69, 120-130.

Mario García-Valdez, Leonardo Trujillo, Juan Julián Merelo-Guérvos, Francisco Fernández-de-Vega. Randomized Parameter Settings for Heterogeneous Workers in a Pool-Based Evolutionary Algorithm, in Proceedings of Parallel Problem Solving from Nature – PPSN XIII, LCNS Vol. 8672, 2014, pp 702-710.

Mario García-Valdez, Leonardo Trujillo, Juan-J Merelo, Francisco Fernández de Vega & Gustavo Olague. The EvoSpace Model for Pool-Based Evolutionary Algorithms, Journal of Grid Computing (2015), September 2015, Volume 13, Issue 3, pp 329–349.

This work presents the EvoSpace model for the development of pool-based evolutionary algorithms (Pool-EA). Conceptually, the EvoSpace model is built around a central repository or population store, incorporating some of the principles of the tuple-space model and adding additional features to tackle some of the issues associated with Pool-EAs; such as, work redundancy, starvation of the population pool, unreliability of connected clients or workers, and a large parameter space. The model is intended as a platform to develop search algorithms that take an opportunistic approach to computing, allowing the exploitation of freely available services over the Internet or volunteer computing resources within a local network. A comprehensive analysis of the model at both the conceptual and implementation levels is provided, evaluating performance based on efficiency, optima found and speedup, while providing a comparison with a standard EA and an island-based model. The issues of lost connections and system parametrization are studied and validated experimentally with encouraging results, that suggest how EvoSpace can be used to develop and implement different Pool-EAs for search and optimization.

D 5.3 Workshop on Problem Difficulty in GP

As part of the NEO 2015 Workshop (discussed in D 1.3), one of the main talks given at the event was by doctoral student Yuliana Martínez, which discussed problem difficulty in GP and methods for performance prediction developed by the partners and collaborators of this project. Moreover, **the new results have been recently accepted for publication** in the flagship journal of Genetic Programming, see:

Yuliana Martínez, Leonardo Trujillo, Pierrick Legrand and Edgar Galván-López. Prediction of Expected Performance for a Genetic Programming Classifier, in Genetic Programming and Evolvable Machines, December 2016, Volume 17, Issue 4, pp 409–449. doi:10.1007/s10710-016-9265-9.

The estimation of problem difficulty is an open issue in genetic programming (GP). The goal of this work is to generate models that predict the expected performance of a GP-based classifier when it is applied to an unseen task. Classification problems are described using domain-specific features, some of which are proposed in this work, and these features are given as input to the predictive models. These models are referred to as predictors of expected performance. We extend this approach by using an ensemble of specialized predictors (SPEP), dividing classification problems into groups and choosing the corresponding SPEP. The proposed predictors are trained using 2D synthetic classification problems with balanced datasets. The models are then used to predict the performance of the GP classifier on unseen real-world datasets that are multidimensional and imbalanced. This work is the first to provide a performance prediction of a GP system on test data, while previous works focused on predicting training performance. Accurate predictive models are generated by posing a symbolic regression task and solving it with GP. These results are achieved by using highly descriptive features and including a dimensionality reduction stage that simplifies the learning and testing process. The proposed approach could be extended to other classification algorithms and used as the basis of an expert system for algorithm selection.

Keywords: *Problem difficulty, Prediction of expected performance, Genetic programming, Supervised learning.*

Further dissemination of these important results have been made by submitting a proposal to present at the GECCO 2016 (the main conference on evolutionary computation) in the Hot-of-the Press Track (http://gecco-2016.sigevo.org/index.html/Hot+off+the+Press), unfortunately it was not accepted due to the fact that a preliminary version of the work was presented at the conference recently.

WP6: Modelling neurobiological behavior

This work package was modified substantially relative to the original proposal, mainly regarding the date some of the deliverables were completed and regarding the focus of the original proposal. This was mainly due to the fact that students that were originally scheduled to participate could not do so as planned. These issues are clarified further when describing each deliverable of the work package below.

D 6.1 Mental state models derived with GP M30

In this deliverable, the results were completed well before the estimated date, and the results presented in the following publication.

Synthesis of odor tracking algorithms with genetic programming, B. Lorena Villarreal, Gustavo Olague, J.L. Gordillo, in Neurocomputing, Volume 175, Part B, 29 January 2016, Pages 1019–1032, Elsevier.

In that work, the partners were able to model the odor sensation using genetic programming, and derive a realistic sensor of a complex mental process as described in the following:

At the moment, smell sensors for odor source localization in mobile robotics represent a topic of interest for researchers around the world. In particular, we introduce in this paper the idea of developing biologically inspired sniffing robots in combination with bioinspired techniques such as evolutionary computing. The aim is to approach the problem of creating an artificial nose that can be incorporated into a real working system, while considering the environmental model and odor behavior, the perception system, and algorithm for tracking the odor plume. Current algorithms try to emulate animal behavior in an attempt to replicate their capability to follow odors. Nevertheless, odor perception systems are still in their infancy and far from their biological counterpart. This paper presents a proposal in which a real-working artificial nose is tested as a perception system within a mobile robot. Genetic programming is used as the learning technique platform to develop odor source localization algorithms. Experiments in simulation and with an actual working robot are presented and the results compared with two algorithms. The quality of results demonstrates that genetic programming is able to recreate chemotaxis behavior by considering mathematical models for odor propagation and perception system.

D 6.2 Fuzzy models of mental state dynamics M30

This deliverable was modified, since the fuzzy approach proved not to be relevant based on our tests and results. For this reason, the decision was taken to focus on the connectionist model studied in the next deliverable.

D 6.3 Neural models of mental state dynamics M30

This deliverable was also completed before the original date planned in the project. To model the mental process using a connectionist approach (neural network based), we followed a newly proposed approach by the partners entitled Brain Programming, which is based on a hybridization of Genetic Programming and a model of the visual cortex of the brain. This approach allowed the partners to develop a new technique to model and reproduce complex mental process, namely saliency detection and object detection. The main results were presented in:

León Dozal, Gustavo Olague, Eddie Clemente & Daniel E. Hernández Brain Programming for the Evolution of an Artificial Dorsal Stream Cognitive Computation ISSN 1866-9956. DOI 10.1007/s12559-014-9251-6

The main results can be summarized as follows:

This work describes the application of a new strategy called brain programming for automating the design of visual attention (VA) models. Nowadays, a term known as cognitive vision coined within the computer vision and cognitive research communities has been introduced to delimitate the kind of computer vision systems that are robust, resilient and adaptable to the task at hand through the incorporation of cognitive abilities. In particular, visual attention is considered as a critical factor whose main goal is to establish a relationship between the different properties or features of the scene with the aim of selecting the most suitable aspects for the task at hand. This paper follows a main trend in cognitive computation where the visual pathway is modeled through a succession of levels or layers. Here, the VA task is defined with the idea that several areas of the brain are in charge of its functionality in a hierarchical way. To achieve such functionality, we propose that an artificial process, mimicking the natural counterpart, would be charged of looking for a set of complex operations using an optimization/search process. The idea is to include such operations within a VA model that will be evolved according to a specific task. The aim of the whole process is to provide with the best solutions among the space of possible visual attention programs (VAPs) for a given problem. In this way, the article presents a methodology for automating the design of VAPs. Therefore, the final design can be seen as a cognitive vision system that is engaged in a purposive goal-directed behavior. The results obtained on a well-known testbed confirm that the proposal is able to automatically design VAPs that outperform previous man-made systems developed by VA experts, while providing readable results through a set of mathematical and computational structures.

Finally, the approach was extended to consider multiple modeling criteria simultaneously, and to solve the more complex task of object recognition, as presented in:

Evolutionary multi-objective visual cortex for object classification in natural images, Daniel E. Hernandez, Eddie Clemente, Gustavo Olague, in Journal of Computational Science, Available online 2 November 2015, Elsevier.

The main results of that paper can be summarized as:

In recent years computer vision systems have used the human visual system as inspiration for solving different tasks such as object detection and classification. Computational models as the artificial

visual cortex (AVC) have shown promising results in solving such problems. Thus, this paper proposes a new methodology for creating an image descriptor vector for classification, and at the same time, finding the objects' location within the image. Also, this work implements the brain programming paradigm from a multi-objective perspective in order to improve the performance in the object classification task. This methodology is implemented for training the proposed model in order to classify the images from the GRAZ-01 and GRAZ-02 databases. The solutions found in this research match, and in some cases outperform, other techniques of the state-of-the-art for classifying the aforementioned databases.

D 6.4 Differential Equation model of mental state dynamics M32

Regarding this deliverable (and the next), which were meant to be developed by partner ITT, were finally excluded from the project. The main reason was that it was decided that the participants from ITT should focus on the issues regarding data analysis, EEG analysis, and algorithm development of GP-based systems, since these topics proved to be the most promising of all the work. In fact, this can be validated by the amount of publications, many of the in high quality journals, that were produced by this collaboration on these topics. For this reasons, these deliverables are left as future work derived from this project.

D 6.5 Localization of compact invariant sets M36

D 6.6 Workshop Organization M15

This workshop was organized as a special session of the "X Congreso Español de Metaheurísticas, Algoritmos Evolutivos y Bioinspirados - MAEB 2015", which was held from the 4th to the 6th of February at the Centro Universitario de Merida from the Univerisdad de Extremadura in Spain. The special session was entitled "Análisis y reconocimiento de patrones basado en modelos y algoritmos bio-inspirados", with the main organizers of the session being Gustavo Olague and Leonardo Trujillo, participants from the partners CICESE and ITT. The session had the participation of 5 oral presentation papers, covering several topics related to the main lines of research of this work package. Full details of the event can be found at:

MAEB 2015 website: http://www.eweb.unex.es/eweb/maeb2015/?Presentaci%C3%B3n

Program of the event: http://www.eweb.unex.es/eweb/maeb2015/userfiles/downloads/Programa%20MAEB%202015_Definitivo.pdf

WP7: Tools integration in GPLAB

During the last year our Spanish partners have focused on publishing in JCR journals the main results of the project developed in collaboration with the other partners. They have also tested some of the developments achieved in the previous years, particularly the use of Map/Reduce methodology to run evolutionary algorithms making use of big amount of data, in specific applications of interest: recommender systems. They have developed a Fuzzy rule based system capable of learning the rules by evolutionary means when processing big amount of data belonging to customers of a specific

application. Similarly, a problem related to sound analysis has been faced using a similar approach. Both applications rely on a proper use of massive versions of distributed evolutionary algorithms. We hope that these new developments will participate in the improvements of the methods already used by the researchers working with big data sets.

The goal of this WP was to assemble the tools developed in the other WPs into GPLAB. GPLAB is a Genetic Programming toolbox for MATLAB initially created and currently developed by Sara Silva. Most of its functions are used as "plug and play" devices, making it a versatile and easily extendable tool. The main objective of WP7 is to provide to the community the numerical products of our collaboration. A parallel objective is to extend and improve GPLAB with the new developed functionalities.

The GP tools developed in WPs 1, 3, 4 and 5, are merged in the toolbox GPLAB in such a way that (i) they use and are used by the other functionalities already implemented, following the same "plug and play" philosophy, and (ii) together they implement specialized functionalities for EEG classification, feature extraction and difficulty prediction. Despite the fact that WP7 is the concluding part of this project, the integration of the tools will be performed as they are developed, beginning in the second year of the project.

Some of the current features of GPLAB are: 3 modes of tree initialization (Full, Grow, Ramped Half-and-Half) + 3 variations on these, several pre-made functions and terminals for building trees, dynamic limits on tree depth or size (optional) , resource-limited GP (variable size populations) (optional), dynamic populations (variable size populations) (optional) , 4 genetic operators (crossover, mutation, swap mutation, shrink mutation), configurable automatic adaptation of operator probabilities (optional) , steady-state + generational + batch modes, with fuzzy frontiers between them , 5 sampling methods (Roulette, SUS, Tournament, Lexicographic Parsimony Pressure Tournament, Double Tournament), 3 modes of calculating the expected number of offspring (absolute + 2 ranking methods), 2 methods for reading input files and for calculating fitness (symbolic regression and parity problems + artificial ant problems) , runtime cross-validation of the best individual of the run (optional) , offline cross-validation or prediction of results by any individual (optional) , 4 levels of elitism, configurable stop conditions, saving of results to files (5 frequency modes, optional) , 3 modes of runtime textual output, runtime graphical output (4 plots, optional) , offline graphical output (5 functions, optional) , runtime measurement of population diversity (2 measures, optional) , runtime measurement of average tree level, number of nodes, number of introns, tree fill rate (optional), 4 demonstration functions (symbolic regression, parity, artificial ant, multiplexer) .

Website of the GPLAB toolbox: http://www.gplab.sourceforge.net

As the project developed, WP7 was identified as the most promising for future collaborations, for several reasons. First, the tools developed were implemented and extensively tested in GPLAB, with work on bloat-free-GP (1), local search (2, 3, 7), novelty search (4), performance prediction (5), and fitness case sampling methods (6), with the main results presented in the following publications:

1. Leonardo Trujillo, Luis Muñoz, Edgar Galván-López, Sara Silva. neat Genetic Programming: Controlling Bloat Naturally, Information Sciences 333:21-43 (2016).
2. Emigdio Z-Flores, Leonardo Trujillo, Oliver Schütze, and Pierrick Legrand. 2015. A Local Search Approach to Genetic Programming for Binary Classification. In Proceedings of the 2015 on

Genetic and Evolutionary Computation Conference (GECCO '15), Sara Silva (Ed.). ACM, New York, NY, USA, 1151-1158.
3. Mauro Castelli, Leonardo Trujillo, Leonardo Vanneschi, Sara Silva, Emigdio Z-Flores, and Pierrick Legrand. 2015. Geometric Semantic Genetic Programming with Local Search. In Proceedings of the 2015 on Genetic and Evolutionary Computation Conference (GECCO '15), Sara Silva (Ed.). ACM, New York, NY, USA, 999-1006.
4. Enrique Naredo, Leonardo Trujillo, Pierrick Legrand, Sara Silva, and Luis Muñoz. 2016. Evolving genetic programming classifiers with novelty search. *Inf. Sci.* 369, C (November 2016), 347-367.
5. Yuliana Martínez, Leonardo Trujillo, Pierrick Legrand and Edgar Galván-López. Prediction of Expected Performance for a Genetic Programming Classifier, Genet Program Evolvable Mach (2016) 17: 409.
6. Yuliana Martínez, Enrique Naredo, Leonardo Trujillo, Pierrick Legrand and Uriel López. A comparison of fitness-case sampling methods for genetic programming, to appear in the Journal of Experimental & Theoretical Artificial Intelligence.
7. Leonardo Trujillo, Emigdio Z-Flores, Perla S. Juárez Smith, Pierrick Legrand, Sara Silva, Mauro Castelli, Leonardo Vanneschi, Oliver Schütze and Luis Muñoz. Local Search is Underused in Genetic Programming. In Rick Riolo et al. editors, Genetic Programming Theory and Practice XIV, Ann Arbor, USA, 2017. Springer.

While most of these results are relevant to other WPs, they were fully implemented and integrated into the GP-LAB toolbox, particularly Version 3 (Enrique Naredo > Novelty Search, Yuliana Martinez > Prediction of performance, Emigdio Z. Flores > Local search…).

The second reason why this WP is considered to be important for future work is because much of the developed tools have proven to be useful in other domains of pattern recognition and data science. This has allowed the partners to quickly propose new and interesting solutions to difficult real-world problems that were not the focus of the original work, but which quickly became apparent that our developed tools could solve, including modelling electronic circuits, predicting atmospheric properties, medical imaging, and modelling mechanical systems; see for instance:

- J.R. Cárdenas Valdez, Emigdio Z. Flores, José Cruz Nuñez Pérez and Leonardo Trujillo, Local Search Approach to Genetic Programming for RF-PAs Modeling Implemented in FPGA, in NEO 2015, Results of the Numerical and Evolutionary Optimization Workshop NEO 2015 held at September 23-25 2015 in Tijuana, Mexico. Schütze, O., Trujillo, L., Legrand, P., Maldonado, Y. (Eds.), series Studies in Computational Intelligence, Springer, pp. 67-88 (2017).
- Mauro Castelli, Ivo Goncalves, Leonardo Trujillo and Ales Popovic. An Evolutionary System for Ozone Concentration Forecasting, to appear in Information Systems Frontiers.
- Mauro Castelli, Leonardo Trujillo, Leonardo Vanneschi, and Aleš Popovič. 2016. Prediction of relative position of CT slices using a computational intelligence system. Appl. Soft Comput. 46, C (September 2016), 537-542.
- José Enrique Hernández-Beltrán; Víctor H. Díaz-Ramírez; Leonardo Trujillo and Pierrick Legrand. Restoration of degraded images using genetic programming, Proc. SPIE 9970, Optics and Photonics for Information Processing X, 99700K (September 14, 2016)
- Josúe Zárate, Leonardo Trujillo, Salvador de Lara, Mauro Castelli, Emigdio Z-Flores, Luis Muñoz, Ales Popovic. Automatic Modeling of a Gas Turbine using Genetic Programming: An Experimental Study, to appear in Applied Soft Computing

Thirdly, we have gone beyond the use of a single software tool, and have applied our approaches to other platforms and with other computing paradigms, including cloud computing, Map Reduce, ECJ

(Java) and DEAP (Python), and GSGP (C++). Therefore, future collaborations will focus on developing hybrid systems, based on the strengths and weaknesses of each computing platform or software system. Results on these topics were reported in the following works:

- Daniel Lanza, F. Chávez, Francisco Fernández, M. Garcia-Valdez,Leonardo Trujillo and Gustavo Olague, Profiting from Several Recommendation Algorithms Using a Scalable Approach, in NEO 2015, Results of the Numerical and Evolutionary Optimization Workshop NEO 2015 held at September 23-25 2015 in Tijuana, Mexico. Schütze, O., Trujillo, L., Legrand, P., Maldonado, Y. (Eds.), series Studies in Computational Intelligence, Springer, pp. 357-376 (2017).
- F. Chávez, F. Fernández, C. Benavides, D. Lanza, J. Villegas, L. Trujillo, G. Olague, G. Román, Deploying Massive Runs of Evolutionary Algorithms with ECJ and Hadoop: reducing interest points required for face recognition. International Journal of High Performance Computing Applications. ACCEPTED October 2016.
- Mauro Castelli, Leonardo Trujillo, Leonardo Vanneschi, Sara Silva, Emigdio Z-Flores, and Pierrick Legrand. 2015. Geometric Semantic Genetic Programming with Local Search. In Proceedings of the 2015 on Genetic and Evolutionary Computation Conference (GECCO '15), Sara Silva (Ed.). ACM, New York, NY, USA, 999-1006.
- Perla Juárez-Smith and Leonardo Trujillo. 2016. Integrating Local Search within neat-GP. In *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion* (GECCO '16 Companion), Tobias Friedrich (Ed.). ACM, New York, NY, USA, 993-996.
- Chávez, F., Fernández, F., Benavides, C., Lanza, D., Villegas, J., Trujillo, & Román, G. (2016, March). ECJ+ HADOOP: An Easy Way to Deploy Massive Runs of Evolutionary Algorithms. In European Conference on the Applications of Evolutionary Computation (pp. 91-106). Springer International Publishing.
- F. Fernández de Vega, F. Chávez, J.A. García, M.J. Díaz, P.A. Castillo, J.J. Merelo, C. Cotta. A Cross-Platform Assessment of Energy Consumption in Evolutionary Algorithms Towards Energy-Aware Bioinspired Algorithms. Proceedings PPSN 2016. PP. 548-557.
- Mauro Castelli, Leonardo Vanneschi and Leonardo Trujillo. Energy Consumption Forecasting using Semantics Based Genetic Programming with Local Search Optimizer, Computational Intelligence and Neuroscience 2015 (2015).

Fourthly, Emigdio Z. Flores built a new software, a symbolic model builder application, named LeafGP, as a complement of GPLAB. It is a visual application intended for the generation of models based on input user data. These models are presented in a visual fashion where different metrics can be observed. The engine behind this software is the powerful Genetic Programming paradigm which is capable of building complex symbolic models with no previously knowledge from data. Different visualizations are offered with several metrics that range from the data to the extracted models. A report generation option is offered which is built based on the evolved models. See **LeafGP_001.png** to **LeafGP_006.png**.

Finally, we started to work on automatic Genetic Improvements of codes, a new hot topic on our community, in order to allow an automatic improvement of our previous development. Victor Lopez, a student of ITT, invited during the last mobility to the University of Bordeaux, is in charge of this topic under the supervision of Leonardo Trujillo and Pierrick Legrand. Here is a summary of the notions, extracted from the paper: Víctor R. López-López, Leonardo Trujillo, Pierrick Legrand, and Gustavo Olague. 2016. Genetic Programming: From Design to Improved Implementation. In

Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion (GECCO '16 Companion), Tobias Friedrich (Ed.). ACM, New York, NY, USA, 1147-1154.

Genetic programming (GP) is an evolutionary-based search paradigm that is well suited to automatically solve difficult design problems. The general principles of GP have been used to evolve mathematical functions, models, image operators, programs, and even antennas and lenses. Since GP evolves the syntax and structure of a solution, the evolutionary process can be carried out in one environment and the solution can then be ported to another. However, given the nature of GP it is common that the evolved designs are unorthodox compared to traditional approaches used in the problem domain. Therefore, efficiently porting, improving or optimizing an evolved design might not be a trivial task. In this work we argue that the same GP principles used to evolve the solution can then be used to optimize a particular new implementation of the design, following the Genetic Improvement approach. In particular, this paper presents a case study where evolved image operators are ported from Matlab to OpenCV, and then the source code is optimized an improved using Genetic Improvement of Software for Multiple Objectives (GISMOE). In the example we show that functional behavior is maintained (output image) while improving non-functional properties (computation time). Despite the fact that this first example is a simple case, it clearly illustrates the possibilities of using GP principles in two distinct stages of the software development process, from design to improved implementation.

For these reasons, we are confident that the work integrated during the development of WP 7 can become the basis for future collaborative projects between the partners, particularly aiming at the RISE call published by H2020.

2 - PROJECT ACHIEVEMENTS

2.1 Scientific highlights and research achievements

See the previous section for details.

Highlights:
Based on the progress from each work package, and evidenced by the following list of publications, we can summarize our main highlights and achievements as follows:

- Successful mobility program, will all of the partners participating in either sending or receiving students and/or researchers.
- Important synergy and transfer of knowledge between the partners
- Substantial scientific progress, in GP classification, features extraction and brain modelling. With over 60 published or submitted publications including high ranked journal papers, all of them either completely or partly based on the findings and progress from our project.
- An international workshop: NEO 2015, http://neo.cinvestav.mx/NEO2015/
- A book: Post Proceedings of NEO 2015 has been published by Springer in the Studies in Computational Intelligence series.
- 7 PhD defenses.
- New installations for signal recording and BCI experiments with the French partners, enhancing our ability to build new data sets. This was possible due to the importance of our joint collaboration between all the partners.
- New campaign of data recording, new protocol, synchronized data.
- New improvements for the software tool GPLAB, addressing a lack of local search and optimization, for improved performance, particularly in real-world problems. (see **operational_structure_GPLAB.gif**).
- New improvements on running massive instances of EAs and pool based models for EAs. (See **Report_part1_part2_UEX.pdf** and journal papers on this topic).
- An important development of software (see below).

Software:

- **GPLAB**, see section 1, WP7 and **operational_structure_GPLAB.gif**. Main developer Sara Silva.
- **LeafGP**, see **LeafGP_001.png** to **LeafGP_006.png**. Main developer Emigdio Z. Flores.
- **Analysis of the biological signals** (see **Interf_Luis.png**). Main developer Luis Herrera.
- **VCN-EEG analysis**, see **Software_EEG_analysis.pdf** and **RapportStage_ExpeVigilance_2015.pdf**. Main developer Pierrick Legrand.
- **ECJ tool** allows to lunch the algorithms over a Hadoop infrastructure. Those interested in the tool can ask for it to: fcofdez@unex.es. See **Report_part1_part2_UEX.pdf**. Main developer Francisco Fernandez de Vega.

All code derived from this project will be released as open-source as we did up to now. For instance, see https://github.com/saarahy/neat-GP-LS and https://github.com/emigdioz/LeafGP.

2.2 Transfer of knowledge and Training activities (workshops)

PhD defenses:

Laurent Vezard
Speciality: Doctorat en mathématiques appliquées et calcul scientifique.
Title: Réduction de dimension en apprentissage supervisé. Applications à l'étude de l'activité cérébrale.
Summary:
L'objectif de ce travail est de développer une méthode capable de déterminer automatiquement l'état de vigilance chez l'humain. Ainsi, les signaux electroencéphalographiques (EEG) de 58 sujets dans deux états de vigilance distincts (état de vigilance haut et bas) ont été recueillis à l'aide d'un casque à 58 électrodes posant ainsi un problème de classification binaire. Dans un premier temps, un prétraitement basé sur l'utilisation d'une décomposition en ondelettes discrètes et d'un algorithme génétique a été proposé. Dans un deuxième temps, la méthode du Common Spatial Pattern a été couplée à un algorithme génétique afin de sélectionner un sous-groupe d'électrodes discriminantes.
Enfin, un algorithme de CSP parcimonieux basé sur l'utilisation des travaux existant sur l'analyse en composantes principales parcimonieuse a été développé. Ces travaux ont abouti sur l'obtention d'un modèle permettant de prédire de manière rapide et fiable l'état de vigilance d'un nouveau sujet.
(Codirection : M. Chavent, P. Legrand, F. Faïta-Ainseba) .
See **VEZARD_LAURENT.pdf**.

Cynthia Beatriz Perez Castro
Cómputo Evolutivo como Enfoque en la Descripción del Contenido de la Imagen Aplicado al Reconocimiento de Objetos. Tesis Doctoral, Universidad de Extremadura. Departamento de Tecnologías de los Computadores y de las Comunicaciones. 2014. 149 páginas. (Codirection of Dr. Gustavo Olague and Dr. Francisco Fernández)

León Felipe Dozal Garcia
Evolución de Una Ruta Dorsal Artificial Utilizando Programación Cerebral Para la Atención Visual. Tesis Doctoral, CICESE. Departamento de Ciencias de la Computación. 06-06-2014. 102 paginas. (Invited Dr. Leonardo Trujillo as part of the committee)

Blanca Lorena Villareal Guerra
Integration of Smell in Autonomous Robotics for Search and Rescue Applications. Tesis de Doctorado. ITESM Campus Monterrey. Doctor of Philosophy in Information Technologies and Communications. Major in Intelligent Systems. School of Engineering and Information Technologies Graduate Programs. 2014. 144 pages. (Codirection of Dr. Gustavo Olague and Dr. José Luis Gordillo).

Gustavo Miguel Jorge dos Reis
Una aproximación genética a la transcripción automática de música. (Codirection of Francisco Fernández de Vega, Anibal Ferreira). 22-07-2014.

In June 2016, two members of our project defended their PhD, which are directly related to the ACOBSEC project.
Enrique Naredo, “Genetic programming based on Novelty search”.
See **Enrique.pdf** and **Enrique_Thesis.pdf**.
Yuliana Martinez, “Prediction of Performance and Problem difficulty in Genetic Programming”. See **Yuliana.pdf** and **Yuliana_tesis.pdf**.

PhD in progress:

Emigdio Z. Flores, PhD student form ITT involved in the project, provided an overview of the work related on his mobility and his actual work. The corresponding slides are available in the following files: **Emigdio.pdf**, **D43_FULL.pdf**.

A PhD student from Mexico, **Arturo Ocampo**, is being co-advised by **G. Olague** and **F. Fernández de Vega**. The student is working in image recognition by means of evolutionary algorithms.

We would thank all our PhD students for their participation to the project and the quality of their scientific contributions.

Workshops:

Through the mobility program, we have carried out several seminars related to WP1 between January and June, at the University of Bordeaux, INRIA, University of Extremadura, and Lisbon with our European partners, and at ITT and CICESE on the Mexican side (See for example **ACOBSEC_WP1_ITT_DOC_CORRFILT_R1.0.pdf**).
Participants include: Laurent Vezard, Frederique Faïta Ainseba, Marie Chavent, Yuliana Martinez, Enrique Naredo, Emigdio Z. Flores, Pierrick Legrand, Sara Silva (Europe); Francisco Fernandex, Francisco Chavez, Leonardo Trujillo, Daniel Lanza, Gustavo Olague, Eddie Clemente, León Dozal, Daniel Hernández, Victor López, Arturo Sotelo (Mexico).
As part of the work package 1 a Special Session was held at the EVOLVE 2014 conference, held on July 1 – 4th 2014 in Beijing China; see:
http://www.evolve-conference.org/2011-03-31-04-23-33

Seminars/workshops related to WP2 were organized in France (INRIA and University of Bordeaux). The main participants in these events were: Yuliana Martinez (ITT), Marie Chavent (INRIA), Frederique Faïta-Ainseba (UB), Francisco Fernandez (UEX), Enrique Naredo (ITT), Emigdio Z. Flores (ITT), Pierrick Legrand (INRIA-UB).

Seminars/workshops related to WP5 were organized in France (INRIA and University of Bordeaux) and in Spain (University of Extremadura). The main participants in these events were: Yuliana Martinez (ITT), Francisco Chavez (UEX), Francisco Fernandez (UEX), Enrique Naredo (ITT), Emigdio Z. Flores (ITT), Pierrick Legrand (INRIA-UB), Luis Muñoz (ITT), Leonardo Trujillo (ITT), Enrique Mediero (UEX), Daniel Lanza (UEX).
As part of the project, some talks has been delivered by members of the UEX team:
F. Fernández de Vega: Improving Computational Creativity. 07-05-2014. Jerez de la Frontera, Spain.
F. Fernández de Vega. IEEE WCCI 2014 Tutorial : Creative Evolutionary Algorithms. 07-07-2014. Beijing, China.
F. Fernández de Vega: An unconventional execution of the Evolutionary Algorithm: Building bridges to the arts.
F. Fernández de Vega. Implicitly Parallel Evolutionary Algorithms. 8/10/2015. Intelligent Distributed Conference 2015. Guimaraes, Portugal.

NEO 2015

The project team members, especially Leonardo Trujillo and Pierrick Legrand, co-organized the Workshop on Numerical and Evolutionary Computation NEO 2015 held from the 23rd to the 25th of September 2015 (see http://neo.cinvestav.mx/NEO2015/) in collaboration with CINVESTAV-IPN, one of the leading research centers in Mexico. At this event, one of the main tracks focused on Genetic Programming, and included the participation of some of the leading researchers in the field, including Dr. Leonardo Vanneschi (winner of the 2015 Award for Outstanding Contributions in Evolutionary

Computation) and Dr. Mario Graff. Moreover, all of the student participants were key participants at the event, giving several talks and working as organizers. From this event, a post-proceedings book has being edited, along with collaborators Dr. Oliver Schütze and Dr. Yazmin Maldonado (see http://neo.cinvestav.mx/NEO2015/index.php/post-proceedings), has been published in September 2016 by Springer in the Studies in Computational Intelligence series.

NEO 2015, Results of the Numerical and Evolutionary Optimization Workshop NEO 2015 held at September 23-25 2015 in Tijuana, Mexico. Schütze, O., Trujillo, L., Legrand, P., Maldonado, Y. (Eds.), 17 chapters, series Studies in Computational Intelligence, Springer, 444 pages (2017).

Other relevant Conferences:

The uex team has organized the following conferences, workshops and special sessions:

- MAEB 2015: X Spanish Conference on Metaheuristics, Evolutionary and Bioinspired Algorithms. February 2015.
- Special Sessions on Art and Creativity at CEC 2014, CEC 2015, WCCI 2016.
- EVOPAR 2016, 2015, 2014. Part of EvoApplications, the European Conference on Applications of Evolutionary Algorithms.
- ACM GECCO 2015 DETA Track.

Teaching / Recording sessions:

F. Faïta gave a seminar to 3 Mexican PhD students (E. Naredo, Y. Martinez, E. Flores), regarding the state of the art in BCI signal measurements with EEG, covering aspects related to hardware and experimental setups. Moreover, the students participated in the setup and acquisition of new experimental data.
F. Faïta and P. Legrand gave seminars and teaching session to the student focusing on aspects related to pre-processing of EEG data.
Francisco Fernandez and Daniel Lanza gave teaching session to Mexican researchers on tools for big data processing, analysis and management.
During his mobility to ITT, Pierrick Legrand gave 5 lectures to Mexican students and researchers.

2.3 Dissemination of results (journal papers, book chapters, conference proceedings and other publications)

WP1:

NEO 2015, Results of the Numerical and Evolutionary Optimization Workshop NEO 2015 held at September 23-25 2015 in Tijuana, Mexico. Schütze, O., Trujillo, L., Legrand, P., Maldonado, Y. (Eds.), 17 chapters, series Studies in Computational Intelligence, Springer, 444 pages (2017). 10.1007/978-3-319-44003-3

Enrique Naredo, Leonardo Trujillo, Pierrick Legrand, Sara Silva, and Luis Muñoz. 2016. Evolving genetic programming classifiers with novelty search. *Inf. Sci.* 369, C (November 2016), 347-367. 10.1016/j.ins.2016.06.044

Leonardo Trujillo, Luis Muñoz, Edgar Galván-López, Sara Silva. neat Genetic Programming: Controlling Bloat Naturally, Information Sciences 333:21-43 (2016).

Perla Juárez-Smith and Leonardo Trujillo. 2016. Integrating Local Search within neat-GP. In *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion* (GECCO '16 Companion), Tobias Friedrich (Ed.). ACM, New York, NY, USA, 993-996. 10.1145/2908961.2931659

Hidalgo, J Ignacio and de Vega, Francisco Fernandez, Parallel Bioinspired Algorithms on the Grid and Cloud, Journal of Grid Computing, Volume 13. DOI: 10.1007/s10723-014-9322-7

Leonardo Trujillo. Book Review: Kenneth O. Stanley and Joel Lehman Why Greatness Cannot Be Planned - The Myth of the Objective Springer, 2015, Softcover ISBN 978-3-319-15523-4, to appear in Genetic Programming and Evolvable Machines.

Semantic Learning Machine: A Feedforward Neural Network Construction Algorithm Inspired by Geometric Semantic Genetic Programming, Ivo Gonçalves, Sara Silva, Carlos M. Fonseca, in Progress in Artificial Intelligence: 17th Portuguese Conference on Artificial Intelligence, EPIA 2015, Coimbra, Portugal, September 8-11, 2015. Proceedings. Springer, 811 pages. DOI: 10.1007/978-3-319-23485-4_28

Leonardo Trujillo and Luis Munoz and Enrique Naredo and Yuliana Martinez. NEAT, There's No Bloat. In M. Nicolau et al. editors, 17th European Conference on Genetic Programming, volume 8599, pages 174-185, Granada, Spain, 2014. Springer (http://rd.springer.com/chapter/10.1007%2F978-3-662-44303-3_15)

Yuliana Martinez and Leonardo Trujillo and Enrique Naredo and Pierrick Legrand. A Comparison of Fitness-Case Sampling Methods for Symbolic Regression with Genetic Programming. In Alexandru-Adrian Tantar et al. editors, EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation V, volume 288, pages 201-212, Peking, 2014. Springer. (http://rd.springer.com/chapter/10.1007%2F978-3-319-07494-8_14)

Emigdio Z-Flores and Leonardo Trujillo and Oliver Schuetze and Pierrick Legrand. Evaluating the Effects of Local Search in Genetic Programming. In Alexandru-Adrian Tantar et al. editors, EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation V, volume 288, pages 213-228, Peking, 2014. Springer (http://rd.springer.com/chapter/10.1007%2F978-3-319-07494-8_15)

E. Galvan-Lopez, C. Harris, L. Trujillo, K. Rodriguez-Vazquez, S. Clarke and V. Cahill. Autonomous Demand-Side Management System Based on Monte Carlo Tree Search. IEEE International Energy Conference (EnergyCon), pages 1325 - 1332, Dubrovnik, Croatia, May 13 - 16, 2014. IEEE Press. (http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=6850585)

Vijay Ingalalli and Sara Silva and Mauro Castelli and Leonardo Vanneschi. A Multi-dimensional GeneticProgramming Approach for Multi-class ClassificationProblems. In M. Nicolau and K. Krawiec and M. I.Heywood and M. Castelli and P. Garci-Sanchez and J.J. Merelo and V. M. R. Santos and K. Sim editors,17th European Conference on Genetic Programming,volume 8599, pages 48-60, Granada, Spain, 2014.Springer.

Luis Munoz and Sara Silva and Leonardo Trujillo. M3GP: Multiclass Classification with GP. In Penousal Machado and Malcolm I. Heywood and James McDermott and Mauro Castelli and Pablo Garcia-Sanchez and Paolo Burelli and Sebastian Risi and Kevin Sim editors, 18th European Conference on Genetic Programming, volume 9025, pages 78-91, Copenhagen, 2015. Springer

Emigdio Z-Flores, Leonardo Trujillo, Oliver Schütze and Pierrick Legrand. 2014. A Local Search Approach to Genetic Programming for Binary Classification. In Proceedings of the seventeenth annual conference on Genetic and evolutionary computation conference (GECCO '15), Ana Esparcia (Ed.). ACM.

Mauro Castelli, Leonardo Trujillo, Leonardo Vanneschi, Sara Silva, Emigdio Z-Flores and Pierrick Legrand. 2014. Geometric Semantic Genetic Programming with Local Search. In Proceedings of the seventeenth annual conference on Genetic and evolutionary computation conference (GECCO '15), Ana Esparcia (Ed.). ACM.

R.C. Guitiérrez-Urquidez, G. Valencia-Palomo, O.M. Rodríguez-Elías, L. Trujillo. Systematic selection of tuning parameters for efficient predictive controllers using a multiobjective evolutionary algorithm, Applied Soft Computing 31:326–338 (2015).

Victor H. Diaz-Ramirez, Andres Cuevas, Vitaly Kober, Leonardo Trujillo, Abdul Awwal. Pattern recognition with composite correlation filters designed with multi-objective combinatorial optimization, Optics Communications, 338(1): 77-89 (2015).

de Vega, F Fernandez and Chavez F de la O and Trujillo, L and Mediero E Manuel and Munoz, L. A Hybrid ECJ+Boinc Tool for Distributed Evolutionary Algorithms, in Research in Computing Science, pp 120--130, 2014.

Eddie Clemente and Francisco Chavez and Francisco Fernandez de Vega and Gustavo Olague . Self-adjusting focus of attention in combination with a genetic fuzzy system for improving a laser environment control device system, Applied Soft Computing, volume 32, pages 250 - 265, 2015, issn 1568-4946, doi = http://dx.doi.org/10.1016/j.asoc.2015.03.011, http://www.sciencedirect.com/science/article/pii/S1568494615001647,

Francisco Chavez de la O and Eddie Clemente and Daniel E. Hernandez and Francisco Fernandez de Vega and Gustavo Olague, A Multi-objective Evolutionary Algorithm for Interaction Systems Based on Laser Pointers, Applications of Evolutionary Computation - 18th European Conference EvoApplications 2015, Copenhagen, Denmark, April 8-10, 2015, Proceedings, pages 504--516, 2015, DBLP:conf/evoW/2015a, http://dx.doi.org/10.1007/978-3-319-16549-3_41, doi = {10.1007/978-3-319-16549-3_41, http://dblp.uni-trier.de/rec/bib/conf/evoW/ChavezCHVO15.

J.M Narvaez and L. Trujillo and F. Fernandez de Vega, and S. Silva and P. Legrand, Diseñando problemas sintéticos de clasificación con superficie de aptitud deceptiva, Proceedings MAEB 2015, 2015.

F. Fernandez de Vega and F. Chavez and A. Mora and C. Fernandez and J.J Merelo, Estudio preliminar sobre visualización y clasificación de la calidad de la emisión de sonido en el clarinete, Proceedings MAEB 2015, pages 51-58, 2015.

F. Chavez and D. Lanza and F. Fernandez de Vega and G. Olague and L. Trujillo, Integrando ECJ con Hadoop, Proceedings MAEB 2015, 511-518, 2015.

F. Chavez, E. Clemente, D. Hernandez,F. Fernandez de Vega and G. Olague. Reconocimiento de punto láser ensistemas de interacción mediante algoritmos multiobjetivos, Proceedings MAEB 2015 pages 567-574, 2015.

WP3:

Emigdio Z-Flores, Leonardo Trujillo, Arturo Sotelo, Pierrick Legrand and Luis Coria. Regularity and Matching Pursuit Feature Extraction for the Detection of Epileptic Seizures, to appear in Journal of Neuroscience Methods.

Pierrick Legrand, Laurent Vezard, Marie Chavent, Frédérique Faïta-Ainseba, Leonardo Trujillo, Feature extraction and classification of EEG signals. The use of a genetic algorithm for an application on alertness prediction. In "Guide to Brain-Computer Music Interfacing", Eduardo Miranda, Julien Castet and Benjamin Knapp Eds., Springer.
http://www.springer.com/computer/hci/book/978-1-4471-6583-5

Laurent Vézard, Pierrick Legrand, Marie Chavent, Frédérique Faïta-Ainseba, Julien Clauzel, and Leonardo Trujillo. Classification of EEG signals by evolutionary algorithm. In "Advances in Knowledge Discovery and Management Vol.4", Studies in Computational Intelligence. Volume 527, pp133-153.
http://www.springer.com/engineering/computational+intelligence+and+complexity/book/978-3-319-02998-6

Laurent Vezard, Pierrick Legrand, Marie Chavent, Frédérique Faïta-Aïnseba, Leonardo Trujillo. Detecting mental states of alertness with genetic algorithm variable selection, in Applied Soft Computing (2015). 32: 113-131.

Arturo Sotelo, Enrique D. Guijarro and Leonardo Trujillo. Seizure States Identification in Experimental Epilepsy using Gabor Atom Analysis, in Journal of Neuroscience Methods (2015).

WP4:

Pierrick Legrand, Laurent Vezard, Marie Chavent, Frédérique Faïta-Ainseba, Leonardo Trujillo, Feature extraction and classification of EEG signals. The use of a genetic algorithm for an application on alertness prediction. In "Guide to Brain-Computer Music Interfacing", Eduardo Miranda, Julien Castet and Benjamin Knapp Eds., Springer.
http://www.springer.com/computer/hci/book/978-1-4471-6583-5

Laurent Vézard, Pierrick Legrand, Marie Chavent, Frédérique Faïta-Ainseba, Julien Clauzel, and Leonardo Trujillo. Classification of EEG signals by evolutionary algorithm. In "Advances in

Knowledge Discovery and Management Vol.4", Studies in Computational Intelligence. Volume 527, pp133-153.
http://www.springer.com/engineering/computational+intelligence+and+complexity/book/978-3-319-02998-6

Laurent Vezard, Pierrick Legrand, Marie Chavent, Frédérique Faïta-Aïnseba, Leonardo Trujillo. EEG classification for the detection of mental states, Applied Soft Computing 32:113-131 (2015).

WP5:

Yuliana Martínez, Enrique Naredo, Leonardo Trujillo, Pierrick Legrand and Uriel López. A comparison of fitness-case sampling methods for genetic programming, in the Journal of Experimental & Theoretical Artificial Intelligence.

Enrique Naredo, Paulo Urbano, Leonardo Trujillo. The Training Set and Generalization in Grammatical Evolution for Autonomous Agent Navigation, in Soft Computing.

Yuliana Martínez, Leonardo Trujillo, Pierrick Legrand and Edgar Galván-López. Prediction of Expected Performance for a Genetic Programming Classifier, Genet Program Evolvable Mach (2016) 17: 409. http://link.springer.com/article/10.1007/s10710-016-9265-9

Francisco Fernández de Vega, Leonardo Trujillo, Francisco Chávez, Enrique Mediero and Luis Muñoz. 2014. A Hybrid ECJ+BOINC Tool for Distributed Evolutionary Algorithms, Research in Computing Science, Advances in Computing Science, Control and Communications, Mireya Garcia et al. (Eds.), 69, 120-130.

Mario García-Valdez, Leonardo Trujillo, Juan Julián Merelo-Guérvos, Francisco Fernández-de-Vega. Randomized Parameter Settings for Heterogeneous Workers in a Pool-Based Evolutionary Algorithm, in Proceedings of Parallel Problem Solving from Nature – PPSN XIII, LCNS Vol. 8672, 2014, pp 702-710.

Mario García-Valdez, Leonardo Trujillo, Juan-J Merelo, Francisco Fernández de Vega & Gustavo Olague. The EvoSpace Model for Pool-Based Evolutionary Algorithms, Journal of Grid Computing, Volume 13. DOI: 10.1007/s10723-014-9319-2

WP6:

Evolutionary multi-objective visual cortex for object classification in natural images, Daniel E. Hernandez, Eddie Clemente, Gustavo Olague, in Journal of Computational Science, 2 November 2015, Elsevier.

Dozal García, L. F., G. Olague Caballero(*), E. H. Clemente Torres y D. E. Hernández Morales. (2014) Brain programming for the evolution of an artificial dorsal stream. Cognitive Computation. http://link.springer.com/article/10.1007%2Fs12559-014-9251-6

Hernández Morales, D. E., G. Olague Caballero(*), E. H. Clemente Torres y L. F. Dozal García. (2014) Optimizing a conspicuous point detector for camera trajectory estimation with brain programming. En: O. Schütze et al, Editores. "Evolve - a Bridge Between Probability, Set Oriented Numerics, and Evolutionary Computation III". Springer. ISBN: 978-3-319-01460-9. 121-140 p. http://link.springer.com/chapter/10.1007/978-3-319-01460-9_6

Olague Caballero, G.(*), L. F. Dozal García, E. H. Clemente Torres y A. I. Ocampo Torres. (2014) Optimizing an artificial dorsal stream on purpose for visual attention. En: O. Schütze et al, Editores. "Evolve - a Bridge Between Probability, Set Oriented Numerics, and Evolutionary Computation III". Springer. ISBN: 978-3-319-01459-3. 141-166 p.
http://link.springer.com/chapter/10.1007%2F978-3-319-01460-9_7

Olague Caballero, G.(*), E. H. Clemente Torres, L. F. Dozal García y D. E. Hernández Morales. (2014) Evolving an artificial visual cortex for object recognition with brain programming. En: O. Schütze et al, Editores. "Evolve - a Bridge Between Probability, Set Oriented Numerics, and Evolutionary Computation III". Springer. ISBN: 978-3-319-01459-3. 97-119 p.
http://link.springer.com/chapter/10.1007/978-3-319-01460-9_5

B. Lorena Villarreal, Gustavo Olague, José Luis Gordillo: Odor Plume Tracking Algorithm Inspired on Evolution. MCPR 2014: 321-330.
http://link.springer.com/chapter/10.1007%2F978-3-319-07491-7_33

Gustavo Olague. Evolutionary Computer Vision. Springer.
10.1007/978-3-662-43693-6

de Vega, F Fernandez and Cruz, C and Navarro, L and Hernandez, P and Gallego, T and Espada, L. Unplugging Evolutionary Algorithms: an experiment on human-algorithmic creativity, in Genetic Programming and Evolvable Machines, volume 15, number 4, pages 379--402, 2014, Springer.
http://link.springer.com/article/10.1007%2Fs10710-014-9225-1#page-1

WP7:

Leonardo Trujillo, Emigdio Z-Flores, Perla S. Juárez Smith, Pierrick Legrand, Sara Silva, Mauro Castelli, Leonardo Vanneschi, Oliver Schütze and Luis Muñoz. Local Search is Underused in Genetic Programming. In Rick Riolo et al. editors, Genetic Programming Theory and Practice XIV, Ann Arbor, USA, 2017. Springer.

Daniel Lanza, F. Chávez, Francisco Fernández, M. Garcia-Valdez,Leonardo Trujillo and Gustavo Olague, Profiting from Several Recommendation Algorithms Using a Scalable Approach, in NEO 2015, Results of the Numerical and Evolutionary Optimization Workshop NEO 2015 held at September 23-25 2015 in Tijuana, Mexico. Schütze, O., Trujillo, L., Legrand, P., Maldonado, Y. (Eds.), series Studies in Computational Intelligence, Springer, pp. 357-376 (2017).

Josúe Zárate, Leonardo Trujillo, Salvador de Lara, Mauro Castelli, Emigdio Z-Flores, Luis Muñoz, Ales Popovic. Automatic Modeling of a Gas Turbine using Genetic Programming: An Experimental Study, in Applied Soft Computing. 10.1016/j.asoc.2016.11.019

J.R. Cárdenas Valdez, Emigdio Z. Flores, José Cruz Nuñez Pérez and Leonardo Trujillo, Local Search Approach to Genetic Programming for RF-PAs Modeling Implemented in FPGA, in NEO 2015, Results of the Numerical and Evolutionary Optimization Workshop NEO 2015 held at September 23-25 2015 in Tijuana, Mexico. Schütze, O., Trujillo, L., Legrand, P., Maldonado, Y. (Eds.), series Studies in Computational Intelligence, Springer, pp. 67-88 (2017).

F. Chávez, F. Fernández, C. Benavides, D. Lanza, J. Villegas, L. Trujillo, G. Olague, G. Román, Deploying Massive Runs of Evolutionary Algorithms with ECJ and Hadoop: reducing interest points required for face recognition. International Journal of High Performance Computing Applications. October 2016. 10.1007/978-3-319-31153-1_7.

Yuliana Martínez, Enrique Naredo, Leonardo Trujillo, Pierrick Legrand and Uriel López. A comparison of fitness-case sampling methods for genetic programming, in the Journal of Experimental & Theoretical Artificial Intelligence.

Mauro Castelli, Ivo Goncalves, Leonardo Trujillo and Ales Popovic. An Evolutionary System for Ozone Concentration Forecasting, in Information Systems Frontiers. 10.1007/s10796-016-9706-2

Mauro Castelli, Leonardo Trujillo, Leonardo Vanneschi, and Aleš Popovič. 2016. Prediction of relative position of CT slices using a computational intelligence system. Appl. Soft Comput. 46, C (September 2016), 537-542.

José Enrique Hernández-Beltrán; Víctor H. Díaz-Ramírez; Leonardo Trujillo and Pierrick Legrand. Restoration of degraded images using genetic programming, Proc. SPIE 9970, Optics and Photonics for Information Processing X, 99700K (September 14, 2016). 10.1117/12.2237092

Víctor R. López-López, Leonardo Trujillo, Pierrick Legrand, and Gustavo Olague. 2016. Genetic Programming: From Design to Improved Implementation. In Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion (GECCO '16 Companion), Tobias Friedrich (Ed.). ACM, New York, NY, USA, 1147-1154. 10.1145/2908961.2931693

Perla Juárez-Smith and Leonardo Trujillo. 2016. Integrating Local Search within neat-GP. In *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion* (GECCO '16 Companion), Tobias Friedrich (Ed.). ACM, New York, NY, USA, 993-996. 10.1145/2908961.2931659

Chávez, F., Fernández, F., Benavides, C., Lanza, D., Villegas, J., Trujillo, & Román, G. (2016, March). ECJ+ HADOOP: An Easy Way to Deploy Massive Runs of Evolutionary Algorithms. In European Conference on the Applications of Evolutionary Computation (pp. 91-106). Springer International Publishing.

F. Fernández de Vega, F. Chávez, J.A. García, M.J. Díaz, P.A. Castillo, J.J. Merelo, C. Cotta. A Cross-Platform Assessment of Energy Consumption in Evolutionary Algorithms Towards Energy-Aware Bioinspired Algorithms. Proceedings PPSN 2016. PP. 548-557. *10.1007/978-3-319-45823-6_51*

Mauro Castelli, Leonardo Vanneschi and Leonardo Trujillo. Energy Consumption Forecasting using Semantics Based Genetic Programming with Local Search Optimizer, Computational Intelligence and Neuroscience 2015 (2015).

Laredo, JLJ and Bouvry, Pascal and Gonzalez, DL and de Vega, F Fernandez and Arenas, MG and Merelo, JJ and Fernandes, Carlos M. Designing robust volunteer-based evolutionary algorithms, in Genetic Programming and Evolvable Machines, volume 15, number 3, pages 221--244, 2014, Springer.
http://link.springer.com/article/10.1007%2Fs10710-014-9213-5#page-1

D. Lanza, F. Chávez de la O, F. Fernández de Vega, C. Benavides, J. Villegas, Speeding up evolutionary approaches to face recognition by means of Hadoop, Proceedings NEO 2015, Tijuana, Mexico

Hidalgo, J Ignacio and de Vega, Francisco Fernandez, Parallel Bioinspired Algorithms on the Grid and Cloud, Journal of Grid Computing, Volume 13. DOI: 10.1007/s10723-014-9322-7

Mauro Castelli, Raul Sormani, Leonardo Trujillo and Ales Popovic. Predicting Per Capita Violent Crimes in Urban Areas: an Artificial Intelligence Approach, in Journal of Ambient Intelligence and Humanized Computing.

Mauro Castelli, Leonardo Trujillo, Leonardo Vanneschi, Ales Popovic. Prediction of energy performance of residential buildings: a genetic programming approach, Energy and Buildings 102:67-74 (2015).

Synthesis of odor tracking algorithms with genetic programming, B. Lorena Villarreal, Gustavo Olague, J.L. Gordillo, in Neurocomputing, Volume 175, Part B, 29 January 2016, Pages 1019–1032, Elsevier.

-3. PROJECT MANAGEMENT

All of the partners involved in the project have been very active; all of them have either sent or received mobilities over this first stage. Moreover, all of the participants have produced scientific results of very high quality thanks to the work derived from this project; this is clear by the large number of publications already derived from the collaborations made possible by this project.

The mobilities carried out during the project have had undeniable effects on the participants from each partner institution. In particular, for the students there have been several objective and subjective benefits. Regarding the objective measures of success, the students have participated in several high-quality publications derived from the project, including 31 journal papers, 2 books, 28 conference papers and 1 book chapter. The students have also had the opportunity to present their work in several high profile conferences, in international summer (winter) schools, as well as several workshops and seminars organized by each partner institution. Also, of the students that have participated (to the mobilities), 2 have received their doctoral degree, 2 have received their masters, and 2 have received their engineering degrees. Secondly, and maybe more importantly, the participation of the students in the project, especially those from Mexico, has allowed them to discuss their work with important researches in the field that they would otherwise not have met, has provided them with useful experience and skills, particularly in developing collaborative research and presenting their results to multidisciplinary audiences, and finally it has given them a richer and deeper understanding of how scientific research can be carried out between different researchers, areas, labs, institutions and countries, preparing them for the future world-wide challenges that young researches will surely face in a world that continues to become smaller and more deeply intertwined.

-4. ANNEXES

https://www.math.u-bordeaux.fr/~plegra100p/FINAL_REPORT_ACOBSEC/Annexes.rar