ECOLE
Experience-based Computation:
Learning to Optimise

D14 - Skills Report

ECOLE
Experience-based Computation: Learning to Optimise
766186

H2020 MSCA-ITN
30th September 2019

1. Introduction

The management team suggested programs of training in the proposal that are now being adhered to by the PhD students with several completed. Their progress on the courses are discussed at thesis meetings and more recently in their career development plans. ESRs should be horizon scanning and proactively seeking to complete the courses as well as any additional opportunities for implementing training in an activity that utilizes skills learnt

1.1. Recruitment

Recruitment has been completed as outlined in length previously in D25 – Progress Report, the gender split shown in Table 1 below. All ESRs have now met the requirement of admission the nominated university and the first secondments are underway.

Table 1 - Gender Balance of ECOLE ESRs

Institution	Total Applicants	Shortlisted		Recruited	
		Male	Female	Male	Female
University of Birmingham	14	4	1	2	0
Universiteit Leiden	95	3	3	2	1
NEC	200	3	1	1	0
Honda	31	6	2	1	1
Total	**340**	**16**	**7**	**6**	**2**

2. Research Progress

2.1. Publications

The management team have defined a publication procedure for approval of manuscripts prior to the submission of the journal. This allows all beneficiaries to comment on published work in addition to ensuring that any potential intellectual property is not disclosed prior to protection. For the overview of the publication process see APPENDIX A and example of consent for publication form see APPENDIX B.

Publications have already been prepared by the ESRs and their supervisors. The publications (list of publications and the outcomes can be found in Table 2) have been submitted to the 2019 IEEE Symposium Series on Computational Intelligence (IEEE SSCI 2019) in Xiamen, China. In addition, publications have also been submitted to the workshop on Learning and Mining with Industrial Data (LMID 2019). The workshop has been organized by the ECOLE supervisory board and topics of interests included but not limited to low-resource machine learning, mining graph and networked data[1]. LMID aims to bring together researchers and practitioners from academia and industry to discuss challenges of learning and mining with industrial data.

[1] Workshop home page available from ECOLE website - https://ecole-itn.eu/lmid-2019/

Table 2 - Manuscripts in Preparation

ESR Authors	Manuscript Name	Submitted	Outcome
Sibghat Ullah, Hao Wang, Thomas Back, Bernhard Sendhoff, Stefan Menzel	An Empirical Comparison of Meta-Modeling Techniques for Robust Design Optimisation	SSCI	Accepted
Sneha Saha, Thiago Rios, Leandro Minku, Xin Yao, Zhao Xu, Bernhard Sendhoff, Stefan Menzel	Optimal Evolutionary Optimisation Hyperparameters to Mimic Human User Behaviour	SSCI	Accepted
Thiago Rios, Stefan Menzel	On The Efficiency of Latent Representation of Point Cloud Auto-Encoders in Evolutionary Target Shape Matching Optimisation	SSCI	Accepted
Thiago Rios, Patricia Wollstadt, Bas van Stein, Thomas Back, Zhao Xu, Bernhard Sendhoff, Stefan Menzel	Scalability of Learning Tasks on 3D CAE Models Using Point Cloud Autoencoders	SSCI	Accepted
Gan Ruan, Leandro Minku, Stefan Menzel, Bernhard Sendhoff, Xin Yao	When and How to Transfer Knowledge in Dynamic Multi-Objective Optimisation	SSCI	Accepted
Stephen Friess, Peter Tino, Stefan Menzel, Bernhard Sendhoff, Xin Yao	Learning Transferable Variation Operators in Continuous Genetic Algorithm	SSCI	Accepted
Jiawen Kong, Wojtek Kowalczyk, Duc Anh Nguyen, Stefan Menzel, Thomas Back	Hyperparameter Optimisation for Improving Classification Under Class Imbalance	SSCI	Accepted
Jiawen Kong, Stefan Menzel, Wojtek Kowalzyk, Thomas Back	How the Data Complexity Affects the Choice of Resampling Technique Under Class Imbalance	LMID	Rejected
Sneha Saha, Thiago Rios, Stefan Menzel, Bernhard Sendhoff, Thomas Back, Xin Yao, Zhao Xu and Patricia Wollstadt	Learning Time-Series Data of Industrial Design Optimization Using Recurrent Neural Networks	LMID	Accepted
Gan Ruan, Leandro Minku, Stefan Menzel, Bernhard Sendhoff, Xin Yao	A Computational Study on the Computational Cost and Quality Of Knowledge Transfer in Dynamic Multi-Objective Optimization	LMID	Accepted
Stephen Friess, Peter Tino, Stefan Menzel, Bernhard Sendhoff, Xin Yao	Extraction and Analysis Of Variable Preferences in Stochastic Optimisation	LMID	Rejected
Giuseppe Serra, Zhao Xu, Xin Yao, Peter Tino	Explainable Embedding with Review Data	LMID	Rejected

2.2. Work packages

Table 3 - Update on Work Packages

WP1 – Experience-guided Optimisation in Automotive Product Design **Honda Research Institute**
In WP1, we develop techniques to accumulate, explore and exploit experience across a series of automotive-like product development circles. Specific questions are concerned, first, with the design of efficient representations, which we research and propose based on the latent parameters of a geometric deep learning framework. Second, we will extract and utilize user design preferences in a multi-criteria optimization framework, where the criteria relate to aesthetics and aerodynamics. One aspect of realizing WP1 is to research and develop an efficient representation for optimizing 3D designs. Traditionally, free form deformation is a state-of-the-art approach to be utilized in 3D design optimization and showed a range of successes in the development of vehicles, racecars and aircraft domain. Intuitive user interactions and a fair trade-off between optimization parameters and design flexibility allows a reasonable integration to CAE and CAD tools. Recent development in geometric deep learning, namely point cloud and voxel autoencoders, enabled the unsupervised learning and extraction of a low dimensional embedding. Especially point cloud autoencoders have been used on public benchmark data sets of 3D geometries for generating efficient latent representations that allow introducing meaningful design variations with a low number of parameters. In WP1, we implemented a point cloud autoencoder and successful validated it against these published results from literature. To evaluate the potential for practical use cases we set-up 3D vehicle design optimizations for a comparison of optimization runs using the point cloud autoencoder with state-of-the-are free form deformation. We showed that if the autoencoder is trained on a meaningful set of vehicle designs an optimization based on the autoencoder provides a good convergence behavior with high performance. Of course, a limiting factor is the meaningful database on which the learning of the autoencoder strongly depends. The results of this study have been accepted for publication and presentation at IEEE Symposium Series on Computational Intelligence (SSCI) 2019 [1]. Besides the successful application of the point cloud autoencoder in 3D design optimization it is important to understand its potential and limits in more depth. Therefore, we focused on evaluating the scalability of the autoencoder towards high-dimensional 3D design data. By choosing a different loss functions during the training of the deep network we observed an increased performance during the learning with similar results to networks published in literature. Furthermore, performance comparisons on different sizes of the latent space and layer sizes have been carried out. These experiments were finally extended by an analysis of the differentiation of geometric features by correlation calculation. These experiments provide valuable insights in the usage of point cloud encoders for 3D geometric data and have been accepted for publication and presentation at IEEE SSCI 2019 [2]. As a contribution to WP1, we set-up a geometric representation based on unsupervised deep learning which can be successfully utilized in design optimization and for the learning of aerodynamic data in the next steps. In addition, the datasets, that are contributing to D1.1, generated during the

optimization and for the training of the autoencoder are valuable data for model building and algorithm tests to WP2 and WP3. Both methods used in the developed software tools are important ingredients towards the multi-criteria software environment targeting D1.2.

As a second aspect for realizing WP1, we implemented and set-up experiments with time series models for including user preferences in the development process. In a series of first experiments, we conducted a human user study comprising an interactive software framework allowing the user to modify control points of a free form deformation system to visually align a given 3D shape with a target 3D shape. Each user modification has been recorded to create a time series of user guided (preferred) optimizations towards a target shape. In the next step, we evaluated the user recordings and set-up a computational target shape matching optimization using evolutionary optimization. By tuning the hyperparameters of the evolutionary optimization we could align the user modifications with the automated process. The outcome of this work is twofold. First, we could set-up an evolutionary optimization according to the different (three) user experience levels. Second, we can now utilize the tuned evolutionary optimization to generate a larger set of time series automatically that circumvents the challenge of generating a large amount of time series by hand. Such a large amount is required to train modern recurrent neural network models like long short term memory networks. The results of this study have been accepted for publication and presentation at IEEE SSCI 2019 [3].

Based on this work, we generated a large amount of time series for a target shape matching of vehicle models using the evolutionary optimizer. The recorded optimization steps were afterwards utilized to train a recurrent network that allowed us to predict based on a window of past modifications steps either the next potential steps or the next three potential steps. We could show that the network predictions outperformed a data driven base model and point to promising next steps when even more data would be available for designing a cooperative engineering system. The results of this study have been accepted for publication and presentation at the Workshop on Learning and Mining with Industrial Data (LMID) which is part of the IEEE International Conference on Data Mining (ICDM) 2019 [4]. The developed methods and software pieces are important ingredients for building towards the software package D1.3.

References

[1] Thiago Rios, Patricia Wollstadt, Bas van Stein, Thomas Bäck, Zhao Xu, Bernhard Sendhoff, Stefan Menzel, „Scalability of Learning Tasks on 3D CAE Models using Point Cloud Autoencoders”, IEEE Symposium Series on Computational Intelligence, China, 2019 (accepted)

[2] Thiago Rios, Thomas Bäck, Bas van Stein, Bernhard Sendhoff, Stefan Menzel, “On the Efficiency of a Point Cloud Autoencoder as a Geometric Representation for Shape Optimization”, IEEE Symposium Series on Computational Intelligence, China, 2019 (accepted)

[3] Sneha Saha, Thiago Rios, Leandro L. Minku, Xin Yao, Zhao Xu, Bernhard Sendhoff, Stefan Menzel, “Optimal Evolutionary Optimization Hyperparameters to Mimic Human User Behavior”, IEEE Symposium Series on Computational Intelligence, China, 2019 (accepted)

[4] Sneha Saha, Thiago Rios, Stefan Menzel, Bernhard Sendhoff, Thomas Bäck, Xin Yao, Zhao Xu, Patricia Wollstadt, „Learning Time-series Data of Industrial Design Optimization using

Recurrent Neural Networks", ICDM Workshop: Learning and Mining with Industrial Data (LMID), China, 2019 (accepted)

WP2 – Learning in Robust and Efficient Multi-Objective Optimisation
Leiden

Update on work packages: all ESRs who are involved in this work package have conducted thorough literature reviews on topics including Robust optimization, Algorithm Configuration, Imbalance data classification and Geometric Deep Learning, which are essential for the aim of this WP. ESR1 has two conference papers (IEEE SSCI '19) accepted on applying autoencoders to target shape matching optimization. ESR3 has one paper accepted in the same venue, where the performance of data-driven modelling methods are compared in the robust optimization setup. ESR5 investigated the problem of classifying extremely imbalanced data set, the result of which will also appear in the same conference. ESR4 participated in this research with the aim to improve the classification quality via hyper-parameter optimization, together with ESR5.

Update on deliverables: In total, four papers are accepted and to appear in the proceedings of IEEE SSCI '19 conference. In addition, the source codes related to those publications are well tested, verified and documented, which are available upon requests.

WP3 – Big Data Analytics and Optimisation using Novel Machine Learning Approaches
NEC

Update on work packages:
All ESRs involved in WP3 have worked on the challenge research problems in the areas of big data analytics and optimization. Some of the achievements have been accepted for presentation at the selective conferences. For WP3.1, ESR5 worked on the topic of hyperparameter optimization in the context of class-imbalance data. Both classification algorithms and resampling techniques involve multiple hyperparameters that highly influence the prediction performance. ESR5 empirically analyzed the effect of a variety of optimization methods for the imbalanced classification problems with interesting results for the community.
In WP3.2, ESR3 focused on integrating observed structures of time series into representation learning for robust time series analysis. ESR3 worked on a novel machine learning method to combine Bayesian models with deep learning to capture the complicated dynamics of multiple correlated time series. For WP3.3, ESR8 has investigated the literature of probabilistic models and deep learning methods for text mining, as well as the recent advances on Bayesian inference. The current work focused on development of a deep generative model, which combines good properties of probabilistic topic models and neural networks, for learning interpretable representations of users and products from text labeled graphs, e.g. review networks. The human-understandable embedding provides insights for explainability of downstream learning tasks, such as product recommendation and product feature optimization. ESR8 also contributed to the work of multi-domain adaptive sentiment analysis to learn customer preference. For WP3.4, ESR7 addressed the challenging topic in dynamic multi-objective optimization DMO: the changing Pareto optimal solution sets. He empirically analyzed when and how transfer learning works or fails in DMO, and proposed an

improved method Tr-DMOEA with a linear kernel function to solve the problem. The initial results are promising and better than state-of-the-art methods.

Update on deliverables:
In total, there are three papers accepted for presentation at the IEEE Symposium Series on Computational Intelligence (IEEE SSCI 2019), and one paper accepted in the book of New Frontiers in Mining Complex Patterns published by Springer Series on Studies in Computational Intelligence. The related experimental results and source codes have been validated and documented.

WP4 - Core Knowledge and Transferable Skills Training
Birmingham

Update on work packages:
All training modules for the ECOLE project have been completed by students apart from PCST1, 4 & 6. The academic courses were the priority for completing as it gave the students the necessary knowledge to begin their PhDs, other courses will be achieved over the lifetime of the project. Cross-domain Knowledge training was suggested as a theme for the Birmingham Summer School so the students could learn about the application of natural computation to other fields e.g. neurobiology and behavior. However, due to the time of year, supervisor availability was not optimal to deliver a talk and alternatively, a talk on "Technology and Business Transfer" was included to demonstrate an alternative perspective on research.

Update on deliverables:
The training program as stated in the proposal is mostly completed with PCST courses noted for future. Skills report and career development reports have been completed.

WP5 - Project Coordination and Management
Birmingham

Update on work packages:
Recruitment has been completed. The first secondments occurred over the summer and all students are expected to have rotated by end of September 2019, (see section 4.1 for updated secondment plan). Visa issues and tax issues have been dealt with and information regarding secondments is now available to the ESRs. The Birmingham Summer School took place in July with invited talks from professionals outside of ECOLE including industry, academia and professional services. The next training workshop will be held on the 14th-15th October, hosted by HRI.

Update on Deliverables:
Consortium agreement, supervision policy are complete; project reports and project meetings are planned at defined intervals for the lifetime of the project.

WP6 – Communication and Dissemination
Birmingham

Update on Work packages:
The website is now completed as per the original requirements. Summer School took place in July with a wide ranging program including talks with academic and industrial themes in

addition to a talk on the translation of research and a hands on statistics workshop. Further, the ESRs developed their outreach and communication strategy with reps being selected to have oversight of contents on the channels. It was also decided that the website will hold a blog that each ESR will contribute to on a regular basis. We are preparing to do a public engagement event and the most appropriate one for the project will be decided on at the management meeting in October. A portfolio of publications are now being prepared (Section 2.1) and students are attending the IEEE conference on Data Mining and the Symposium Series on Computational Intelligence in December. During the Leiden and Birmingham meetings at the Birmingham workshop in May 2019, it was indicated that ESR 7's research may benefit from a placement in TATA. This is because TATA may have a suitable case study for dynamic transfer optimisation, which is ESR 7's research topic. We will further discuss this possibility during a meeting at TATA on 2nd October 2019.

Update on Deliverables:

Project website is completed but additional features (blog) to be added, the ESRs have a proactive role for the dissemination of project objectives firstly via social media then at engagement events. Summer School is complete with actions to follow up on and career events will be organized in future.

3. Thesis Plan and Career Development

3.1 Training Plan

The training activities presented in the proposal follow the Vitae Researcher Development Framework a summary of which can be found in Table 4. To date, the ESRs have completed the Academic Modules (AMs) and the Research Skill Training modules (RST). The Personal and Career Skill Training modules (PCST) are still outstanding however, these will be completed in future. The focus previously was to learn the tools to undertake the research and training in future will focus on applying what has been learned.

Table 4 - Training Plan

To date[2], the ESRs have completed roughly 50% of the training activities outlined in the proposal, see table 4 for those that have been completed (green) or to be completed (red).

Module	Learning Objective/description	ESR Number							
		1	2	3	4	5	6	7	8
AM1	Nature Inspired Search and Optimisation: A comprehensive introduction to the field of natural inspired optimisation, covering theories, algorithms and applications								
AM2	Intelligent Data Analysis: A comprehensive introduction to statistical pattern analysis, high-dimensional data mining, and text mining.								
AM3	Machine Learning: Advanced topics in machine learning, covering several forms of supervised, semi-supervised and unsupervised learning, in both theories and applications.								
AM4	Multiple-Criteria Optimisation and Decision Analysis: Theoretical foundations, algorithms, and application techniques of multi-objective optimisation								
AM5	Advances in Data Mining: Recent developments in data mining for classification, regression and clustering and beyond, dealing with massive data sets. Techniques for distributed data mining (e.g., Hadoop).								
AM6	Evolutionary Algorithms: State-of-the-art in evolutionary computation; including efficient optimisation techniques (i.e., small number of function evaluations)								

[2] September 2019

RST1	Research Skills: Including literature review, academic writing, referencing, LATEX, presentations, etc.							
RST2	Reading Groups: In reading group meetings, ESRs will present their work to the rest of the group on a rotating basis and strengthen their communication skills and confidence in presenting their work							
RST3	Research Seminars: In seminars, primarily given by external speakers, ESRs will network with leading professionals and extend their research insights.							
PCST1	Outreach training courses: Half-day course focusing on the effective communication of research to wider public and academic community							
PCST2	Time management course: Half-day course discusses planning techniques and resources for time management, helping students to prioritise to-do lists.							
PCST3	Speed reading workshop: Half-day workshop aiming at improving personal skills for more effective handling of reading material, master the techniques of scanning and skimming and improve the retention of written documents.							
PCST4	Talent pool: A unique five-day programme for ESRs, providing customised professional development training and access to practical, 'real-life' opportunities to develop the next generation of academics, business consultants and entrepreneurs							
PCST5	Enterprise training programme (Medici): A training programme with flexible time, helping ESRs realise the commercial potential of their research, covering business strategies, marketing, finance and business planning, funding opportunities, sale and negotiation skills and networking skills.							
PCST6	Open science programme: Focus on publishing and communicating scientific knowledge, particularly open access and open data strategies and techniques.							
PCST7	Intellectual Property Rights: Introduction to the formal aspects of intellectual properties, including guidelines on how to summarise research results in a way that they can be turned into powerful IPR by IPR professionals.							

PCST8	Technology Evaluation and Transfer: Introduction to the steps of turning research results into technology, technology into innovations and innovations into products and services with a sustainable business value.								
PCST9	Project Management in Research: Planning, Quality and Risks: Introduction to different standard methods for project management, including risk management.								

3.2 Outreach

The ESRs discussed plans for outreach at Summer School. The plan is to maximize content on existing channels with ESRs social media reps selected and dedicated to posting content on Twitter and Instagram. Twitter will have alternative week output. Instagram will contain a collection of pictures from everyday life of ESRs and publish when required. We are also developing a blog on the website in which it was agreed to have a blog post every two weeks written by an ESR, cycling through everyone. The audience that has been selected is public and specific scientific/industrial audience and content will be tailored as per. The ECOLE leaflets are finalized and will be circulated to appropriate audiences once they have been received.

BIS Conference – 26th June, Seville

The 22nd International Conference on Business Information Systems (BIS) took place in June. ESR5 presented her research about class imbalance in the "Doctoral Consortium" to an audience of PhD students and mentors from other institutions. She gave a short introduction of ECOLE to the audience.

Z2X Festival – 31st August, Berlin

ESR6 recently took part in the Z2X festival organised by the German newspaper "Die Zeit" and participated in a 10 minute talk delivered to young adults active in science, activism and arts. The aim of the event is to encourage exchange of interesting and novel ideas. ESR6 presented a talk on AI with reference to the ECOLE project. The talk titled "Nature Inspired Artificial Intelligence" was received well by the audience and encouraged questions from other participants. *"I have been invited by a group later in this evening who enjoyed my talk to join them for further discussion and beers. Overall I really enjoyed the experience and would likely do a popular talk again when given the opportunity."* ESR6

3.3 Career Development Plans

3.3.1 Thiago Rios

Efficient Representations for High-dimensional CAE Models in Learning-based Automotive Optimization

Thesis Objectives Include:

The objective of the research is to define a novel geometric representation for high-dimensional CAE (Computer Aided Engineering) models, which can be handled by multi-objective evolutionary optimization algorithms and enable the transfer of knowledge between optimization problems.

The research is divided in 4 main phases:

1. Survey of the literature about geometric representations (9 months): the main objective is to become familiarized with the state of the art representations for geometric models and select the best base representations for geometric deep learning tasks. The study is complemented by the review of deep learning architectures for 3D models and their preliminary implementation.
2. Dataset generation and development of surrogate models (9 months): during this phase a dataset of car shapes and corresponding aerodynamic performance is generated, which can be used for training the surrogate models that will be used during the research. Furthermore, the architectures reviewed during the first phase will be fully implemented according to the purposes of the research and the first experiments with the new data will be performed.
3. Analysis of surrogate models (9 months): the third phases concerns the layer-wise analysis of the deep learning models, with the purpose to identify the relation between the activation weights and the geometric features contained in the dataset. Furthermore, characteristics that enable the transfer of knowledge between optimization problems and a metric for measuring the similarity between problems are also part of the objectives for this stage.
4. Multi-objective optimization with transfer of experience (9 months): here, the methods reviewed and implemented in the previous phases are brought together in a framework of evolutionary multi-objective optimization. During this phase the performance of the optimization algorithms with the novel geometric representation and transfer of knowledge will be compared to state the art approaches, in order to verify advantages and limitations of the proposed method for optimization in the automotive domain.

Outreach Activities:

I will prepare, organize and participate in outreach events that come up during project runtime according to achieved project results and planned events in the proposal document. I will actively check for opportunities to present my research in potential events. Further plans for training include sessions on intellectual property rights, technology evaluation and transfer and time management training.

3.3.2 Sneha Saha

Multi-criteria Preference aware Design Optimization

Thesis Objectives Include:

Creating novel virtual or physical objects cooperatively as a team (man-man, man-machine) combines the perspectives of different actors in a development process. These perspectives are aligned with the different backgrounds and targets of each actor for her/his specific domain in this process, yet they altogether strive for a common goal defined by global product specifications. In the field of automotive development these domains include (besides many others) e.g. aesthetic design processes to generate novel shapes as well as engineering processes, here: aerodynamic and structural optimization, for performance compliance.

The expected out of the project is divided into 4 phase-

1. Expertise build up and simplified shape generation process. The deliverables will be written report/paper on assessed topic.
2. Recurrent neural network for learning shape generation. The deliverables will be written report/paper on assessed RNN and optimization software
3. Utilizing experience by Knowledge transfer. The deliverables will be written report/paper on assessed experience transfer and optimization software.

4. Preference aware shape optimization for multi criteria. The deliverables will be written report/paper/PhD-thesis on assessed preference aware optimization, user study and optimization software.

Outreach Activities:

I will prepare, organize and participate in outreach events, which come up during project runtime according to achieved project results and planned events in the proposal document. I will actively check for opportunities to present my research in potential events. Further plans for training include project management skills and software training including Python, Dash and Plotly.

3.3.3 Sibghat Ullah

Uncertainty handling and Robust design in learning-based optimisation

Thesis Objectives Include:

Comparison of uncertainty aware and robust Meta-Model algorithms with a focus on industry compliant indicators robustness assessment (Nov 2018-June 2019: LU)

1. Big data analytics through deep structured learning and representation learning with a focus on Robust Time Series Analysis (July 2019-Mar 2020: NEC)
2. Online learning of experience in dynamic and robust optimisation (Apr 2020-Dec 2020: HRI)
3. Challenging Issues in industrial robust and dynamic optimisation (Jan 2021-Nov 2021: LU)

Outreach Activities:

Plan to continue with the modules outlined in the proposal. Additionally publishing and sub-reviewing publications.

3.3.4 Duc Anh Nguyen

Automatic algorithm configuration for parameter tuning of modelling and optimisation algorithms

Thesis Objectives Include:

1. **Early-stage (Oct,2018 – Jul, 2019: Leiden University):**

- Review, problem analysis, and getting acquainted with the project. Review of Automatic algorithm configuration approaches.
- Literature study on Automatic algorithm configuration.
- Analysis of the project requirements on Automatic algorithm configuration.

2. **In industry (Aug, 2019 - Jun, 2021: NEC, Honda, TATA):**

- Development and test of algorithms.
- Adaptation of such algorithms to the specific requirements of the project.
- Assessment of different algorithms and variations thereof, including the development of new variants as required by the application domain.
- Development of a suitable integration into the overall software approach of ECOLE.
- Investigation of the scalability of algorithms and of their real-world applicability.
- Experimental tests with real-world applications at Honda.
- Evaluation of the results.
- Hyper-parameter tuning of the algorithms, further improvements.
- Comparison of hyper-parameter tuning approaches.

3. **Final-stage (from Jul, 2021: Leiden University):**

- Final results and integration into the ECOLE approach.

3.3.5 Jiawen Kong

Class imbalance classification through semi-supervised and active learning for experience-based optimisation

Thesis Objectives Include:

1. Hyperparameter Optimisation for improving classification under class imbalance – 1st year
2. Novel method to classify feasible and infeasible meshes based on the deformation parameters (feasible and infeasible mesh dataset is imbalanced) – 2nd year
3. Classify models according to thresholds of forces and momentum – 2nd year
4. Converged and not converged simulation. Normally the convergence analysis highly relies on the human experience, developing novel experienced-based approach to do the convergence analysis is the key objective of this past of research. – 3rd year

Outreach Activities:

I went to a workshop for PhD student in BIS conference and gave a presentation there to introduce not only my recent research but also the ECOLE project.

3.3.6 Stephen Friess

Learning in Evolutionary Search

Thesis Objectives Include:

Aim of my research is to develop new methods which realize a notion of learning in evolutionary search akin to transfer learning in machine learning. Of vital important for the success within the ECOLE project will be to benchmark them on state-of-the-art industry problems.

1. 07/19 –10/19: Development of a concept using a continuous GA as base.
2. 11/19 –03/20: Extension of the framework to modern algorithms and benchmarking on problems of shape optimization in the automotive industry.
3. 04/20 –12/20: Extension of the framework to consider problems in coordination with industrial partner NEC, potentially network optimization / neuroevolution.
4. 01/21 –09/21: Additional time spent in the industry.
5. 10/21 –09/22: Write-up and thesis submission.

Outreach Activities:

Plans for further training include software such as Python and skills training such as communication, dissemination which will be employed in the ECOLE research blog. I hope to also learn about patent law and intellectual property rights.

3.3.7 Gan Ruan

Transfer learning in evolutionary dynamic optimization

Thesis Objectives include:

1. Now – 15/8/2019: Figure out true reason(s) that transfer learning performs not very well on dynamic multi-objective optimization.

2. 16/8/2019 – 30/9/2019: Make improvements on transfer learning in dynamic multi-objective optimization.
3. 1/10/2019 – 31/12/2019: Try to find real-world optimization problems in Honda and solve them with transfer learning based optimization algorithms.
4. 1/1/2020 – 30/3/2020: Develop a transfer learning based dynamic algorithm to make it adaptive to dynamics.
5. 1/4/2020 – 30/6/2020: Try to find real-world optimization problems in TATA Steel and solve them with transfer learning based optimization algorithm.
6. 1/7/2020 – 30/9/2020: Try to apply online learning methods in dynamic optimization and explore whether there are some evolutionary algorithms which could be potentially used in online learning methods.
7. 1/10/2020– 31/12/2020: Try to find real-world optimization problems in NEC and solve them with online learning and transfer learning based optimization algorithms.
8. 1/1/2021 – 30/9/2021: Do further research with machine learning in optimization.

Outreach Activities:

I have been involved in the research presentations and discussions with other PhD students and postdocs of my supervisor in the school of computer science, University of Birmingham bi-weekly. In the future, I would like to participate some machine learning workshops or related activities that are beneficial for my research within the project. I am also planning on attending modules and seminars to improve my skills including communication, presentation, writing, entrepreneurship and so on. After taking them, some practical training and application of these skills are planned to strengthen them.

3.3.8 Giuseppe Serra

Explainable Deep Generative Models for Preference Learning on Products

Thesis Objectives Include:

Developing statistical machine learning and text mining methods for automatically detecting customers' opinions on aspects and features of products from large collections of unstructured data.

1. Dec 2018 – Aug 2019: Literature review about text mining methods for representation learning. Developing interpretable embedding for users and items using review data
2. Sept 2019 – Feb 2021: Research and development of deep generative models and Bayesian inference for explainability and uncertainty estimation in deep learning scenarios
3. Mar 2021 – Dec 2021: Research and investigation of explainable deep generative models for real applications with industry impact. Thesis and dissemination preparation

Outreach Activities:

Involved in website development for the LMID 2019 Workshop (dissemination activity of ECOLE). For the future: ECOLE blog managed along with the other ESRs

4 ECOLE Management

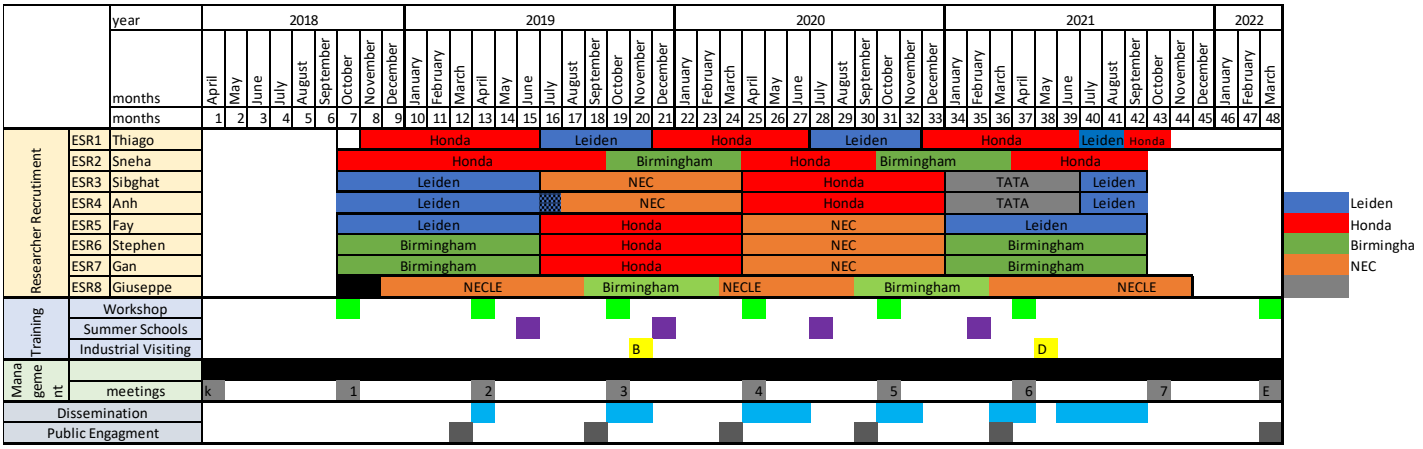
4.1 Updated secondment plan

Secondment plans have been amended for some of the ESRs with changes approved via the portal. Prior to updating the portal, all changes were related to beneficiaries and the consortium agreed for the changes to be approved. This is to ensure all secondment hosts are aware of the scheduling for timetabling and induction purposes. All changes have now been approved on the portal by the EU project officer.

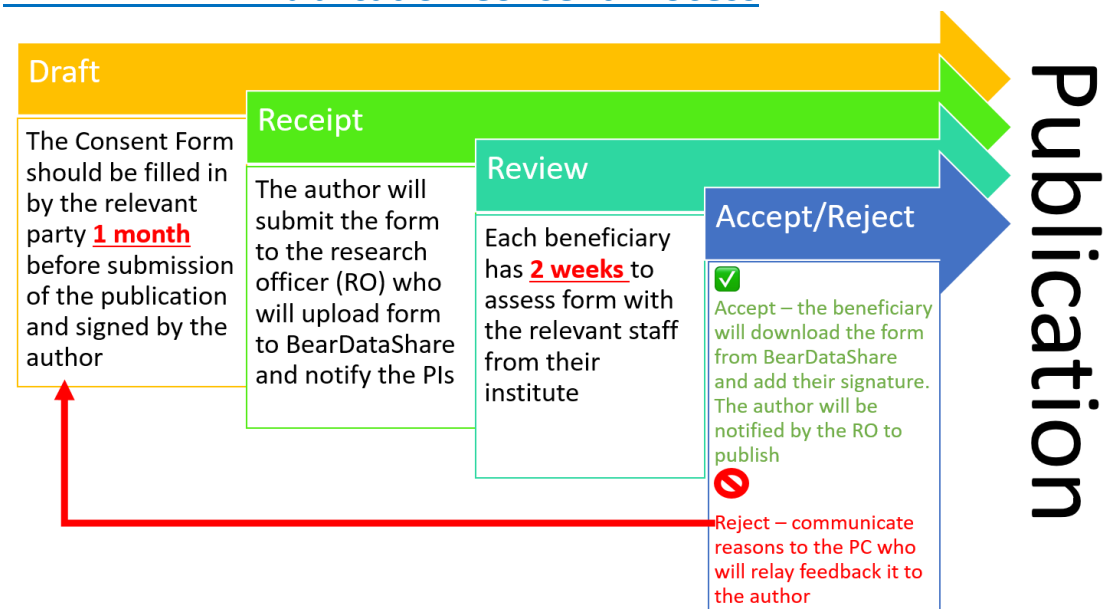
Reasons for amendments include:

Sneha Saha passed English test in June 2019 to allow completion of conditions for entry to a University of Birmingham PGR course. Due to the length of time taken for this to be re-sat and confirmed, admission to university was delayed as was application for visa. Duc Anh Nguyen travelled in July for his wedding, returned in August to begin the next secondment (1 month delay). For both Giuseppe Serra and Thiago Rios secondment plans have been altered to two visits of 6 months instead of one consecutive visit of 12 months.

Table 5 - Updated Secondment Plan



APPENDIX A – Publication Consent Process



Draft

The consent form template is available online and has been circulated to all members previously. The contents and criteria were confirmed at the Management Meeting in Birmingham in May 2019. The consent form will be drafted by the author of the paper and emailed to the Project Coordinator (PC).

Receipt

On the day of receipt, the Research Officer (RO) will upload the form onto BearDataShare and share the link with all PIs via a calendar invite/task stating the deadline for review. The folders will be labelled on BearDatashare as per the date they were received.

Review

Each PI has 2 weeks to discuss the submission internally and prepare feedback. PC will send reminders at 1 week from the deadline, any issues with reviewing should be communicated at this stage.

Accept

If there are no concerns regarding the submission, approval will be indicated on the spreadsheet on BearDataShare. If the agreement is unanimous, the Research Officer (RO) will notify the author after the deadline of the outcome and submission can proceed.

Reject

If a beneficiary requests that there is to be a revision, feedback should be sent to the RO who will collate feedback and share with the PC who will communicate revisions to the author. The author will amend the form if necessary and the process will be started again.

Manuscripts

After approval is granted, manuscripts should be prepared. Finished manuscripts should be sent to the RO who will circulate to the PIs to view. This should **be no less than 2 weeks** before final submission to the journal.

APPENDIX B – Publication Consent Form

The process for publication consent has been circulated previously and a copy can be found on BearDataShare, please familiarize yourself with this. This form should be filled in by the corresponding author of an intended publication and signed by the PIs of all ECOLE beneficiaries. Please note that the form should be **completed at least 4 weeks** prior to the intended submission date, otherwise consent cannot be guaranteed. Completed manuscripts should be shared **no less than 2 weeks** prior to submission deadline.

SECTION 1 – PUBLICATION INFORMATION

Submission Details:

Beneficiary submitting this form:	Choose an **item**.
Intended date of submission:	
Event: (Place/time)	

Publication Details:

Authors and their affiliation:
Title of intended paper:
Draft Abstract:

Patent Details[3]:

Intention to file patent: (select answer then fill out relevant sections of this table)	Choose an item.
If answered **Yes** to previous: What is the novel part/idea/method you would like to apply for?	
If answered **No** to previous: Please state reason for not filing IP	

[3] ***For further clarification on ownership, please see Section 3 "Rights and Obligations related to Background and Results" p29-39 of the ECOLE Grant Agreement***

If answered **No**: Please state which algorithm/methods/benchmark functions are used for this publication and if these algorithms/methods/ benchmark functions are known in the literature (provide reference if known)?
If answered **No**: Please provide any other relevant information including any appropriate references

Acknowledgement:
"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 766186".

Declaration:
We/I, the authors, have not used any unauthorized materials from any other beneficiaries and partners in this paper

Signature: (by the corresponding author) _____

SECTION 2 – BENEFICIARIES CONSENT

We agree to the submission and publication of this paper:

Signatures: (PIs of all other beneficiaries)

• Choose an **item**.
• Choose an **item**.
• Choose an **item**.
• Choose an **item**.