Development of Machine Learning algorithms to improve the optimization of recurrent combinatorial problems

Contact: Emmanuel Rachelson
Département d'Ingénierie des Systèmes Complexes
emmanuel.rachelson@isae-supaero.fr

Topic

Recurrent combinatorial optimization problems — ie. problems that require the resolution of several successive instances — are a major industrial challenge. For instance, matching airline crews and flights for a given day, electricity production planning, re-scheduling of satellite imaging are highly combinatorial problems where the problem for day D and the one for day D+1 are variants of the same initial problem and where finding a (quasi-)optimal solution is important. Often, the re-optimization process is time-constrained, leading to an optimality versus computation time compromise.

This research internship falls within the "Learning4Opt" project at ISAE-SUPAERO and will be tutored by E. Rachelson (ISAE-SUPAERO professor) and L. Mossina (PhD student).

The question we try to answer in the Learning4Opt project is the following. Given a series of optimization problems that are related to each other and need to be solved in sequence, how can one use the experience from solving the previous N problems to improve the resolution of the N+1th one?

In practice, we focus on recurrent combinatorial optimization problems such as:

- Unit commitment in energy production

- Project planning

- Assignment problems

- Satellite planning problems

- and standard other Operations Research benchmarks.

Each of these problems has a unique description as a combinatorial optimization task. A sequence of problems P_0 ... P_n is then defined by varying the data (each instance in the sequence has different values for the input data, e.g. the daily power demand in the first example above).

Recent advances in our team have introduced a new method called NaiBX for the problem of multi-label classification, with reduced computational cost. This method was inpired by the problem of selecting of a subset of variables among all the optimization variables, given the knowledge of the varying data. This has lead to the development of the NaiBX general purpose C++ library for multi-label classification. Applied to recurrent combinatorial problems, NaiBX serves to shrink down the size of new optimization problems given the solutions to past problems. Given some training data obtained from solving P_0... P_n, NaiBX learns how to predict a subset of variables in P_{n+1} that can be frozen to a heuristic value, thus defining a reduced problem, easier to solve. This method of Multi-Label Classification has been studied, developed and implemented and is readily available in our lab.

Different tasks will be available for the intern and will be chosen in agreement with the tutors. These tasks are:

- *Improvement of NaiBX*. Several improvements are required for NaiBX. The C++ library currently only allows to model variables that are Normally distributed or Bag-of-Words features. A first topic within this task will be to extend the library to a variety of other parametric distributions. Another topic deals with the inclusion of a model selection criterion

to allow for the automated selection of these distributions, per variable. An important third topic consists in extending NaiBX with an commitee-learning approach in order to improve its overall prediction performance.

- *Empirical evaluation*. Two directions are desirable for the empirical evaluation of the application of NaiBX to optimization problems. First, no extensive evaluation of the coupling between NaiBX and a variety of different optimization problems has been done so far. So the first topic in this task consists in evaluating how NaiBX performs on those problems. In particular: how much speed-up we can expect on each type of problem, how much optimality loss, what type of problems benefit more than others from NaiBX, what symmetry or quasi-symmetry properties can NaiBX improve, etc. The second topic deals specifically with the unit-commitment problem and is an attempt at finely evaluating the benefits of the application of NaiBX to real-world problems (in particular with the question of scaling up to large production networks).

- *New research leads*. The proposition of a formal framework and of new algorithms for the resolution of recurrent combinatorial problems is an exciting perspective and new contributions are welcome. Some starting ideas are already available in the team.

Student profile

Master of Science (or equivalent) in Machine Learning, Operations Research or a related discipline.

- Basic Probability theory and Statistics,

- Curiosity for Machine Learning algorithms and Combinatorial Optimization,

- Some confidence in programming (C++ would be a plus).

- Good level in English (the working languages in our lab are English and French).

Context

ISAE-SUPAERO is a worldwide reference in aerospace higher education programs and research for aerospace engineering. Located in the lively city of Toulouse, European capital for Aeronautics and Space, it is nested within the university scientific campus in a rich research environment. Within the Department of Complex Systems Engineering, the Decision Systems group works on innovative resolution methods in artificial intelligence and operations research and applies them to different problems in industrial engineering and robotics (for instance).

Toulouse, hosting the Airbus headquarters, is known worldwide as one of the aeronautics industry's capitals. Often ranked among the best living places for students, Toulouse is a lively city with a strong scientific environment.

This intership is funded by the "Programme Gaspard Monge pour l'Optimisation (PGMO)", in collaboration with EDF R&D.

Salary: 554.40€/month (net), housing opportunities on campus.

Duration: 6 months.

Informations and applications: emmanuel.rachelson@isae-supaero.fr.