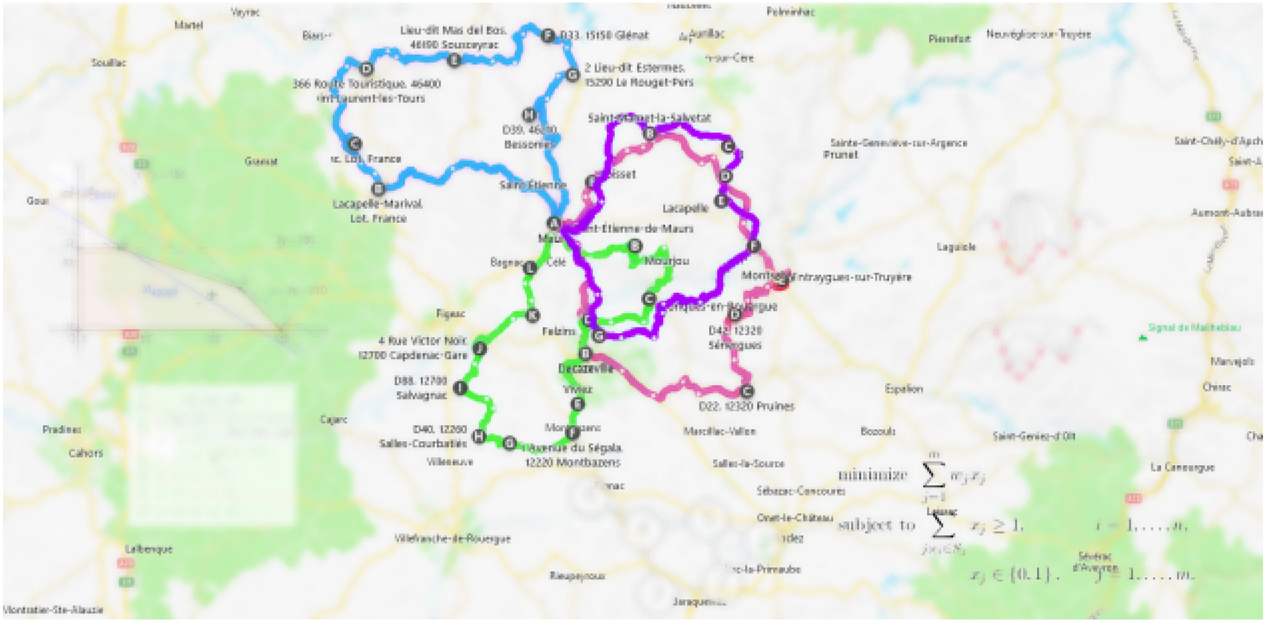
Advanced Algorithms **Preparation** Project **1 - PBL Loop** **2 - PBL Loop** **3 - PBL Loop** **4 - PBL Loop**

Learning objectives | Resources for students | Resources for Tutors | Tout afficher | Tout réduire

INTRODUCTION



Since the 90s, there has been a truly global awareness of the need to reduce energy consumption and greenhouse gas emissions. The first commitments emerged when the Kyoto Protocol was signed in 1997. However, this protocol only entered into force in 2005, and many scientists believed that the efforts to slow down global warming were not enough. Since then, other more ambitious commitments have seen the light of day (France's commitment to a 75 % reduction in emissions by 2050, for example, commitments made by certain large cities such as Paris). But the task is complicated. The government and local authorities are unable to force companies and individuals to change their habits in order to meet these goals. Therefore, action is primarily focused on changing behaviour. Saving and recycling raw materials, improving means of transport and the energy performance of buildings should become priorities.

TOPIC

Instructions

QUIZ

ADEME (French Environment and Energy Management Agency) has recently launched a call for expressions of interest to promote the execution of demos and experiments of new mobility solutions for people and goods, adapted to different kinds of territories.

Your CesiCDP structure is already well established in the field. With the help of many partners, you have carried out several studies on Smart Multimodal Mobility. New transport technologies, despite being more cost-effective and cleaner, You are part of the team set up by CesiCDP to answer the call from ADEME. The challenge is to win new markets with very attractive financing schemes to keep developing your business activity.

CesiCDP decided to focus its study on the management of delivery routes. The algorithmic problem consists in calculating on a road network a route allowing to connect a subset of cities between them, then to return to its starting point, so as to minimize the total duration of the route. This optimization should take into account the expected traffic on each axis for the different time slots.

The idea is to propose a method from Operations Research to generate a delivery route corresponding to this problem.

Even though the scope is yet to be defined, you have outlined a basic version of the problem. But you hesitate in adding additional constraints in order to make it more realistic and get ADEME's full attention. One should expect the problem to be harder to address like this.

- ***- Basic version***
  - Selecting the model and code in Python capable of solving large instances (several thousand cities)
  - A statistical study regarding the algorithm's experimental performance

- ***Additional constraints***

  Here is a (non-comprehensive) list of constraints that could be incorporated into your study scope. For some of them, advanced versions are also provided. Please note that implementing one constraint and any of its advanced versions is the same as implementing two constraints.

  - Delivery time slot for each item
    - No deliveries allowed beyond the time slot
    - Waiting on site for the time slot to open is a possibility
  - *k* trucks simultaneously available to make deliveries. The route calculation will have to include the allocation of items (and therefore the delivery points) to the different trucks available, and instead of minimising the total time, one should minimize the date when the last truck returns to base.
    - Truck capacity (2 or 3 dimensions) and item footprint
    - Some items can only be delivered by certain trucks
  - Each item has a specific collection point
  - The travel time of an edge varies over time (which is equivalent to varying its length), to represent the variation in traffic flow

  These constraints can be divided into two categories:

  - Constraints that do not change the space for solutions, only their values. For example, by taking into account time slots where waiting is required (if the truck is ahead of schedule). In the case of a neighbourhood-based method, the neighbours of a solution will not be changed by incorporating this constraint, only the costs will be different
  - Constraints that change the space for solutions. For example, some items that require a specific type of truck for being delivered. Adding this constraint will render some solutions invalid, and consequently limit the space for solutions.

  Other constraints can be addressed, if this is justified by an industrial application (not necessarily resulting from the current context).

The deadline for submitting projects to ADEME is set [see date with your tutor]. In order to meet deadlines and prepare all the necessary reports, you have defined the following roadmap:

1. Formal modelling
2. Algorithmic design and implementation

Each of these steps will result in a deliverable that you will present to your manager for follow-up and discussion. These deliverables will take the form of a Notebook presenting both the approach and the corresponding code (when an implementation is required). Preference should be given to the story-telling aspect when writing each of these Notebooks.

The code in the different deliverables does not have to be according to standards of a code to be mass-produced (no need for great scalability). On the other hand, it must remain readable, commented, and favour performance. Following the PEP recommendations is strongly recommended:

https://openclassrooms.com/fr/courses/235344-apprenez-a-programmer-en-python/235263-de-bonnes-pratiques - FR

Expected deliverables

- ***Livrable Modélisation (check)***
    - Deadline: [see date with your tutor]
    - Assessment: No
    - Goal: Modelling the problem using a Jupyter Notebook
    - Description:

      This first deliverable presents the problem you are addressing as well as its context, rephrases the problem in a formal way, and studies its theoretical properties, especially complexity ones. On the one hand, you must propose a formal representation of the data, problem and goal to be optimised (remember to incorporate the additional constraints that you decide to address). On the other hand, you must rely on this formal representation to demonstrate the theoretical complexity of the problem in question. It is strongly recommended to include bibliographic references of scientific research articles or books in this Notebook.

      The problem-solving methods are not to be covered during this check phase. You will have the opportunity to study them in detail in the following loops.

      The status is discussed with the project tutor, allowing you to properly validate your modelling prior to implementation. Items may be subject to change before being incorporated into the final project deliverable.

- ***- Final project deliverable***
    - Deadline: see course schedule
    - Assessment: Yes
    - Goal: Presenting the entire approach adopted, the technical execution and drawing a conclusion on the achieved results
    - Description :

      This deliverable is divided into two parts:

        - **PART 1: Modelling**
            - Includes updated formal modelling elements
            - Describes the problem-solving method chosen: details on the algorithm used, i.e. the metaheuristic technique chosen and the problem modelling according to the formality of this metaheuristic technique (neighbourhood, crossover operations…)
        - **PART 2: Implementation and operation**
            - Implementation :

              This algorithm's implementation

              A demonstration of how this implementation works, on different test cases (no need to go into detail)
            - Operation: Experimental study

your analysis.

The methodology used in this experimental design, the analyses that you will propose based on the statistical results achieved, and the proposals for improvement that you will derive from them, should be justified in detail.

- ***Presentation***
  - Deadline: see course schedule
  - Assessment: Yes
  - Goal: Demonstrating the ability in English to orally present a work in a professional context
  - Description :

    The presentation should be result-oriented, with a demonstration of how the code runs (on cases that are small enough to keep the presentation fluid) and a presentation of results. As the presentation is in English, it is likely that at least one of the members of the jury is not an IT specialist. It is therefore recommended to pay special attention to the popularisation aspect, while maintaining the accuracy and scientific rigour whenever possible.

    The following points should be addressed in the presentation:

    - Presentation of the actual problem with the constraints under consideration, and industrial challenges
    - Concise and popularised presentation of the problem-solving approach
    - Technical demonstration on instances that are small enough to keep the presentation fluid
    - Popularised presentation of the statistical study, focusing on the most relevant result

    The oral presentation is done in English and should allow the following items to be assessed:

  -

GRAMMATICAL ACCURACY	[B2] Shows a relatively high degree of grammatical control. Does not make mistakes which lead to misunderstanding.
SPOKEN FLUENCY	[B2] Can produce stretches of language with a fairly even tempo; although he/she can be hesitant as he/she searches for patterns and expressions, there are few noticeably long pauses. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without imposing strain on either party.
PHONOLOGICAL CONTROL	[B1] Pronunciation is clearly intelligible even if a foreign accent is sometimes evident and occasional mispronunciations occur.

Assessment

Terms

- Project assessment (final deliverable) - Macro chart:

A	Functional program with theoretical study of complexity and statistical study
B	Functional program without theoretical study and with statistical study
C	Functional program without studies
D	Program not working

Please note: an A cannot be granted if instances are too small, and a B cannot be granted if there is no statistical study

🕶 EDUCATIONAL GUIDE

Tutoring help

Problem feasibility

point (see further below).

Objective function

The basic version of the problem, which is the TSP, only considers one vehicle. Once several vehicles are involved, the objective must be redefined. This is not done in the subject because students need to understand for themselves that minimising the total distance is no longer the right indicator, and that they have to redefine the objective function they want to consider (and justify this in their various renderings).

Depending on the constraints under consideration, one might be interested in minimising the delivery time, minimising the number of vehicles used for delivering over a given period of time, minimising deviations from delivery time slots, or even minimising a complex fitness function (a rather complicated case), etc. All these variants make sense from an industrial point of view, but it is important to ensure that students are confident about their choices.

In particular, if students are relying on the scientific literature, or using data sets, they must check that the resources are indeed dealing with the same objective as them.

Complexity

The project statement mentions the graph length and the number of customers requiring delivery as parameters of the problem complexity. In fact, only the second parameter is relevant. The path to be considered between two consecutive customers is the shortest path between the two delivery points. The complexity lies in the order in which customers are visited, and in the allocation to a truck if this constraint is considered.

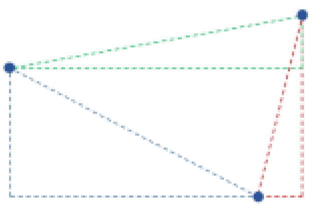
The NP-Completeness proof of the decision problem associated with the TSP on an incomplete graph will be discussed in the last exercise of the algorithmic complexity Workshop. The correction of this exercise is the proof to be provided in the first deliverable. It's not necessary to tell them this, but if students see it for themselves, that's even better. On the other hand, they must explain how the fact of incorporating additional constraints preserves the NP-Completeness of the problem (no need for a polynomial-time reduction, a simple reasoning on the solution space is enough).

Instance generation

The optimization problem considers an incomplete graph. To address this case, the approach that students are likely to consider, **but which should be avoided**, consists in incorporating, in the cost calculation, the transition time between two cities, calculated by shortest path (which is equivalent to completing the graph with edges weighted by these shortest path lengths).

The **preferred** approach is to consider that the shortest path between two cities problem is a simple problem (Dijsktra's or Bellman-Ford algorithm, depending on the context), and that the only important thing is to choose the order in which the cities are visited (and the allocations to the vehicles, in the case of the VRP). In this case, an industrialised solution should incorporate this shortest path calculation, but the goal here is to propose an optimization algorithm prototype, and to study its behaviour.

To do this, simply considering a complete graph that meets the triangle inequalities is enough. Therefore, students will have to generate such a graph in order to assess their algorithm's behaviour on this type of graph. The most common approach is simply to randomly generate spatial coordinates (x, y), and use the Pythagorean theorem to calculate the distance between each pair of vertices:



However, students must justify this choice, by providing an explanation like the one shown here.

Work distribution

Given that the workload is substantial, it is highly recommended that students split the work. For example, this distribution can be done as follows:

- The problem's data structure

- The generation of random instance sets

The recommended group length is 4 to 5 people. If this is not possible, it is better to favour larger groups, even if it means being a little more demanding on the renderings. A group of 3 students is less likely to finish the project on time.

Standard corrections

- - Modelling Deliverable (Check)

  Grille de check du livrable Modélisation - EN [xls]

- - Check phase chart for the implementation part

  This chart can be used to take stock with the groups on the implementation part (during the execution phase of loop 4 or at the end of prosit outcome and feedback)

  It comprises 2 parts, so it can be used once or twice.

  Grille de check de la partie Implémentation - EN [xls]

- - Final project deliverable

  Grille d'évaluation du livrable final - EN [xls]

- Presentation

  Grille d'évaluation anglais [xls]