GEP - Lliurable 2: Temporal Planning

Tomàs Osarte Segura

March 2024

1 Introduction

This document is designed to outline a detailed timeline for the completion of the graduation thesis titled "Search & Analysis for New Heuristics for Solving NP-Hard Problems with Deep Reinforcement Learning." The project commenced on February 15, 2024, and is scheduled for completion by June 29, 2024, culminating in the presentation and defense of the thesis.

2 Tasks descriptions

In this section, we will comprehensively detail the individual tasks, systematically grouped to clearly delineate the distinct phases of the project. This structured approach ensures each phase is distinctly defined, allowing for a more organized and efficient workflow. By categorizing tasks in this manner, we aim to provide clarity on the progression and dependencies of the project's various components. This organization not only facilitates a better understanding of each phase but also aids in efficient project management.

2.1 Study (ST)

These tasks pertain to activities that necessitate in-depth study or research on a specific topic. For this set of tasks it's only required to have internet connection and a device to search all the necessary information.

- **DRL applied to NP-HARD problems (ST1):** This task involves a comprehensive review and acquisition of knowledge about cutting-edge techniques in deep reinforcement learning that are currently employed to tackle NP-hard problems. Given the significance of this phase in the development of new heuristics, we have allocated approximately 45 hours to this task. This duration reflects the critical importance of understanding the state-of-the-art in DRL applications for NP-hard problems, ensuring a robust foundation for the project's subsequent phases.
- **Explainability for Deep Reinforcement Learning (ST2):** This task plays a crucial role in offering insights and understanding of the heuristics developed through our project. While the primary aim is not to devise new techniques for explainability, it is still vital to ensure that the heuristics are interpretable and their workings are transparent. Consequently, we have estimated that this task will require approximately 25 hours. This time allocation is designed to adequately address the explainability aspect of our heuristics, balancing the need for clarity without diverting significant resources from the core objectives of the project.

2.2 Development (D)

These tasks are essential components of the thesis development phase and form a crucial segment of the project's overall activities. Given the intensive computational demands typically associated with Deep Reinforcement Learning (DRL) frameworks, these tasks are likely to require substantial computational resources to yield effective results within a reasonable timeframe. Therefore, securing the most powerful computing device available is a strategic priority, ensuring the project's computational needs are met efficiently and effectively. This approach is aimed at optimizing the performance and outcome of the research.

- **DRL frameworks for solving single instances (D1):** The development of these frameworks is crucial for evaluating and analyzing which framework is most effective for the selected problems. Recognizing the importance of this phase, we plan to invest a significant amount of time in testing various frameworks. This comprehensive approach will provide a diverse range of options for the subsequent generalization phase. Accordingly, we have allocated 45 hours to this phase, ensuring thorough exploration and assessment of different frameworks to identify the most suitable ones for our project needs. The time provided do not take into account the training phase since is very unpredictable and almost do not require human attention.
- **Selected DRL frameworks for generalization of the problem (D2):** The phase following D1 is likely to be one of the most critical stages of the project, as it involves the development of a new heuristic with generalization capabilities, which represents the one of ultimate goals of this research. Given the significance of this task, we plan to dedicate approximately 45 hours to it, ensuring enough time for the meticulous development and refinement of a heuristic that not only addresses the specific problems at hand but also demonstrates a broader applicability and effectiveness. The time provided do not take into account the training phase since is very unpredictable and almost do not require human attention.
- **Explainability framework (D3):** This task, while necessitating meticulous development, is perceived as less demanding compared to the other two development tasks, primarily because innovation is not a central component in this stage. Based on this rationale, we have decided to allocate approximately 25 hours to this task. This time frame is carefully considered to ensure sufficient attention to detail and quality, while acknowledging that the primary focus here is on execution rather than innovation.

2.3 Evaluation (E)

This set of tasks focuses on evaluating developments generated earlier in the project. They will likely require benchmarks and computational resources, though not as extensively as in phases D1 or D2.

- **Single instance frameworks (E1):** The evaluation of the single instance framework will involve a comparative analysis of various statistics to determine the most suitable framework for the task. As this phase is not particularly time-intensive, we anticipate completing it in under 15 hours.
- **Generalization frameworks (E2):** The evaluation of the generalization framework will encompass a comparative analysis using state-of-the-art metrics and results initially explored in ST1. We expect this task to be less time-intensive compared to others and anticipate its completion in under 15 hours.

- **Explainability framework (E3):** The evaluation of the explainability framework presents a challenge due to the less obvious nature of its metrics. However, we do not intend to allocate extensive time to this task. It is anticipated that a thorough evaluation can be efficiently completed in under 10 hours.

2.4 Selection (SE)

This set of tasks are the ones that are considered to be selections among various options. This tasks do not require of any extra resource.

- **Single instance (SE1):** Selecting the optimal frameworks for solving single instances, NP-Hard problems is a crucial task that heavily relies on the outcomes of D1 and E1. Provided the preceding work is executed accurately, this selection process should not require more than 5 hours.

- **Generalization (SE2):** Selecting the optimal framework, if necessary, for solving NP-Hard problems could be necessary if more than one framework is developed. This task mostly relies on the outcome of D2 and E2 and given their results is should not take more than 5 hours to select this best heuristic.

2.5 Analysis (A)

This group of tasks is responsible for analyzing and synthesizing ideas from various topics. Being primarily a reasoning process, it requires no resources other than the documentation generated by related tasks.

- **Heuristic (A1):** The analysis of the newly developed heuristic is a vital task, closely related to E2. It demands a deep and thorough understanding of the heuristic's intricacies. We estimate that this task will take approximately 5 hours to complete.

- **Explainability (A2):** The task of analyzing the explainability of the heuristic also presents a unique challenge, as it lacks clear metrics and requires innovative thinking. We estimate that this task will require approximately 5 hours to complete.

2.6 Follow up (F)

These tasks involve the essential follow-up activities for the project, ensuring thorough oversight and adherence to the schedule. The only requirement for these tasks is access to an internet connection or a suitable location for meetings.

- **Meetings (F1):** Follow-up meetings are scheduled weekly throughout the project's duration and are crucial for reviewing recent achievements and proposing new subobjectives. Each meeting is expected to last approximately one hour to have a total amount of dedicated time more or less of 25 hours.

- **Correction sessions (F2):** These sessions, particularly during the development phases, are aimed at ensuring the proper development of the respective frameworks, which can be challenging to construct. Given the emphasis on achieving the highest quality for the thesis, we estimate that these sessions will cumulatively amount to approximately 15 hours in total.

2.7 Documentation (DC)

These tasks are focused on documentation, requiring only a computer to compile and prepare the necessary documentation.

- **Project Management (DC1):** To maintain an organized approach, the entire project management process is meticulously documented. This ensures a clear understanding and establishment of procedures, objectives, scope, and other critical concepts vital for achieving positive outcomes and also serves as a staring point for the D2. Documenting these details is a time-intensive process, and we estimate it will take approximately 40 hours to complete. This task can be effectively broken down into several sub-tasks to provide a clearer understanding of the various components involved in the project management documentation:

  - **Contextualization & scope (DC1.1):**This sub-task is dedicated to providing a comprehensive overview of the project, including its objectives and background. The aim is to clearly define the project's purpose and its expected outcomes. This portion is anticipated to take approximately 20 hours, reflecting the depth of detail required to accurately capture the project's essence.
  - **Temporal planning (DC1.2):** This involves creating a detailed plan that outlines the scope and timeline of each task, along with the resources required for their completion. The objective is to establish a clear roadmap of the project's workflow, ensuring efficient time management and resource allocation. This part is estimated to take around 10 hours, signifying its importance in project management.
  - **Budget and sustainability analysis (DC1.3):** This sub-task focuses on assessing the financial aspects of the project, including budget allocation and cost management. Additionally, an analysis of the project's sustainability is conducted to understand its long-term viability and environmental impact. This analysis is crucial for ensuring the project's feasibility and is estimated to be completed in approximately 10 hours.

- **Memory (DC2):** This graduation thesis is grounded in an academic perspective, signifying that the project documentation holds substantial importance and impact. Acknowledging this, it is essential to emphasize that effective documentation is a meticulous and time-consuming task. It involves not only the recording of data and findings but also a comprehensive articulation of the project's objectives, methodology, theoretical framework, analysis, and conclusions.Furthermore, the documentation must be coherent, well-structured, and adhere to academic standards, ensuring it effectively communicates the research process and insights gained. Given the depth and breadth required in this documentation, we estimate that this process will require approximately 45 hours.

Group	**Total Hours: 400**			
ID	**Task**	**Hours**	**Resources**	**Role**
Study (ST) - Total Hours: 70				
ST1	DRL applied to NP-HARD problems	45	Internet device	Researcher
ST2	Explainability for DRL	25	Internet device	Researcher
Development (D) - Total Hours: 115				
D1	DRL frameworks for single instances	45	Computational Resources	Developer
D2	DRL frameworks for generalization	45	Computational Resources	Developer
D3	Explainability framework	25	Computational Resources	Developer
Evaluation (E) - Total Hours: 40				
E1	Evaluation of single instance frameworks	15	Benchmarks, Computational Resources	Evaluator
E2	Evaluation of generalization frameworks	15	Benchmarks, Computational Resources	Evaluator
E3	Evaluation of explainability framework	5	Benchmarks	Evaluator
Selection (SE) - Total Hours: 10				
SE1	Selection of frameworks for single instances	5	-	Decision Maker
SE2	Selection of generalization frameworks	5	-	Decision Maker
Analysis (A) - Total Hours: 10				
A1	Analysis of the heuristic	5	Documentation	Analyst
A2	Analysis of explainability	5	Documentation	Analyst
Follow up (F) - Total Hours: 40				
F1	Follow-up meetings	25	Internet, Meeting Space	Project Manager
F2	Correction sessions	15	Internet, Meeting Space	Quality Assurance
Documentation (DC) - Total Hours: 85				
DC1.1	Contextualization & Scope	20	Computer	Documenter
DC1.2	Temporal Planning	10	Computer	Documenter
DC1.3	Budget and Sustainability Analysis	10	Computer	Financial Analyst
DC2	Project Documentation	45	Computer	Documenter

Table 1: Project Tasks Overview

3 PERT & Gantt charts

To offer a clearer perspective on the workflow and timing of the thesis, we present a PERT chart for a comprehensive visualization of task dependencies and a Gantt chart to track task progression over time. The PERT chart exposes the interdependencies of tasks, providing a roadmap for project execution, while the Gantt chart delineates the temporal development of tasks, articulated on a weekly basis. These tools are instrumental in streamlining project management and are referenced in Figures 1 and 2, respectively.

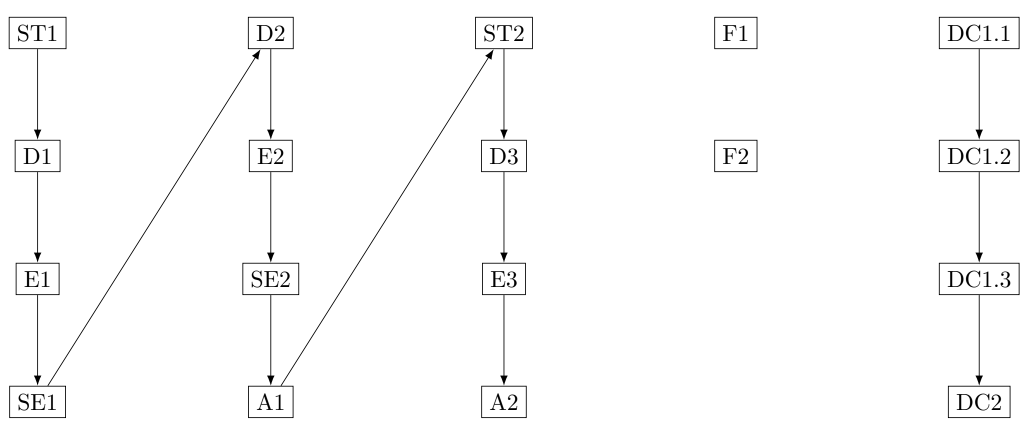

Figure 1: Task Dependency Chart

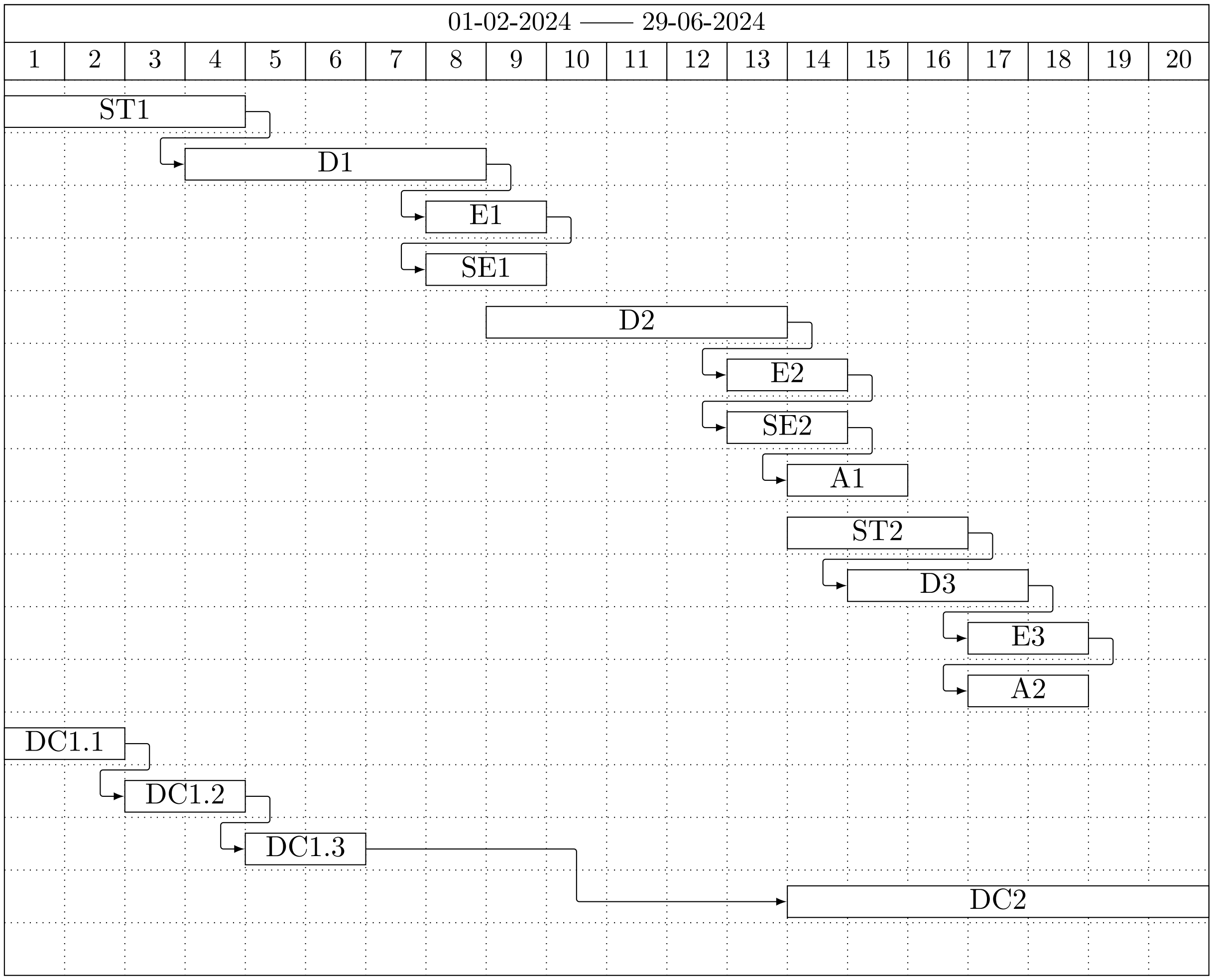

Figure 2: Task Dependency Chart

Risk management: Challenges and Mitigation Strategies in Reinforcement Learning Frameworks

Reinforcement learning (RL) frameworks are powerful tools but can face several challenges and risks during development and deployment. Since this is the main source of problems that we can face in this thesis, here is provided a list of possible obstacles and how will they be tackled:

- **Sample Efficiency:** RL algorithms often require numerous samples to learn effective policies. This inefficiency can extend training durations and limit real-world scalability.

  Mitigation Strategies:

  - Utilize experience replay to recycle past interactions.
  - Prioritize experiences by significance.
  - Employ model-based methods to harness environment dynamics.
  - Adopt transfer learning from related tasks.

- **Exploration-Exploitation Tradeoff:** RL must balance the act of exploring new actions with exploiting known strategies to maximize rewards, which is a nuanced challenge.

  Mitigation Techniques:

  - Implement epsilon-greedy or softmax policies.
  - Use Upper Confidence Bound (UCB) methods for action prioritization.
  - Engage in Thompson sampling for action selection.
  - Adapt multi-armed bandit algorithms for decisions.

- **Generalization:** RL frameworks may not generalize effectively to new environments or tasks, affecting policy transferability.

  Strategies to Improve Generalization:

  - Ensure training on diverse datasets.
  - Apply regularization techniques.
  - Use transfer learning for broader applicability.
  - Employ ensemble methods and domain randomization.

- **Training Stability:** The RL training process can be volatile, particularly with deep neural networks, leading to issues like vanishing or exploding gradients.

  Mitigation Strategies:

  - Utilize gradient clipping to prevent gradient issues.
  - Apply batch normalization for stable training.
  - Implement learning rate scheduling for better convergence.
  - Use suitable weight initialization and experience replay.

- **Resource Constraints:** RL frameworks demand substantial computational power and datasets, which can be prohibitive in resource-limited settings.

  Overcoming Resource Constraints:

  - The solutions here often involve significant investment in computational infrastructure.

The challenges associated with Reinforcement Learning (RL) frameworks have been considered and are reflected in the time allocated to each task within the project. Concerning resource constraints, we are proactively exploring solutions to acquire powerful computational resources at minimal cost. Our strategy includes seeking partnerships with entities equipped with high-performance computing facilities. One potential collaborator could be the Barcelona Supercomputing Center, which could offer access to advanced computational capabilities. By leveraging such partnerships, we aim to mitigate the limitations imposed by resource constraints, thereby ensuring the scalability and success of our RL frameworks without incurring excessive expenses. This forward-thinking approach is designed to align with our commitment to efficiency and fiscal responsibility throughout the development of the project.