PHD PROPOSAL IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Trustworthy Planning & Scheduling

Advisors:
Sylvie Thiébaux (Sylvie.Thiebaux@gmail.com)
Emmanuel Hebrard (hebrard@laas.fr)
Joao Marques-Silva (joao.marques-silva@irit.fr)

Gross salary: 2402€ per month

Duration: 36 months, October 2022 start
Application deadline: September 4 2022, then until filled

DESCRIPTION

Two PhD positions are available in the framework of the **Horizon EU Project TUPLES**: « Trustworthy Planning and Scheduling with Learning and Explanations. »

Automated planning and scheduling is a core area of AI. Its aim is to build systems that assist humans in choosing, organising and optimising courses of actions to achieve complex objectives. There is a pressing need for such systems in industry and public services, yet current planning and scheduling approaches are not sufficiently trusted to be widely adopted, as they do not satisfy essential properties of trustworthy AI systems such as transparency, explanability, robustness, safety and scalability.

TUPLES is a new 3 year Horizon EU Research & Innovation Action starting in October 2022, aiming to develop **scalable, yet transparent, robust and safe algorithmic solutions for planning and scheduling**. The cornerstones of TUPLES' scientific contributions will be

1) **hybrid approaches combining symbolic Planning and scheduling methods with data-driven methods** to benefit from the scalability and modelling power of the latter, while gaining the transparency, robustness, and safety of the former and
2) **rigorous explanations and verification approaches** for ensuring the transparency, robustness, and safety of plans and schedules obtained by data-driven and hybrid methods.

Each of the positions will focus on one of the two following topics :

Hybrid Methods for Planning and Scheduling
Deep learning has become the method of choice for perception tasks in computer vision or natural language processing. However, whether and how deep learning can help with planning and scheduling tasks, is very much an open question. Deep learning has mainly been used in reinforcement learning, which requires very large amounts of data and CPU, is unsafe, and doesn't easily capture and exploit existing knowledge of the problem. New approaches are needed that combine reasoning, learning and optimisation for planning and scheduling. For instance, new neural network architectures such as ASNets or STRIPS-HGN exploit the structure of planning problems to learn generalised heuristics and policies applicable to problems of any size. New innovative ways to guide search for constrained programming using learnt heuristics have been applied to real-world problems. There are many

opportunities to improve on these results. Depending on the applicant's interest the PhD will focus on one or more of the following :

a) Closely ntegrating learning with search. For instance, a search procedure guided by a neural network heuristic or policy could generate its own examples to improve the guiding policy/heuristic at the same time as solving the problem.
b) More effective deep learning architectures for planning and scheduling. For instance, architectures that exploit the relational structure of problems representations.
c) Handling constraints in deep learning for planning and schduling.
d) Predict+Optimise for planning and scheduling. Predict+Optimise is a paradigm to solve optimisation problems for which some of the input parameters must be predicted using ML.

Explananations and Verification of Learnt Policies and Schedules

The field of eXplainable AI (XAI) is arguably one of the pillars of trustworthy AI. Model agnostic approaches such as LIME and SHAP represent one of the most visible XAI approaches, but are unsound, which makes them unsuited to high-risk or safety-critical situations. Intrinsically explainable ML models often produce explanations that aren't sufficiently concise and exhibit redundancy. Formal XAI approaches offer guarantees of rigor unequalled by other approaches but currently suffer from a high computational complexity. Past work on XAI, including formal XAI, focused mostly on explaining "point" decisions, e.g. of ML classification algorithms. While this is still a challenging research subject, it is substantially easier than explaining a series of interacting decisions, as is required by planning and scheduling algorithms. Depending on the applicant's interest, the PhD will focus on one or more of the following :

a) Improving the practical efficiency of formal XAI approaches by devising heuristics and exploit user-interaction to navigate the space of explanations.
b) Extending formal XAI approaches to explain series of decisions made by plans, schedules, and policies learnt by ML or hybrid methods.
c) Designing scalable approaches for the related problem of verifying the robustness and safety of learnt plans, schedules and policies.

In both cases, the PhD research will involve **collaboration with TUPLES project partners** (Saarland University, KU Leuven, the University of Bologna, Airbus, Optit, and SciSports). The results will be **demonstrated and validated on some of the use cases** proposed by TUPLES' Industry partners, which concern aircraft manufacturing, aircraft operations, sport management, waste collection, and energy management.

SELECTION CRITERIA

- A Master's degree or equivalent in a relevant discipline (computer science, applied mathematics, or an appropriate branch of engineering)
- Prior background (relevant courses, Master's thesis) in artificial intelligence, operations research, and/or formal methods
- Excellent programming skills in a variety of programming languages
- Excellent formalisation skills
- Excellent spoken and written English
- Relevant publications would be a plus

APPLICATION PROCEDURE

Formal applications should include detailed cv, a motivation letter, university transcripts. Samples of published research by the candidate and reference letters will be a plus.
> applications should be submitted via this form: https://www.laas.fr/ost/node/234
or sent by email to: Sylvie.Thiebaux@gmail.com