

PhD project

Monitoring and constraining adaptive systems

The PhD candidate will be part of the Department of Information and Computing Sciences (Intelligent Systems group) of Utrecht University, in close partnership with the Department of Computer Science of the Vrije Universiteit Amsterdam (Computational Intelligence group).

“Where do you think you’re going?”

Keeping a close eye on adapting systems.

What will you do?

In the context of adaptive systems, we will design systems that allow for *monitoring* such systems in order to check if the adapting system obeys various types of constraint. Constraints may be the result of limited resources, laws and regulations, ethical or societal considerations, or any other type of constraint that is important for the environment at large in which the adaptive system operates. These constraints themselves may change over time: new laws may be introduced, safety regulations may become more strict, etc. The monitoring system will model these constraints in a human-intuitive way in order to facilitate updating the constraints by human agents, as well as to allow for explaining which constraints the adaptive system is or isn’t obeying.

The monitoring system will be able to detect and react to violations of constraints, and preferably predict that violations are about to occur in order to warn the human agent about this well in advance. Either the human agent or the monitoring system can subsequently provide input to the learning algorithm of the adaptive system to steer the system back on track.

Why is it important for Hybrid Intelligence?

Since we allow an adaptive system to change itself, we need to trust that it does not evolve into a system that violates the constraints. A monitoring system can thus be used to certify that an adaptive system adheres to certain constraints. Moreover, since the monitoring system models the constraints in a human-intuitive way, these can easily be inspected and changed. Finally, being able to predict the behaviour of an adaptive system allows for analysing and explaining it. These aspects are all important to facilitate communication and collaboration between human and artificial adaptive agents.

How will you do it?

Our monitoring system will consist of the following components:

- A 'normative' component that models the constraints that the adaptive system should obey. Depending on the type of constraint, these could be expressed in representations varying from pure logic-based to probabilistic-based. We will consider different existing representations and investigate their suitability for capturing different types of constraint.
- A 'learning' component that learns and is able to predict the actual behaviour of the adaptive system with respect to the constraints laid down in the normative component.

For our components we will build on the following existing techniques: Bayesian networks; Sum-product networks; and Deep generative modelling (e.g., VAEs, flow-based models). The former two can be both learned and handcrafted and can therefore be used in either component. The Deep generative models allow for detecting out-of-distribution examples and can generate new samples that allow for predicting the adaptive system's behaviour. In addition, our system will have available a set of metrics to compare the actual and desired behaviour and decide whether or not a warning should be issued. We will investigate existing metrics and design new ones tailored to the task at hand.

We are looking for....

an enthusiastic, creative and highly motivated candidate who has:

- a Master's degree or equivalent in Computer Science, Mathematics, (technical) Artificial Intelligence, or related disciplines;
- a solid background in maths and programming, demonstrated by successful completion of associated courses, and a keen interest in Artificial Intelligence;
- skills and interest in both modelling and (machine) learning, with a desire to combine theoretical and applied research in an interdisciplinary setting;
- excellent English communication skills (written and orally);
- a wish to contribute to the Department's teaching programmes;
- a talent for planning and organizing their work, and for meeting deadlines.

For more information....

- on this position: please contact



dr. Silja Renooij or
(s.renooij@uu.nl)



dr. Jakub Tomczak
(j.m.tomczak@vu.nl)

- on the Hybrid Intelligence project: <https://www.hybrid-intelligence-centre.nl/>