DCPF | Hub

Automatic Planning Laboratory

Our research revolves around automatic planning in all of its forms: time management, robots with autonomous capabilities, AI assistants (Automated Operations Department), etc. Our team is currently working on automated time management, that is, developing optimal schedules based on statistical data.

Current Task

A user has several blocks of free time available (for example, one per day). He has to complete N tasks, maximizing the remaining free time. Our problem: given the graph of dependencies between tasks "OrderGraph", which defines a partial order of task execution (for instance, if the lecture must precede the assignment, this graph will include a directed edge from lecture to assignment) and the performance influence graph “InfluenceGraph”, which indicates how task execution times are influenced by tasks previously completed, create an optimal schedule. Additionally, we need to create an “InfluenceGraph” based on the user's statistical data and adjust it based on the accuracy of our predictions.

Notation

In the OrderGraph, the directed edge (a, b) indicates that task "a" must be completed before task "b".

In the InfluenceGraph, the directed edge (a, b, w_sameday, w_sameweek) denotes the following relationships:

1) If "a" gets completed the same block before "b", then the modifier w_same (w_same > 0) is applied to the execution time of "b"
2) If "a" gets completed the same cycle before "b", then the modifier w_sameweek (w_sameweek > 0) is applied to the execution time of "b".
P.S.: For example, we might consider blocks to be daily 3-hour periods and cycles to be weeks.

Adversarial Machine Learning Laboratory

Our research interests include machine learning, neuroscience, computer science, and artificial intelligence. For the time being, our team is researching neural network hacking, applications, and countermeasures.

Current Task

The intention is to create a neural network that could perform Poisoning/Model Stealing/Backdoor attacks based on empirical information (black box analysis) about a different neural network. The Poisoning attack aims to sabotage model performance permanently, the Model Stealing attack aims to get personal information about the model's performance, and the Backdoor attack involves hijacking control over the network. Right now, we are constructing a simulation of empirical communication between neural networks.

Join us

We are searching for curious researchers interested in the problem we are solving. A background in artificial intelligence is desirable, but simultaneous training is also

possible.
Basic requirements:
Having sufficient free time, research experience (from lab work to a small personal project), a sufficient understanding of mathematics and programming, the capacity to read scientific material in English, and a general interest in the topic of the research.

Jailbreaking Lab

Our research examines the behavior of neural networks that are meant to breach security (“jailbreak”), by leaving some limited area ("jail") or by finding vulnerabilities within the area. Also, our research focuses on countering such “jailbreak”-networks. This area (“jail”) can be a part of computer memory, a process, a virtual machine, a simulation, etc. We're currently working on escapes from two-level software sandboxes, one of which contains a vulnerability.

Current Task

Design a neural network capable of taking over the second layer of a software sandbox via the vulnerability, breaching it, and escaping into the first layer.

Join us

We are searching for curious researchers interested in the problem we are solving. A background in artificial intelligence is desirable, but simultaneous training is also possible.
Basic requirements:
Having sufficient free time, research experience (from lab work to a small personal project), a sufficient understanding of mathematics and programming, the capacity to read scientific material in English, and a general interest in the topic of the research.