

HABILITATION THESIS REVIEWER'S REPORT

Masaryk University

Applicant

RNDr. Vít Musil, Ph.D.

Habilitation thesis

Bridging Continuous Learning and Discrete Optimization in Artificial Intelligence

Reviewer

Dr. Pasquale Minervini

Reviewer's home unit, institution

University of Edinburgh

This thesis is an extremely interesting research contribution to learning with discrete optimisation and finite-memory strategy synthesis, covering backpropagating through discrete combinatorial solvers, learning ILP constraints, and efficient policy evaluation. It is a collection of eleven papers, all published at leading venues (ICLR'20,'26; ICML'21,'24; CVPR'20, with a best-paper nomination; ECCV'20; UAI'21,'22; IJCAI'22,'23; ICAPS '26), organised in two streams -- differentiable programming for combinatorial algorithms, and strategy synthesis in stochastic and adversarial environments -- and introduced by a clear and well-written synthesis chapter that also summarises the author's contribution to each paper.

A core work in this dissertation is the blackbox-backprop paper, which proposes using a local surrogate gradient using a perturbed solver call. This contribution is extremely interesting, as it allows to back-propagate through arbitrary combinatorial solvers, like Dijkstra's shortest path algorithm, optimal allocation and scheduling algorithms, graph matching algorithms, etc.; when it was published it was very inspiring (e.g., our <https://arxiv.org/abs/2209.04862> and <https://arxiv.org/abs/2106.01798> are blackbox-backprop in disguise). The CombOptNet paper then focuses on the natural problem of learning the constraints (which in the blackbox-backprop paper are assumed as given) as well as costs; and the LPGD paper relates several update rules through Lagrange-Moreau envelopes. Then, ideas in these papers are applied to back-propagate through graph-matching algorithms and ranking losses.

Application papers are also extremely interesting; e.g., the RaMBO paper is based on the observation that ranking is itself a linear program over permutations, so an off-the-shelf argsort becomes a differentiable layer with $O(n \log n)$ forward and backward passes. The paper proposes several contributions around this, including a family of recall losses whose closed form finally explains why optimising the logarithm of a rank had been found "beneficial" in earlier work. The result is competitive with the state of the art in image retrieval and gives consistent gains in object detection. BB-GM wraps an unmodified Lagrangian-decomposition graph-matching solver and still advances the state of the art on Pascal VOC (80.1% vs. 73.2% for the reproduced DGMC) and on SPair-71k. The DiffMJX paper moves the same philosophy to differentiable physics: the diagnosis that gradient errors in penalty-based simulators come from time discretisation of stiff ODEs (and not from time-of-impact effects) is really interesting, and Contacts-from-Distance applied only in the backward pass via STE at the ODE level is a very nice idea.

The remaining contributions are also interesting. The Regstar paper focuses on whether finite-memory strategies achieve the optimal value of adversarial patrolling games in the limit and gives a polynomial DP algorithm that evaluates the protection value and its gradient by aggregating exponentially many attack paths in a backward, Dijkstra-like search. This turns strategy synthesis into gradient ascent and scales to instances that were out of reach before. The changing-environment paper identifies a new phenomenon, the "security hole" created by switching between two individually optimal strategies, and shows how to bound it, how to warm-start the new strategy from the old one, and how to mitigate the hole by randomised switching. The memory-assignment paper analyses where controller memory should be placed; reading conflicting gradient signs of near-maximal attacks as evidence that a state must be split is simple and, empirically, robust across five benchmark families. The two IJCAI papers extend the framework to recurrent-reachability objectives and to periodic maintenance, where randomised finite-memory schedules act as a compact generative model from which long deterministic cycles (100–200 nodes, exceeding twelve working shifts) are sampled.

Reviewer's questions for the habilitation thesis defence (number of questions up to the reviewer)

- In blackbox-backprop, λ is chosen by hand to balance $\langle w \rangle$ against $\langle dL/dy \rangle$; with LPGD's error bound now scaling as $1/\tau$, is there a principled or adaptive way to set it?
- Modern-day LLMs are mainly trained to use external tools and solve complex tasks via policy gradient methods (GRPO etc.); is there any ideas from the blackbox-backprop literature that could be useful in how we design and/or train LLMs?
- About CombOptNet, can it be used in practice? E.g., can it learn the constraints behind combinatorial solvers for finding shortest paths in a graph? If yes, what's the scale of the graphs it can handle on today's hardware?
- In CombOptNet, many constraint systems explain the same (c, y^*) pairs; what is actually being learned, and when is a prediction feasible under the true constraints rather than the learned ones?
- LPGD generalises many approaches for backpropagating through combinatorial solvers – does it also allow to design new methods that have not been proposed yet?
- Sect. 7.2 says GD and LPGD "have similar sensitivities to inaccurate solutions" while App G.2 says "LPGD is more sensitive to inaccurate solutions than GD"; which is it, does the answer depend on τ , and what does it imply for heuristic forward solvers?

Not related to the papers, but I'm very curious about these:

- Which single result in the thesis do you like the most, and what is the next problem you want to attack?
- If you received a 10M EUR donation in funding and five years to produce the next breakthrough in AI (with no teaching or admin duties), what would you work on?

Conclusion

The habilitation thesis entitled *Bridging Continuous Learning and Discrete Optimization in Artificial Intelligence* by RNDr. Vít Musil, Ph.D. **fulfils** requirements expected of a habilitation thesis in the field of Informatics.

Date: 18/09/2026

Signature: