Ottawa, April 18, 2026

To Whom It May Concern,

I confirm that I have read the doctoral thesis entitled SampleEfficient Actor-Critic Algorithms in Reinforcement Learning written by Mr. Michał Nauman under supervision of dr hab. Marek Cygan prof. ucz. (University of Warsaw), presented to the Academic Council of Computer and Information Sciences of the University of Warsaw.

I confirm that I deem the thesis, as presented, sufficient to grant a PhD. It is built from six scientific papers accepted at top-tier conferences. Together, these papers form a body of work exploring techniques for improving the sample efficiency of Actor-Critic methods in deep reinforcement learning, approaching the problem from both theoretical and empirical perspectives.

Historically, advances in (non-deep) reinforcement learning placed a strong emphasis on theoretical approaches. Since the use of neural networks has become ubiquitous, it becomes increasingly clear that the most impactful performance improvements come through empirical insights. As such, the empirical works in this thesis appear to me to be the ones which will have the biggest impact in the field.

I found [P2] particularly insightful, where the authors conduct an exhaustive empirical evaluation on SAC to determine what interventions work best at addressing value overestimation, critic underfitting, and plasticity loss. The authors argue the "bitter lesson" that general-purpose neural network regularizers outperform RL-specific interventions, even though the challenges addressed are specific to RL.

Perhaps the contributions that may end up having the most impact in the field are [P5] and [P6], which tackle the problem of scaling RL networks in a way that reliably improves performance. In [P5] the authors combine many of the insights in the prior works to introduce BroNet, an RL network architecture that improves with scale. In [P6], the authors propose a series of empirical scaling laws for scaling the replay ratio (number of gradient updates per environment step), which can aid in devising better predictors for performance at different levels of compute.

Although I have highlighted only some of the works presented in this thesis, all of them are advancing in a direction that is relevant and important to the study of reinforcement learning.

The output during Michał's PhD is quite impressive, as six top-tier publications are more than what I am accustomed to seeing for PhD students. Given this, as well as the cohesiveness in his work, I would like to recommend an honorary distinction.

If there are any questions or concerns, please do not hesitate to contact me.

Best,

Pablo Samuel Castro
Adjunct professor, Université de Montréal
Core Industry Member, Mila - Québec AI Institute
Senior Staff Research Scientist, Google DeepMind