Meta-Learned Hamiltonian

Seungjun Lee and Woojae Seong

Seoul National University

Summary

We introduce meta-learning algorithms to identify the shared representation of Hamiltonian systems.

Introduction

- A Hamiltonian system is usually governing by a mathematical expression with several physical parameters. The Hamiltonian of an idea pendulum consists of

- shared expression: $H = \frac{p^2}{2ml^2} + mgl(1 - \cos q)$,
- physical parameters: mass m, pendulum length l and gravity constant g,
- state: angle q and angular momentum q.

From observing several pendulums, the shared expression of the Hamiltonian could be established. Then, researchers can readily recognize a new pendulum-like system by adapting new physical parameters on the expression. However, identifying an appropriate Hamiltonian is very important yet extremely hard in most science and engineering problems.

- In meta-learning point of view, our goal is identifying the meta-transferable knowledge of Hamiltonian,

- meta-transferable knowledge: $H_\theta(q, p)$
- task-specific parameter: $\theta_i \sim p(T)$
- state: angle q and angular momentum q.

- Adapting a hypothesized governing expression of Hamiltonian to many related phenomena by adapting new physical parameters to the governing equation is similar to the task-adaptation process in meta-learning. Then, the hypothesized governing expression is corrected by validating newly observed systems that are assumed to be governed by the same physical laws. After many verifications on many related systems, the governing equation of Hamiltonian is established which is a similar process to the meta-learned model with lots of iterations across various tasks.

Contact Information

- Email: tl7qns7ch@snu.ac.kr, wseong@snu.ac.kr

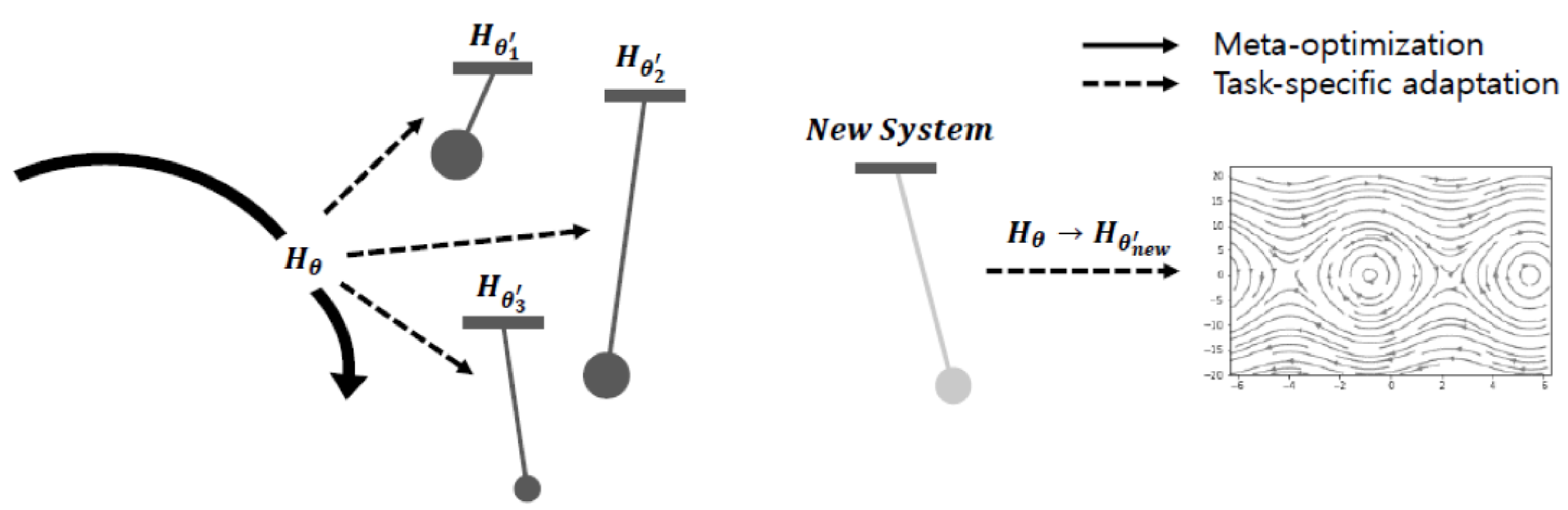


Figure: A hypothesized governing equation of Hamiltonian, usually corrected and established by evaluating many related systems, could be learned using meta-learning as a data-driven method (left). The well-established Hamiltonian can be utilized to predict new system dynamics, which could be viewed as a meta-transfer process by a well-trained meta-learner (right).

Hamiltonian Neural Networks

In Hamiltonian mechanics,

- position: $\boldsymbol{q} = (q_1, q_2, \ldots, q_n)$,
- momentum: $\boldsymbol{p} = (p_1, p_2, \ldots, p_n)$,
- canonical coordinates: $\boldsymbol{x} = (\boldsymbol{q}, \boldsymbol{p})$,
- Hamiltonian equation:

$$\frac{d\boldsymbol{x}}{dt} = \left(\frac{\partial H}{\partial \boldsymbol{p}}, -\frac{\partial H}{\partial \boldsymbol{q}}\right) = \Omega \nabla_{\boldsymbol{x}} H(\boldsymbol{x}), \tag{1}$$

where $H(\boldsymbol{x}) : R^{2n} \rightarrow R$ is the Hamiltonian that is conservative during the process and $\Omega = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$ is a $2n \times 2n$ skew-symmetric matrix.

The Hamiltonian function can be approximated by neural networks, H_θ, called Hamiltonian Neural Networks (HNN) [1]:

$$L_{HNN} = \left\| \frac{d\boldsymbol{x}}{dt} - \Omega \nabla_{\boldsymbol{x}} H_\theta(\boldsymbol{x}) \right\|_2^2. \tag{2}$$

Gradient-Based Meta-Learning

Given task distribution $T_i \sim p(T)$ which is composed with task-specific train set and test set, $D_i = \{D_i^{tr}, D_i^{te}\}$. Model-agnostic meta-learning (MAML) [2] consists of the task-specific adaptation on train sets (inner-loop),

$$\theta = \theta - \alpha \nabla_\theta L_{T_i}\left(D_i^{tr}; \theta\right), \tag{3}$$

and the meta-optimization on test sets (outer-loop),

$$\theta \leftarrow \theta - \beta \nabla_\theta \sum_{T_i \sim p(T)} L_{T_i}(D_i^{te}; \theta_i'). \tag{4}$$

From the *feature reuse* hypothesis [3], they hypothesized that the body of the model behaves as a shared representation across the different tasks, whereas the head of the model behaves as a task-specific parameter. In the modified algorithm called Almost No Inner Loop (ANIL), the inner-loop is replaced by

$$(\theta^{(1)}, \ldots, \theta^{(l-1)}, \theta^{(l)} - \alpha \nabla_{\theta^{(l)}} L_{T_i}\left(D_i^{tr}; \theta\right)), \tag{5}$$

where the θ consists of l layers.

Experiments and results

- Physical parameters: $(m, l, q_0) \in ([0.5, 5], [0.5, 5], [-\pi, \pi])$,
- sampled states: $(q, p) \in ([-2\pi, 2\pi], [-20, 20])$,
- meta-train sets: 10,000 tasks consisting of train set and test set given by 50 randomly sampled point state and their time-derivative pairs $(\boldsymbol{x}, \frac{d\boldsymbol{x}}{dt})$,
- meta-test sets: train set of new systems with $K = \{25, 50\}$ points are given and fine-spaced test sets of new systems consist of 50 equally spaced grids for each coordinate.

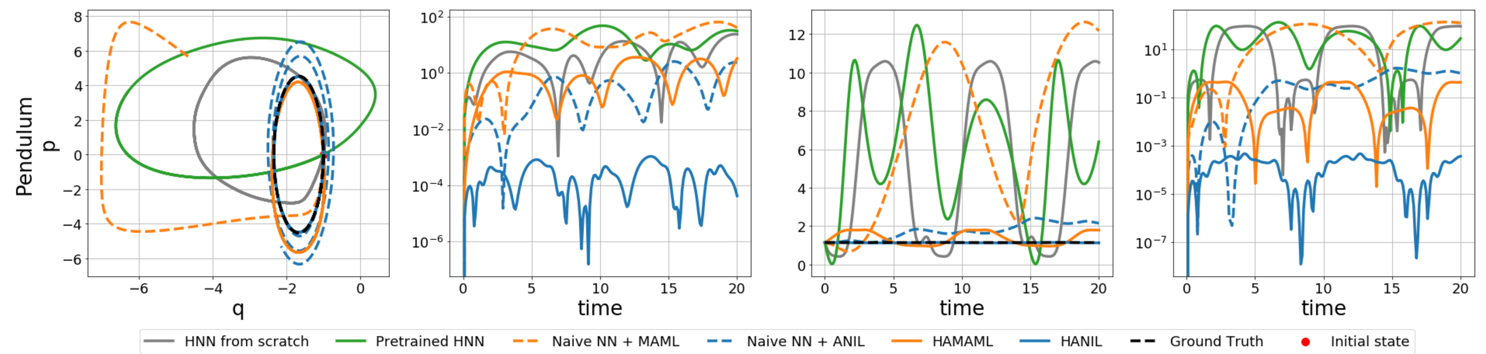
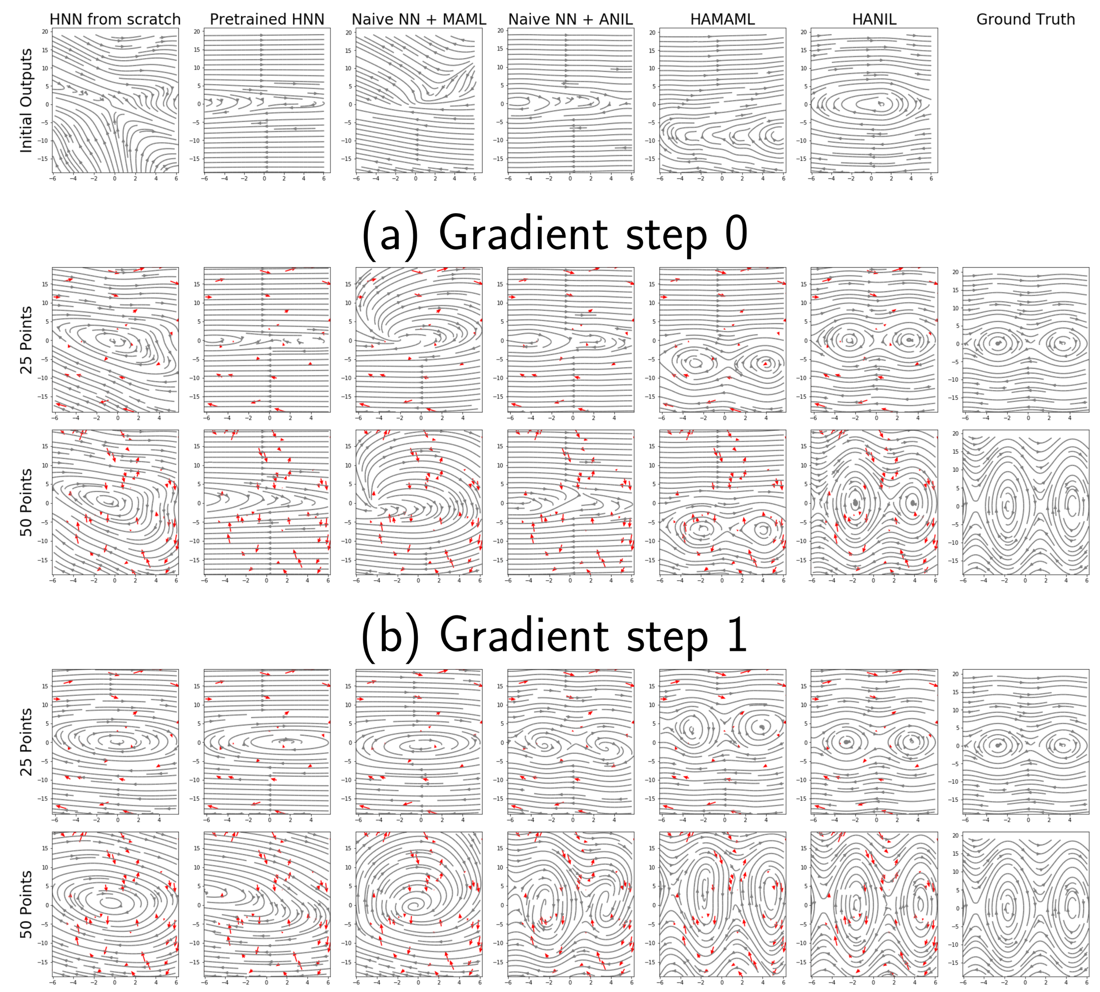


Figure: Predicted state and energy trajectories from the initial states, and their corresponding MSEs starting from an initial state during $20s$.



(a) Gradient step 0

(b) Gradient step 1

(c) Gradient step 10

Figure: Predicted vector fields (gray streamlines) by adapting the learners to observations of new pendulum systems as given point dynamics (red arrows) after the corresponding gradient steps.

Conclusion

From observing the similarity between seemingly unrelated problems, identifying the Hamiltonian and meta-learning, we formulate the problem of identifying the Hamiltonian as a meta-learning problem. The results state that the proposed methods have the ability to extract the meta-learned Hamiltonian representation, which can be considered as physical nature across the observed systems during meta-training.

References

[1] Samuel Greydanus, Misko Dzamba, and Jason Yosinski. Hamiltonian neural networks. In *Neurips*, 2019.

[2] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *ICML*, 2017.

[3] Aniruddh Raghu, Maithra Raghu, Samy Bengio, and Oriol Vinyals. Rapid learning or feature reuse? towards understanding the effectiveness of maml. In *ICLR*, 2020.