Reinforcement Learning (RL)-based Shape Optimization of 2D profile extrusion die geometries [1]

Clemens Fricke[1] Daniel Wolff[2] Marco Kemmerling[3] Stefanie Elgeti[1,2]

Background

Demand:

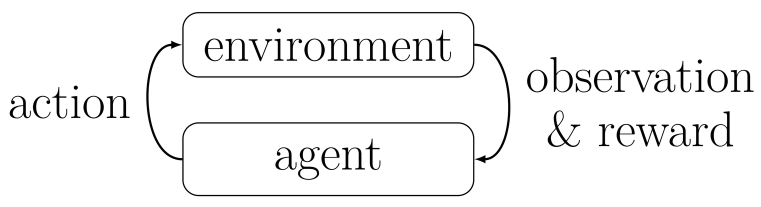
In profile extrusion, the die is a very important component. It ensures that the material is shaped into the desired profile. During the cooling process, the imparted shape will deform due to temperature-based shrinkage as well as inhomogeneities in the material. This can be accounted for in the die design, as the optimization of the die shape currently is computationally very expensive.

Aim:

As seen in [2], there have been advances in RL-based Shape Optimization. RL-based Shape Optimization splits the computational load into an offline part where the agent (optimizer) is trained for a preselected training domain. Inference, that is, the online optimization of a single problem from the domain of feasible parameters, is then much shorter.

Reinforcement Learning

In RL, the agent selects an action (new design variables). This action is used in the environment to produce a new observation (geometry). During training, also a reward (objective function) is supplied. It informs the agent whether the last action improved the geometry.



There are two shape optimization strategies for RL. For the incremental approach, the agent iteratively outputs incremental optimizations of the previous design variables. For the direct approach, in contrast, the agent directly yields the optimal design variables.

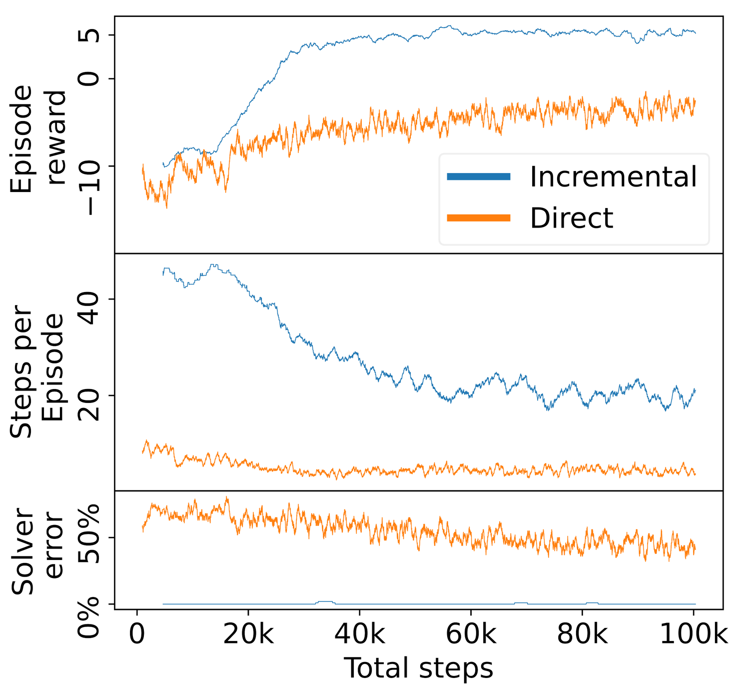
$s_i \longrightarrow$ agent $\longrightarrow \Delta s_i$ — Incremental Strategy

$s_i \longrightarrow$ agent $\longrightarrow s*$ — Direct Strategy

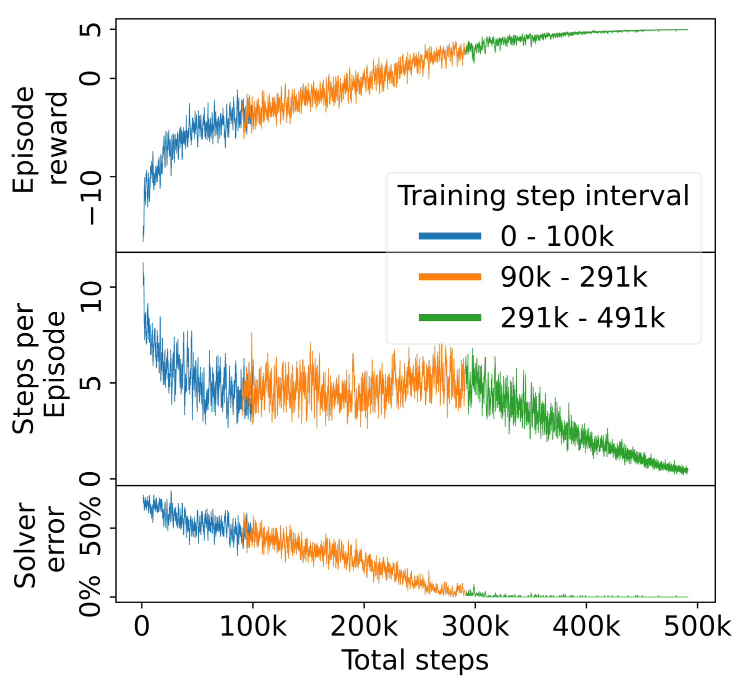
Results - Serial

Incremental vs. Direct:

Incremental converges, but Direct does not converge after 100k iterations.

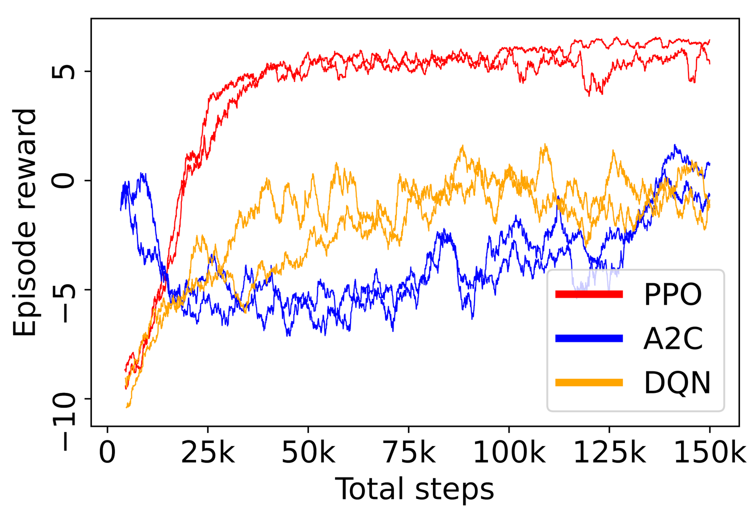


Direct converges after additional training but is overall 10 times slower.

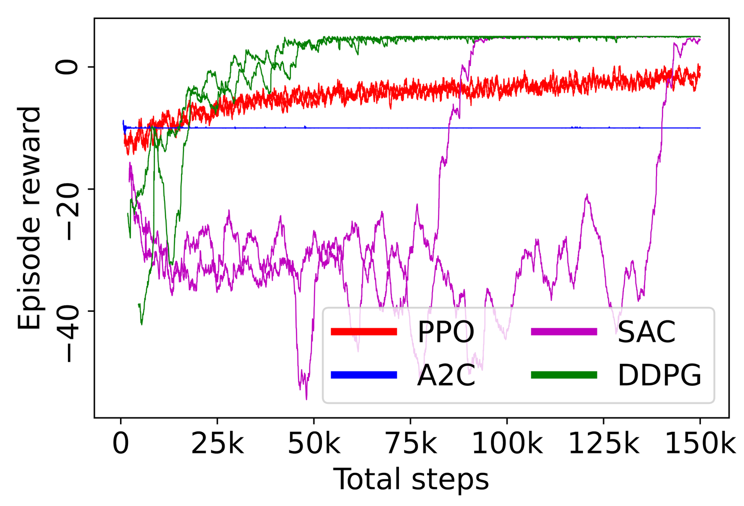


Agent:

We used the PPO agent as our baseline. Depending on the action space, we compare this baseline to other agents (A2C, DQN, SAC, DDPG).



- PPO is the fastest.
- A2C and DQN do not fully converge.

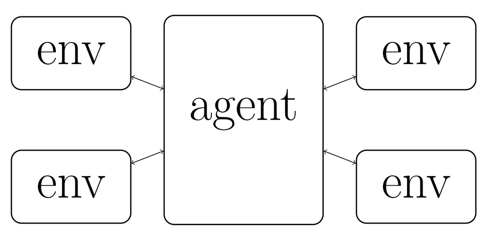


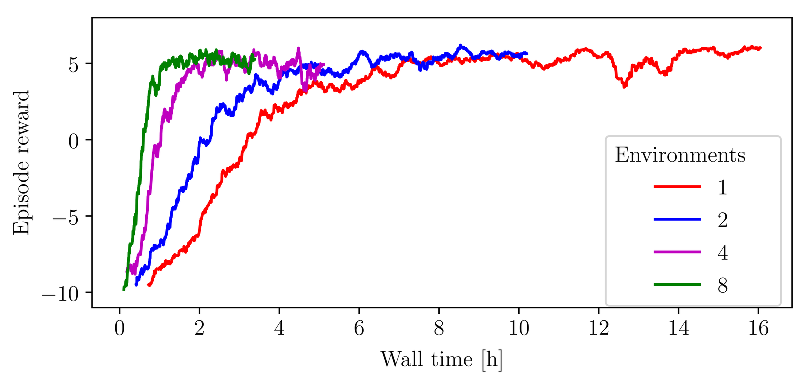
- PPO is slowest to converge, DDPG fastest.
- A2C does not converge, SAC shows inconsistent behavior.

Results - Parallel

Vectorized Environment Training:

To reduce the wall-clock offline training time, the agent learns with multiple environments in parallel.





Results in an increase of computational load by 38%, but also in a wall-clock time speedup of 5.

Future Work

- Apply methodology to 3D shapes.
- Expand observation space for better generalization of learned optimizations.

References

[1] C. Fricke, D. Wolff, M. Kemmerling and S. Elgeti, (2023) Reinforcement Learning based Shape Optimization of 2D profile extrusion die geometries, *Advances in Computational Science and Engineering*, **1**(1):1-35

[2] J. Viquerat, P. Meliga, A. Larcher, E. Hachem; (2022) A review on deep reinforcement learning for fluid mechanics: An update. *Physics of Fluids*; **34**(11): 111301.

[1]Institute of Lightweight Design and Structural Biomechanics, TU Wien, Austria [2]Chair for Computational Analysis of Technical Systems, RWTH Aachen University, Germany
[3]Information Management in Mechanical Engineering, RWTH Aachen University, Germany

Simulations were performed with computing resources granted by RWTH Aachen University and projects thes1136 and jara0185.