Optimizing the microstructure in open-die forgings using reinforcement learning

REINISCH Niklas [1,a *], BAILLY David [1,b] and HIRT Gerhard[1,c]

[1]Institute of Metal Forming, RWTH Aachen University, Intzestraße 10, 52072 Aachen, Germany

[a]niklas.reinisch@ibf.rwth-aachen.de, [b]david.bailly@ibf.rwth-aachen.de, [c]gerhard.hirt@ibf.rwth-aachen.de

Keywords: Open-Die Forging, Reinforcement Learning, Process Optimization, Grain Size

Abstract. The open-die forging process can produce large workpieces with excellent material properties that can be used for heavy-duty applications like turbine shafts. The mechanical properties result from the microstructure, which in turn directly results from the process route. Since in open-die forging processes commonly hundreds of unique forming operations are carried out, numerous process routes lead to the same final geometry but produce different microstructures. This is why the prior design of an optimal pass schedule is essential to ensure good mechanical properties of open-die forgings. In the past reinforcement learning (RL) was already used to design optimized pass schedules for open-die forging that achieve the desired geometry, utilize the available press force, and reduce the number of passes. Furthermore, the design of a single pass schedule only took a few seconds which creates opportunities for the use of RL in control systems e.g. for the microstructure. This is why in this publication an existing RL algorithm is extended so that microstructure can be included in the optimization. Within this process, a microstructure model is integrated into the RL algorithm and the reward function (defines the goal of the training process) was extended in two steps to also rate the achieved average grain size continuously dependent on the temperature of the workpiece. In addition, the RL implementation was changed to ensure the production of the desired final geometry leading to a decrease in the complexity of the optimization problem. Thus, both an improvement of the designed pass schedules and a significant reduction of the training time of the RL algorithm was achieved.

Introduction

Open-die forging is a bulk metal forming process that offers the possibility to produce large mostly longitudinally oriented workpieces while removing casting defects from the component. This includes the closing of casting voids and the transformation of an unfavorable microstructure that can result from casting into a fine homogeneous one. Therefore, open-die forgings exhibit excellent mechanical properties and are often used for heavy duty applications like turbine shafts. During the forging process, the workpiece is incrementally formed using simple tools. Therefore, the final geometry of the component results from the sequence of forming operations, which can comprise up to several hundred individual strokes. This leads to an infinite number of different forging processes, which all result in the same final geometry, but can differ significantly in terms of the microstructure produced, the energy required, or the process duration.

To ensure efficient processes and high component quality in open-die forging targeted process design and, in the long term, process control are of great importance. An optimized process design, considering the microstructure produced, can be achieved e.g. via finite element (FE) simulations coupled with a microstructure model. Since FE simulations of open-die forging processes often require long computing times, a fast design or the control of forging processes is not possible using this approach. In the industry forging processes are until today often designed based on experience or using simple models describing e.g. the geometry development. In addition, commercial software to design pass schedules, such as the program "ForgeBase®" from SMS group GmbH [1] or "easy2forge" from Dr. Fister GmbH [2], are available. Although these programs can quickly design (in some instance optimum) pass schedules for various ingot geometries, considering e.g. plant boundaries, the microstructure generated is not yet included in the design. Overall, a need for new possibilities for rapid design of optimum pass schedules as well as targeted process control arises, whereby besides geometry further workpiece parameters such as grain size are included.

In this publication, a novel approach for the design of optimized open-die forging processes using the machine learning method reinforcement learning (RL) is presented. RL algorithms as well as deep reinforcement learning (DRL) approaches have been successfully applied in various manufacturing applications including metal forming for different process design and control tasks [3,4]. Gamal et al. [5] trained a deep deterministic policy gradient algorithm to control the roll gap in bar and wire hot rolling using real plant measurement data. Dornheim et al. [6] used a RL approach coupled to a FE-simulation of the process to optimize the blank holder force in sheet metal deep drawing. Scheiderer et al. [7] used a soft actor critic algorithm to design an optimal pass schedule for heavy plate rolling that produces a desired grain size.

Since RL algorithms learn and store general rules for optimal behavior in different situations, after a successful training process they can determine solutions for individual problems in a very short time. In addition, (deep) neural networks can be used to store the knowledge learned during the training process, which can map even highly nonlinear correlations that are often present in open-die forging e.g. due to the incremental forming. For this reason, a double deep q-learning (DDQL) algorithm for the optimized design of open-die forging processes was developed in [8]. The algorithm is capable of designing pass schedules that achieve the desired final geometry while making use of the available press force and reducing the number of passes. Since the design of one optimized pass schedule via RL took only 2-4 seconds, the potential of RL for process planning and control has already been demonstrated. Since microstructure has a great influence on the mechanical properties of a forging, in this publication, a microstructure model is incorporated into the existing DRL algorithm to enable consideration of the produced grain size in the design of optimized pass schedules for open-die forging.

Previous Reinforcement Learning in Open-Die Forging

The machine learning approach RL is similar to human trial-and-error learning [3]. In an industrial open-die forging plant, this learning procedure runs unconsciously. Both press and manipulator operators select process parameters such as height reduction or bite ratio on a daily basis, interacting with their environment the open-die forge. In addition, they receive subsequent feedback on the part quality, e.g. good part or scrap, and thus with time learn how different materials and workpiece geometries are forged optimally (cf. Fig. 1, left).

Throughout its training process, the RL algorithm implemented in [8] undergoes the same learning process. Here, a so-called RL agent (cf. Fig. 1, center) interacts with its environment by selecting actions that correspond to process parameters such as bite ratio or height reduction (cf. Fig 1, middle). In contrast to a plant operator, who learns directly from the real forging process, the RL agent uses a virtual forging environment (cf. Fig. 1, right). This incorporates fast process models that can calculate the change in key workpiece properties such as geometry [9], temperature [10], and equivalent strain (in the core fiber) [11] as well the process duration and press force. Since it is often necessary to design several thousands of pass schedules in a RL training process as well as analyzing them concerning the material and process properties achieved, using a real forging plant for training is not feasible. This is due to risks for operators and machines as well as considerable costs incurred from the material used, wear, and the occupancy of staff and plants.

The virtual forging environment uses the process parameters selected by the agent to adjust the current state of the workpiece, including its geometry, temperature, and equivalent strain (in core fiber). An evaluation of this action selection according to a so-called reward function follows. The reward function assigns a scalar reward to each process parameter selection of the algorithm considering the current ingot state and hence, defines the goal of the RL training process. Afterwards, the agent uses the rewards and the information on the state change to adapt its action selection rules stored in a neural network (cf. Fig 1).

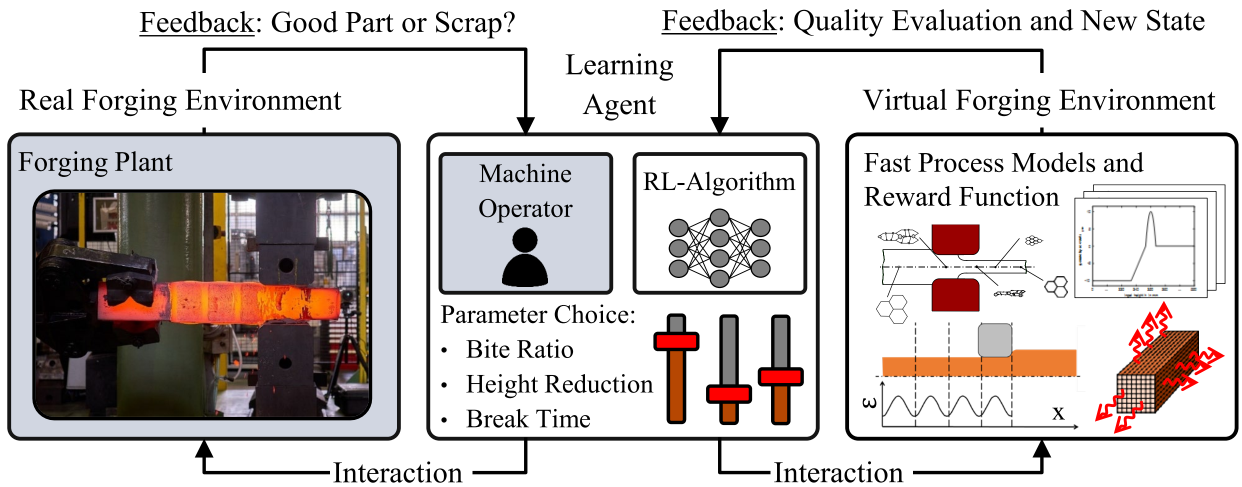

Fig. 1. Graphical representation of the trial-and-error learning procedure; left: human plant operator; right: RL algorithm.

In the given implementation, the RL agent learns how to design efficient forging processes. The goal includes not only to produce the desired final geometry but also to minimize the number of passes and to make the best possible use of the available press force. At the same time, process and plant limits such as the maximum force of the press are considered in the pass schedule design. For this reason, 80 % of the maximum press force is set as the nominal value, which, when adhered to, provides the maximum reward for the algorithm. In contrast, the reward function produces a penalty if the press force limit is exceeded. A detailed description of the implementation as well as the training process of the RL algorithm can be found in [8].

Description of the Use Case

To ensure good mechanical properties in open-die forged components while forging efficiently a targeted adjustment of the microstructure is essential. The RL algorithms presented throughout this publication are all trained to design pass schedules considering forging ingots made of 1.4301 steel (AISI 304), with a starting temperature of 1100°C and an initial grain size of 200 µm. The initial ingots have square initial cross-sections with edge lengths between 180 mm and 220 mm and are drawn out to final cross-sections with edge lengths between 100 mm and 140 mm. The initial length of the forging ingots is 500 mm. In the end, the trained RL algorithms should be able to give optimal pass schedules in seconds for various combinations of initial and final ingot geometries within the above-mentioned ranges. This would enable the use of RL-algorithms not only for fast offline design but also for online process adaption in order to control forging processes.

In a first step, the existing RL algorithm for efficient forging (no consideration of grain size) is trained without modifications. Afterwards, 400 different pass schedules with uniformly distributed combinations of initial and target geometry were designed using the RL agent. A semi-empirical microstructure model [12] was used to analyze the produced grain size along the core fiber of each pass schedule. This creates a comparative measure to rank the effectiveness of the later discussed changes that account for grain size in the optimized pass schedule design. Fig. 2 shows the consolidated results using a heat map. Here, each box represents a pass schedule whose associated initial and target workpiece geometry can be read on the x and y axes, respectively. The coloration of the box represents the average austenitic grain size achieved in the core fiber of the workpiece at the end of the forging.

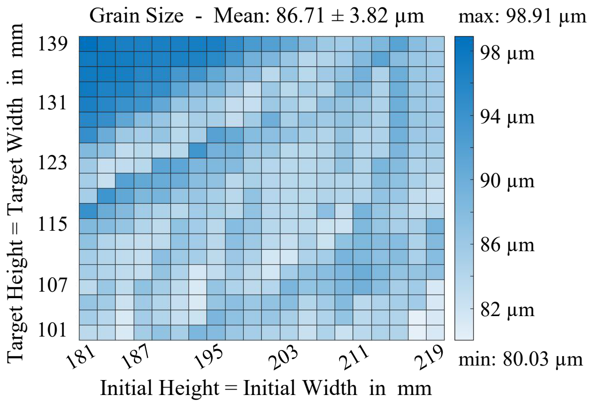

Fig. 2. Heatmap showing the produced average grain sizes of 400 different pass schedules designed by RL without a specific grain size reward.

The generated average grain sizes vary between 80.03 µm and 98.91 µm with a standard deviation of 3.82 µm. Averaged over the 400 pass schedules, the mean average grain size is 86.71 µm. Furthermore, averaged comparative values for the deviation from the target geometry (± 0.23 %), the average number of passes (6.15), and the utilization of the press force (± 5.02 %) are calculated and listed in Table 1.

Extension of the Previous RL Implementation

To enable the RL algorithm (cf. Fig. 1) to consider the production of a specific grain size in the process design, the existing implementation was extended in three steps. The first step consists of the implementation of a microstructure model developed by Karhausen et al [12] into the environment of the algorithm. The model uses semi-empirical equations to estimate recrystallization and grain growth, as well as a substructure modeling to determine the change in austenite grain size during forging. The model input is time-resolved data of temperature and equivalent strain along the core fiber that is already available in the environment of the RL algorithm.

In the second step, the previous reward function (cf. [8]) consisting of a weighted sum of the single rewards for force utilization and geometry, was complemented by an additional term R_{GS} to evaluate the generated mean austenitic grain size in the core fiber at the end of each process (cf. Fig. 3). As shown in Fig. 2, the mean average grain size produced by the previous algorithm was 86.71 ± 3.82 µm with a minimum grain size of 80.03 µm. To force the adapted RL agent to change its pass schedule design pattern for each combination of initial and target geometry, a target grain size of 70 µm, as well as a permissible maximum deviation of ± 10 µm, was chosen. This ensures that the new target grain size does not overlap with the previous grain size spectrum of the efficient forging use case (cf. Fig. 2).

To reflect this, a grain size reward function R_{GS} was designed, which is evaluated only once at the end of each pass schedule and hence, is referred to as a sparse reward. The reward R_{GS} divides into three characteristic sections depending on the mean grain size (cf. Fig. 3). The first section covers the permissible target grain size region of 70 ± 10 µm. Here, the reward increases parabolically and reaches its maximum at the exact target grain size. Outside the permissible grain size range, the reward drops linearly with increasing deviation from the target grain size in section two until it remains constant at its minimum in section three. This lower limit of the reward is called reward clipping. It prevents large gradients from occurring in the training process of the RL algorithm, which can increase the training stability.

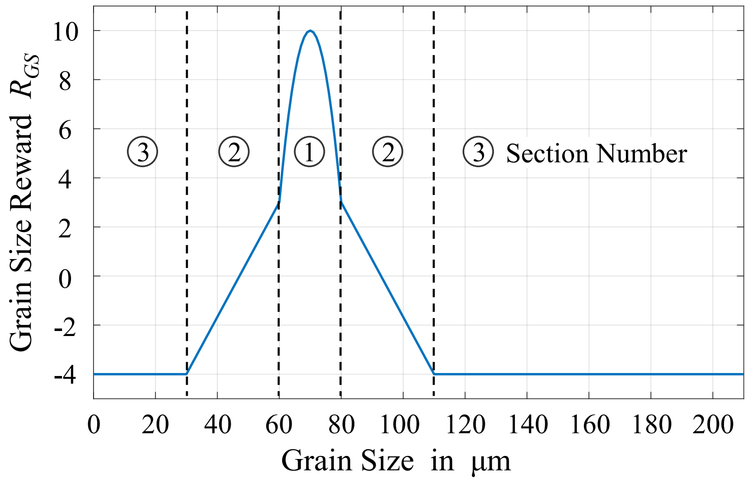

Fig. 3. Graphical representation of the sparse grain size reward R_{GS}.

Finally, the action selection of the algorithm, which previously consisted of height reduction (3% - 30%) and bite ratio (0.3 – 0.9), was extended by the selection of a break time between 0 and 120 seconds before every second pass. This enables the algorithm to influence the microstructure significantly more throughout the forming process. Now, depending on the current ingot state, the RL can e.g. lower the workpiece temperature by cooling or give material-physical processes such as static recrystallization or grain growth the necessary time to complete.

Training of the RL Algorithm – Sparse Grain Size Reward

After the RL algorithm was extended and the use case was defined, a new training process with 100,000 iterations was carried out. This corresponds to 100,000 designed pass schedules that are analyzed concerning their final workpiece properties by the environment using fast process models. The training process took about eight days on four cores of an Intel® Xeon® E5-2667 v3 processor with the algorithm successfully converging against high rewards.

To classify the quality of the generated pass schedules, again 400 different pass schedules were designed using the trained RL agent and then analyzed concerning their achieved grain sizes. Subsequently, averaging was used to generate comparable process parameters such as a mean number of passes or an averaged deviation from the target geometry. On average, the grain size over all 400 pass schedules is 70.99 ± 3.58 µm (cf. Fig. 4) and thus close to the desired target value of 70 µm. In comparison to the previous RL algorithm for efficient forging, the mean grain size decreased significantly throughout all 400 pass schedules (cf. Fig. 2) while the standard deviation remained almost the same. Moreover, nearly all processes produce a grain size within the permissible range between 60 µm and 80 µm. Only a small cluster of five pass schedules (cf. Fig. 4, red frame) exceeds the permissible maximum value of 80 µm with a maximum grain size of 82.05 µm for the combination of an initial height of 189 mm and a final height of 139 mm.

In addition to the new optimization goal of generating a defined grain size, the agent also successfully took into account the objectives for efficient open-die forging. The average number of passes was reduced to 5.9 while the designed pass schedules use the available press force well. This is underlined by a small average deviation of 6.1 % from the target press force, whereby the maximum press force is not exceeded in any pass schedule. In general, the designed pass schedules lead to the desired target geometry well resulting in an average deviation of 0.31% between generated and desired final height. Nevertheless, one of the 400 pass schedules (initial height = 213 mm and target height = 137 mm) failed on the permissible final height range (target height ± 1%) with a deviation of -1.02 % from the target geometry.

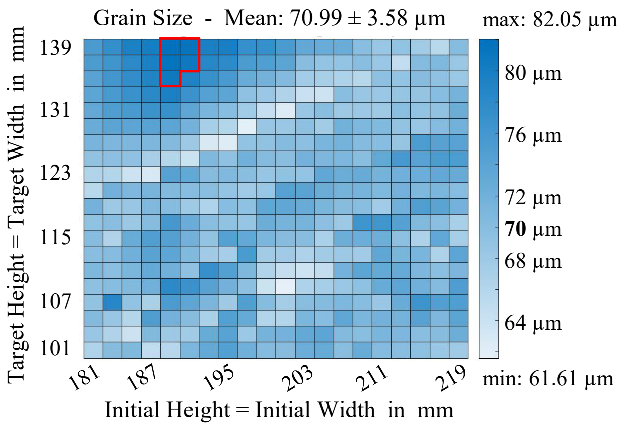

Fig. 4. Heatmap showing the produced average grain sizes of 400 different pass schedules designed by RL with a sparse grain size reward (cf. Fig. 3).

The pass schedules failing slightly on complying with the target grain size and target geometry ranges indicate that the algorithm might not have converged to its full extent yet. One reason for this slow and incomplete convergence could be the increased complexity of the optimization problem. Furthermore, an unclear formulation of the optimization goals within the reward function can result in bad convergence as well as the observed grain sizes outside of the desired range. A longer training process might improve the result quality but due to the already high training duration of several days, it is not a feasible solution. Instead of longer training durations, the following chapter aims at two different approaches to increase the efficiency of the training process: Reducing the complexity of the optimization problem and improving the reward function.

Reduction of Complexity

The first change to the implementation of the RL algorithm aims at reducing the complexity of the optimization problem to be solved. Until now, the algorithm has had to balance four optimization objectives (set final geometry, minimize number of passes, utilize press force, set grain size) and satisfy all of them simultaneously. So far, the final geometry resulted from the selected height reductions and bite ratios, considering the occurring spread. Since the desired final cross section is known upfront and fast process models can predict the geometry change during forging, from now on the desired final geometry is directly ensured in the implementation of the RL environment. As soon as the algorithm selects a combination of height reduction and bite ratio that causes the final height of the workpiece to fall below the desired final height, a correction of the last two passes happens. Here, the height reduction is inversely lowered by backwards calculation of the geometry model ensuring the exact production of the target geometry. The bite ratio and break time selected by the algorithm remain unchanged during this adjustment. This change to the implementation removes one of the four optimization objectives, so that the complexity of the optimization problem decreases significantly, which should lead to improved convergence in the training process. Since hereafter, the desired geometry is always produced perfectly, the associated reward calculated after the last pass is invalid and removed from the reward function.

Furthermore, a restriction of the number of actions selectable by the agent decreases the complexity even further. The discrete action choice of the used RL (double deep Q-learning) agent is based on the estimation of the value $Q(s,a)$ for all possible action choices a in the current state s. The Q value represents a measure of the total reward expected for the remaining process, which the agent tries to maximize. Therefore, the agent always chooses the actions with the highest Q-value, promising the highest total reward gained. So far, the RL agent can choose bite ratios between 0.3 and 0.9 in increments of 0.05, and height reductions between 3% and 30% in increments of 1%. In combination with break times, which the algorithm is allowed to choose between 0 s and 120 s in 30 s increments, 1820 combinations of action choices are possible in each state of the forging. Hence, before every action choice 1820 Q-values need to be calculated, leading to the conclusion that in the current implementation increasing discretization of actions leads to slower action choices and hence to higher training durations. Furthermore, the algorithm must estimate the value *Q(s,a)* for all combinations of state *s* (state of the forging ingot) and action *a* (combination of process parameters) and store it in the neural network. This mapping becomes more difficult as the number of combinations increases. This is why the step size of the action choices bite ratio (0.3 : 0.05 : 0.9 → 0.3 : 0.1 : 0.9) and height reduction (3% : 1% : 30% → 3% : 3% : 30%) were reduced. Through this adjustment, the number of possible actions per state decreases from 1820 to 350 representing another significant reduction in complexity.

Adaption of the Reward Function

Besides the high complexity of the optimization problem, the design of the reward function can be a reason for the slow convergence behavior as well as the miss of the desired grain size range. The definition of the reward function is one of the most important issues in RL, because the reward function does not only influence the achieved result, here e.g. the quality of the generated pass schedules, but also the training process itself. A well-defined reward function contributes to accelerating the convergence of RL algorithms and leads to more stable training processes. It is known that compared to sparse rewards, continuous reward functions promote the convergence of RL algorithms and lead to better training results. In the present context of pass schedule design, a sparse reward corresponds to feedback that is given only once per pass schedule after the last pass, whereas a continuous reward is calculated after every second pass. The current reward function already includes continuous rewards for utilizing the available press force and for minimizing the number of passes, but producing a specific grain size is only rated sparsely (cf. Fig. 3). It is hardly possible to describe the microstructure evolution during open-die forging processes inversely due to the incremental forming and interaction of various material physical effects like recrystallization and grain growth. For this reason, there is no feasible benchmark for the continuous evaluation of the ingot state over the process to achieve a desired grain size at the end of the process.

To overcome this hurdle, a new temperature-dependent continuous reward of the grain size is introduced, which should help the algorithm to reliably achieve the desired final grain size across all combinations of initial and target geometry. This reward structure is based on the assumption that the algorithm can only influence the microstructure as long as the ingot is sufficiently hot. As the workpiece cools during forging, processes such as recrystallization and grain growth are increasingly restricted, and hence, the influence of the agent on the microstructure decreases. This insight enables the creation of a new reward function $R_{GS,C}$ for rating the grain size continuously throughout the pass schedule. A downward-opening parabola whose maximum occurs at the target grain size forms the basis of this function. As can be seen in Fig. 5, on the one hand, the minimum reward is lowered with falling temperatures. On the other hand, the parabola is unilaterally stretched in the direction of larger grain sizes with raising temperatures. Both effects described, apply linearly between 900°C and 1100°C and represent the decreasing influence on the grain size (cf. Fig. 5). Outside this temperature range the reward $R_{GS,C}$ remains constant. The new continuous grain size reward only complements the existing sparse grain size reward. Hence, at the end of a pass schedule still a grain size reward inspired by Fig. 3 is calculated. It is to be noted that $R_{GS,C}$ is deliberately defined as negative. The punishment granted for each pass indirectly represents the goal to also minimize the number of passes, since higher amounts of passes lead to more grain size evaluations resulting in a higher overall punishment.

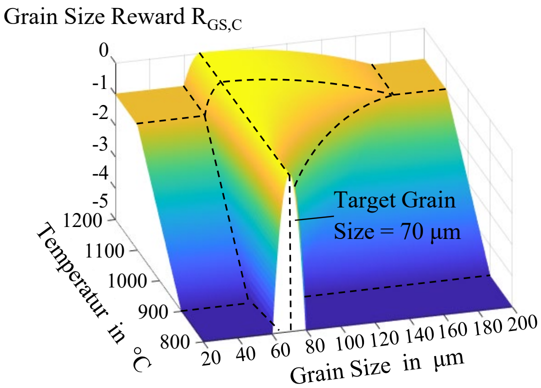

Fig. 5. Graphical representation of the continuous reward function for the grain size $R_{GS,C}$.

Retraining of the adapted RL algorithm – Continuous Grain Size Reward

To be able to assess the influence of the adjustments made, the adapted RL algorithm had to be retrained once again. This training process comprised 50,000 iterations that equal 50,000 designed and analyzed pass schedules, and took only 24 hours on four cores of an Intel® Xeon® E-2236 processor. Again, 400 different pass schedules were designed using the trained RL algorithm and the achieved mean austenitic grain sizes are summarized in Fig. 6.

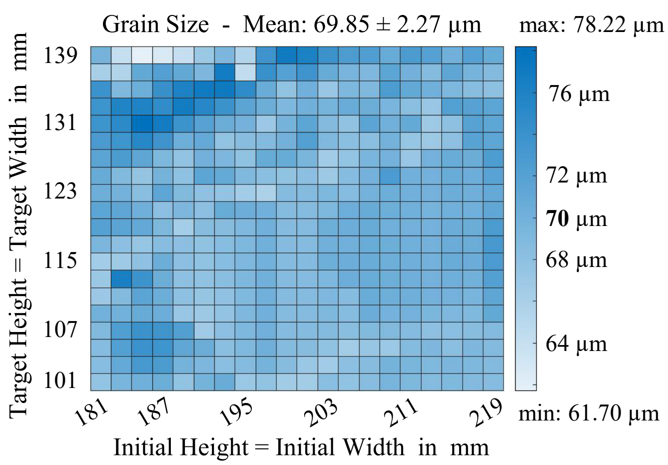

Fig. 6. Heatmap showing the produced average grain sizes of 400 different pass schedules designed by RL with a continuous grain size reward.

Averaged over all 400 pass schedules, the adapted RL algorithm produces a mean grain size of 69.85µm, which is very close to the target grain size of 70µm. The standard deviation results in 2.27µm while none of the 400 pass schedules produces a grain size outside the desired range. In addition, the algorithm continues to meet the goals of efficient open-die forging with an average number of passes of 5.55 and an average deviation of 5.5% between the predicted and nominal press force. At the same time, plant limits such as the maximum press force are observed in all pass schedules. Compared to the previous results from the RL algorithms with no and with sparse grain size rewards (cf. Table 1), a significant improvement in the quality of the designed forging processes was achieved.

Table 1. Summary of the pass schedule quality of all three RL algorithms.

RL algorithm	**no GS reward**	**sparse GS reward**	**continuous GS reward**
Average grain size [µm]	86.71 ± 3.82	70.99 ± 3.58	69.85 ± 2.27
Average number of passes	6.15	5.88	5.55
Avg. deviation nominal Force [%]	5.02	6.13	5.53
Avg. deviation from final height [%]	0.23	0.31	0

In general, the latest RL algorithm shows similar process design patterns compared to its predecessor with sparse grain size reward. Whenever possible, large height reductions are applied to reduce the number of passes. At the same time, break times and bite ratios are selected interdependently so that the press force is utilized well and the desired grain size is achieved. The greatest behavior change is present in the area in which the maximum permissible grain size of 80 µm was previously exceeded (see Fig. 4, red frame). Here, the adapted RL algorithm applies six passes instead of four while concentrating deformation in passes three and four. This leads to a different recrystallization timing, which shifts the final grain size into the desired range.

The significant improvement in the quality of the designed pass schedules, combined with the drastic reduction of the training duration from eight days to 24 hours, demonstrate the feasibility of RL for the design of optimized pass schedules in open-die forging with consideration of microstructure. After the training process, the design of one optimized pass schedule only takes about four seconds. This short calculation time enables the use of RL for the control of open-die forging processes when combined with a suitable process monitoring system (cf. [13]). If deviations from the target process route happen that would lead to e.g. an undesired change in final grain size, the monitoring system can detect those. Subsequently, a RL algorithm could be used to calculate a new optimized pass schedule in real time, starting from the intermediate state of the process and continuing to produce the desired final grain size. To achieve this, further changes to the RL algorithm are necessary, so that it can cope with sensor noise present in the real forging environment as well as a wider range of input states of the forging ingots.

Summary

In this publication, an existing RL algorithm for the design of efficient open-die forging processes was extended to successfully consider microstructure in the process design. In a first extension, a sparse grain size reward was introduced into the RL algorithm leading to an improved microstructure. The average grain size of 70.99 ± 3.58 µm was significantly closer to the target value of 70 µm compared to the original implementation (86.71 ± 3.82 µm). The goals of efficient open-die forging (produce target geometry, use available press force, reduce number of passes) were also met, although slightly worse than before (cf. Table 1). Since individual pass schedules still missed the desired range for the grain size as well as the target geometry, a new continuous reward structure depending on the ingot temperature was added in a second extension. Besides that, the complexity of the underlying optimization problem was reduced so that the RL algorithm is now able to design optimized pass schedules for all combinations of initial and target geometry. On average, the mean austenitic grain size after the forging resulted in 69.85 ± 2.27 µm while all designed pass schedules complied with the specified grain size limits. In addition, the adherence to the optimization objectives of efficient forging also improved, underlining the great potential of utilizing process knowledge for the improvement of reinforcement learning applications. Besides the raise of result quality the duration of the training process decreased from eight to one day and the design of individual optimized pass schedules still takes just a few seconds. This highlights not only the possibility to use RL for the fast design of optimized pass schedules, but also the potential to use RL for the control of open-die forging processes.

Outlook

In the future, forging experiments on the open-die forging press of the Institute of Metal Forming (IBF) as well as microstructure analysis are planned to validate the presented RL approach. In addition, after the extension, the algorithm is still able to design individual optimized pass schedules in a few seconds. This offers the potential to extend the RL algorithm from designing processes in advance to control open-die forgings with regard to microstructure. To achieve this, the RL has to be extended so that deviations from a nominal process, such as interruptions or varying initial temperatures, can be included in the training process and hence, be reacted to.

Acknowledgement.
This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 452539735.

References
[1] Information on https://www.sms-group.com/de-de/plants/open-die-forging-presses
[2] Information on https://easy2forge.com/
[3] A. Esteso, D. Peidro, J. Mula, M. Díaz-Madroñero, Reinforcement learning applied to production planning and control, Int. J. Prod. Res. (2022). https://doi.org/10.1080/00207543.2022.2104180
[4] C. Li, P. Zheng, Y. Yin, B. Wang, L. Wang, Deep reinforcement learning in smart manufacturing: A review and prospects, CIRP J. Manuf. Sci. Technol. 40 (2023) 75-101. http://doi.org/10.1016/j.cirpj.2022.11.003
[5] O. Gamal, M.I.P. Mohamed, C.G. Patel, H. Roth, Data-Driven Model-Free Intelligent Roll Gap Control of Bar and Wire Hot Rolling Process Using Reinforcement Learning, Int. J. Mech. Eng. Robot. Res. 10/7 (2021) 349-356.
[6] J. Dornheim, N. Link, P. Gumbsch, Model-free Adaptive Optimal Control of Episodic Fixed-horizon Manufacturing Processes Using Reinforcement Learning. Int. J. Control Autom. Syst. 18 (2020) 1593–1604
[7] C. Scheiderer, T. Thun, C. Idzik, A.F. Posada-Moreno, A. Krämer, J. Lohmar, G. Hirt, T. Meisen, Simulation-as-a-Service for Reinforcement Learning Applications by Example of Heavy Plate Rolling Processes, Procedia Manuf. 51 (2020) 897-903. https://doi.org/10.1016/j.promfg.2020.10.126
[8] N. Reinisch, F. Rudolph, S. Günther, D. Bailly, G. Hirt, Successful Pass Schedule Design in Open-Die Forging Using Double Deep Q-Learning, Processes 9 (2021) 1084. https://doi.org/10.3390/pr9071084
[9] A. Tomlinson, J.D. Stringer, Spread and elongation in flat tool forging, J. Iron Steel Inst. 193 (1959) 157-162.
[10] M. Wolfgarten, D. Rosenstock, F. Rudolph, G. Hirt, New approach for the optimization of pass-schedules in open-die forging, Int. J. Mater. Form. 12 (2019) 973-983. https://doi.org/10.1007/s12289-019-01471-w
[11] D. Recker, M. Franzke, G. Hirt, Fast models for online-optimization during open die forging, CIRP Annals – Manuf. Technol. 60 (2011) 295–298. http://doi.org/10.1016/j.cirp.2011.03.142
[12] K. Karhausen, R. Kopp, Model for integrated process and microstructure simulation in hot forming, Steel Res. 63 (1992) 247-256. https://doi.org/10.1002/srin.199200509
[13] F. Rudolph, M. Wolfgarten, V. Keray, G. Hirt, Optimization of Open-Die Forging Using Fast Models for Strain, Temperature, and Grain Size in the Context of an Assistance System, Forming the Future. The Minerals, Metals & Materials Series (2021) 1145-1159.