

Bi-phase Curated Reinforcement Learning

Kernel Infilling Stage



Kernel Generation Stage



Single Operations → Fused Operations → Model Structures

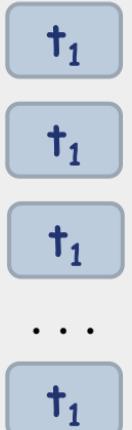
Easy

Data Scheduling

Hard

&
One-shot
Prompt

Policy
Model



Standardized Reward

Reward Model

Successful Compilation

Functional Correctness

Group
Computation

```
import torch
...
import torch.nn as nn
```

Prefix

```
elementwise_add_cpp_source = (
    ...
) elementwise_add = load_inline(
    ...
)
```

```
class ModelNew(nn.Module):
    ...
```

Suffix



PyTorch Reference
Efficient CUDA Kernel



DICE

Response Corresponding Trace

$$t_1 \triangleq (t_1(1), t_1(2), t_1(3), \dots, t_1(T))$$



Decoded Tokens at Each Step