



GnarlyMshtep / pass_at_k



Code

Issues 28

Pull requests 6

Projects 1

Wiki

Security



Edit

New issue



Get verl to work on deltaAI #67

Open

Edits ▾



I can't save models for whatever reason. Here is an example err

```

RUN AND DEB... D No Config... ⚙ ... distributed_c10d.py X
VARIABLES
Locals
> special variables
  async_op = False
  device_ids = None
> group = torch.distributed.distributed_
  opts = <torch.distributed.distributed_>
  work = <torch.distributed.distributed_>
Globals

Exception has occurred: RuntimeError (note: full exception trace is shown but execution is paused at: <module>)
[pytorch/third_party/gloo/gloo/transport/tcp/unbound_buffer.cc:81] Timed out waiting 60000ms for recv operation to complete
File "/u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/torch/distributed/distributed_c10d.py", line 4688, in barrier
  work.wait()
File "/u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/torch/distributed/c10d_logger.py", line 81, in wrapper
  return f(*args, **kwargs)
File "/u/aileyas/code/pass_at_k/verl/utils/checkpoint/fsdp_checkpoint_manager.py", line 304, in save_checkpoint
  torch.distributed.barrier()
File "/u/aileyas/code/pass_at_k/verl/workers/fsdp_workers.py", line 869, in save_checkpoint
  self.checkpoint_manager.save_checkpoint()
File "/u/aileyas/code/pass_at_k/verl/simple_controller/base/decorator_py", line 430, in inner
  return f(*args, **kwargs)
File "/u/aileyas/code/pass_at_k/verl/simple_controller/cnn/base_mx", line 701, in func
  return getattr(self.worker_dict[key], name)(*args, **kwargs)
File "/u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/roy/util/tracing/tracing_helper.py", line 461, in _resume_span
  return method(*self, *args, **kwargs)
File "/u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/roy/_private/function_manager.py", line 689, in actor_method_executor
  return method._remote_actor(*args, **kwargs)
File "/u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/roy/roylet.py", line 1865, in roy_roylet_execute_task.function_executor
  File "/u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/roy/roylet.py", line 1825, in roy_roylet_execute_task
  File "/u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/roy/roylet.py", line 2026, in roy_roylet_execute_task
  File "/u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/roy/roylet.py", line 1918, in roy_roylet_execute_task
  File "/u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/roy/_private/worker_py", line 974, in main_loop
    self.core_workers_run_task_loop()
File "/u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/roy/_private/workers/default_worker.py", line 321, in <module> (Current frame)
  worker.main_loop()
RuntimeError: [pytorch/third_party/gloo/gloo/transport/tcp/unbound_buffer.cc:81] Timed out waiting 60000ms for recv operation to complete

4689   # Otherwise, the backend has synced'ed at CPP level
4690
4691   def monitored_barrier(
4692     group: Optional[ProcessGroup] = GroupMember.WORLD,
4693     timeout=None,
4694     wait_all_ranks=False,
4695   ): ...
4696   """
4697   Synchronize processes similar to ``torch.distributed.barrier``, but consider a configurable timeout.
4698
4699   It is able to report ranks that did not pass this barrier within the provided timeout.
4700   Specifically, for non-zero ranks, will block until a send/recv is processed from rank 0.
4701   Rank 0 will block until all send/recv from other ranks are processed, and will report
4702   failures for ranks that failed to respond in time. Note that if one rank does not reach the
4703   PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
  Debugger client is connected

CALL STACK
/u/pytorch/vthrd/partylgologictrn
barrier distributed_c10d.py 4688
wrapper c10d_logger.py 81
save_checkpoint fsdp_checkpoint_m...
save_checkpoint fsdp_workers.py
inner decorator.py @3301
func base.py 7011
Show 8 More Stack Frames
> Thread-1 PAUSED
> Thread-2 __read_thread) PAUSED
> Thread-3 PAUSED

```

When using /u/aileyas/code/pass_at_k/custom/run_scripts/10/27/dapo_newreward.sh

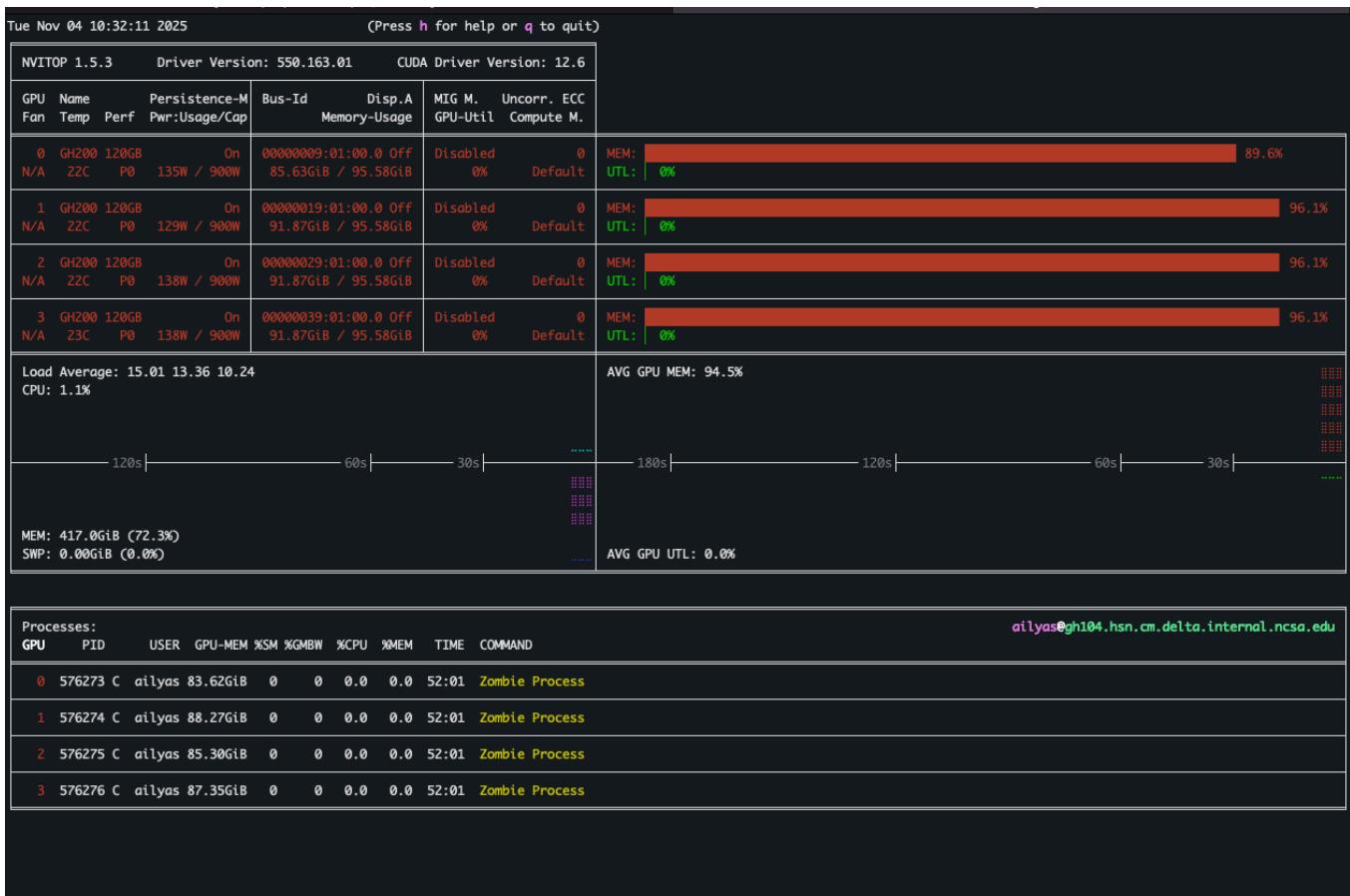
terminal output

```
[notify] 🚀 Detected W&B URL: https://wandb.ai/matan-shtepel-carnegie-mellon-university/dapo/runs/4yw9kh8a
(TaskRunner pid=553716) wandb: WARNING Symlinked 0 file into the W&B run directory, call
```



```
wandb.save again to sync new files.  
Training Progress: 0% | 0/990 [00:00<?, ?it/s]  
(WorkerDict pid=576273) INFO:2025-11-04 09:53:11,110: [Rank 0] Saved model to  
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-  
math_baseline_4096_matan_reward_fn/global_step_1/actor/model_world_size_4_rank_0.pt  
(WorkerDict pid=576274) /u/ailyas/miniconda3/envs/deltaai-env/lib/python3.10/site-  
packages/torch/distributed/fsdp/fully_sharded_data_parallel.py:680: FutureWarning:  
FSDP.state_dict_type() and FSDP.set_state_dict_type() are being deprecated. Please use  
APIs, get_state_dict() and set_state_dict(), which can support different parallelisms,  
FSDP1, FSDP2, DDP. API doc:  
https://pytorch.org/docs/stable/distributed.checkpoint.html#torch.distributed.checkpoint.start.Tutorial: https://pytorch.org/tutorials/recipes/distributed\_checkpoint\_recipe.html.  
(WorkerDict pid=576274) warnings.warn(  
(WorkerDict pid=576273) INFO:2025-11-04 09:54:20,831: [Rank 0] Saved optim to  
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-  
math_baseline_4096_matan_reward_fn/global_step_1/actor/optim_world_size_4_rank_0.pt  
(WorkerDict pid=576273) INFO:2025-11-04 09:54:20,854: [Rank 0] Saved extra_state to  
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-  
math_baseline_4096_matan_reward_fn/global_step_1/actor/extrastate_world_size_4_rank_0.pt  
(WorkerDict pid=576273) INFO:2025-11-04 09:54:21,104: [Rank 0] Saved model config and  
tokenizer class to /work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-  
math_baseline_4096_matan_reward_fn/global_step_1/actor/huggingface  
(WorkerDict pid=576274) INFO:2025-11-04 09:57:23,613: [Rank 1] Saved model to  
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-  
math_baseline_4096_matan_reward_fn/global_step_1/actor/model_world_size_4_rank_1.pt  
(WorkerDict pid=576275) INFO:2025-11-04 09:59:01,449: [Rank 2] Saved model to  
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-  
math_baseline_4096_matan_reward_fn/global_step_1/actor/model_world_size_4_rank_2.pt  
(WorkerDict pid=576273) /u/ailyas/miniconda3/envs/deltaai-env/lib/python3.10/site-  
packages/pydevd_plugins/__init__.py:2: UserWarning: pkg_resources is deprecated as an  
API. See https://setuptools.pypa.io/en/latest/pkg\_resources.html. The pkg_resources  
package is slated for removal as early as 2025-11-30. Refrain from using this package or  
pin to Setuptools<81.  
(WorkerDict pid=576273) __import__("pkg_resources").declare_namespace(__name__)  
(WorkerDict pid=576273) Ray debugger is listening on 141.142.249.220:56515  
(WorkerDict pid=576273) Waiting for debugger to attach (see  
https://docs.ray.io/en/latest/ray-observability/ray-distributed-debugger.html)...  
  
[notify] ! Detected trigger: Ray debugger is listening  
(WorkerDict pid=576273) Debugger client is connected
```

and I am left with these zombie processes and have to allocate a new machine



I also use Claude to test write 20GB to `$HFH/models/checkpoints` folder and it worked just fine!

[Create sub-issue](#) ▾



- 🕒 GnarlyMshtep moved this to Experiments-active in [matan-all-board](#) 3 days ago
- 🕒 GnarlyMshtep added this to [matan-all-board](#) 3 days ago
- 🕒 GnarlyMshtep moved this from Experiments-active to Experiments-inactive in [matan-all-board](#) 2 days ago
- 🕒 GnarlyMshtep moved this from Experiments-inactive to Experiments-active in [matan-all-board](#) yesterday
- 🕒 GnarlyMshtep transferred this issue from [GnarlyMshtep/cot-monitor](#) yesterday



GnarlyMshtep yesterday · edited by GnarlyMshtep

Edits ▾ [Author](#) ...

Testing things which could have caused this

Reward function complexity

I am using a more complex reward function, which uses ray workers. I think this is the biggest thing I changed around the time this stopped working. Let me just use the basic `math_verify` reward.

Basically identical err

```
(TaskRunner pid=1348941) wandb: 🚀 View run at https://wandb.ai/matan-shtepel-carnegie-mellon-university/dapo/runs/aovnj2tv

[notify] DEBUG: Matched wandb run URL in line: (TaskRunner pid=1348941) wandb: 🚀 View run at https://wandb.ai/matan-shtepel-carnegie-mellon-university/dapo/runs/aovnj2tv
[notify] DEBUG: Extracted URL: https://wandb.ai/matan-shtepel-carnegie-mellon-university/dapo/runs/aovnj2tv
[notify] DEBUG: Sending wandb notification to https://ntfy.sh/mshtepel-wandburl-ml-runs
[notify] DEBUG: Data: {'event': 'wandb', 'run_id': '20251104_105210_8289', 'wandb_url': 'https://wandb.ai/matan-shtepel-carnegie-mellon-university/dapo/runs/aovnj2tv', 'timestamp': '2025-11-04T10:55:47.538658'}

[notify] 🚀 Detected W&B URL: https://wandb.ai/matan-shtepel-carnegie-mellon-university/dapo/runs/aovnj2tv
(TaskRunner pid=1348941) wandb: WARNING Symlinked 0 file into the W&B run directory, call wandb.save again to sync new files.
Training Progress: 0% | 0/990 [00:00<?, ?it/s]
(WorkerDict pid=1371189) INFO:2025-11-04 11:00:57,047:[Rank 0] Saved model to /work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/model_world_size_4_rank_0.pt
(WorkerDict pid=1371192) /u/ailyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/torch/distributed/fsdp/fully_sharded_data_parallel.py:680: FutureWarning: FSDP.state_dict_type() and FSDP.set_state_dict_type() are being deprecated. Please use APIs, get_state_dict() and set_state_dict(), which can support different parallelisms, FSDP1, FSDP2, DDP. API doc:
https://pytorch.org/docs/stable/distributed.checkpoint.html#torch.distributed.checkpoint.sta
.Tutorial: https://pytorch.org/tutorials/recipes/distributed_checkpoint_recipe.html .
[repeated 3x across cluster]
(WorkerDict pid=1371192) warnings.warn( [repeated 3x across cluster]
(WorkerDict pid=1371189) INFO:2025-11-04 11:02:01,738:[Rank 0] Saved optim to /work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/optim_world_size_4_rank_0.pt
(WorkerDict pid=1371189) INFO:2025-11-04 11:02:01,764:[Rank 0] Saved extra_state to /work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/extrastate_world_size_4_rank_0.pt
(WorkerDict pid=1371189) INFO:2025-11-04 11:02:02,014:[Rank 0] Saved model config and tokenizer class to /work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/huggingface
(WorkerDict pid=1371191) INFO:2025-11-04 11:02:22,352:[Rank 2] Saved model to /work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/model_world_size_4_rank_2.pt
(WorkerDict pid=1371190) INFO:2025-11-04 11:04:13,290:[Rank 1] Saved model to /work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/model_world_size_4_rank_1.pt
(WorkerDict pid=1371189) /u/ailyas/miniconda3/envs/deltaai-env/lib/python3.10/site-packages/pydevd_plugins/__init__.py:2: UserWarning: pkg_resources is deprecated as an API. See https://setuptools.pypa.io/en/latest/pkg_resources.html. The pkg_resources package is slated for removal as early as 2025-11-30. Refrain from using this package or pin to Setuptools<81.
(WorkerDict pid=1371189) __import__("pkg_resources").declare_namespace(__name__)
```

```
(WorkerDict pid=1371189) Ray debugger is listening on 141.142.249.224:32861
(WorkerDict pid=1371189) Waiting for debugger to attach (see
https://docs.ray.io/en/latest/ray-observability/ray-distributed-debugger.html)...
```

```
[notify] ⚠️ Detected trigger: Ray debugger is listening
(WorkerDict pid=1371189) Debugger client is connected
```



GnarlyMshtep added a commit that references this issue [yesterday](#)

[#67 simpler reward](#)

2cadb40



GnarlyMshtep yesterday

Author

...

Changes to Conda env.

```
(deltaai-env) → pass_at_k git:(57-code-baseline) conda list --revisions | tail -n 10
+libcurand-dev-10.3.7.77 (nvidia/linux-aarch64)
+libcusolver-11.7.1.2 (nvidia/linux-aarch64)
+libcusolver-dev-11.7.1.2 (nvidia/linux-aarch64)
+libcusparse-12.5.4.2 (nvidia/linux-aarch64)
+libcusparse-dev-12.5.4.2 (nvidia/linux-aarch64)
+libffi-3.4.4 (defaults/linux-aarch64)
+libgcc-14.2.0 (conda-forge/linux-aarch64)
+libgcc-devel_linux-aarch64-11.2.0 (defaults/linux-aarch64)
+libgcc-ng-14.2.0 (conda-forge/linux-aarch64)
+libglib-2.78.4 (defaults/linux-aarch64)
+libgomp-14.2.0 (conda-forge/linux-aarch64)
+libiconv-1.16 (defaults/linux-aarch64)
+libnpp-12.2.5.30 (nvidia/linux-aarch64)
+libnpp-dev-12.2.5.30 (nvidia/linux-aarch64)
+libnvfatbin-12.6.77 (nvidia/linux-aarch64)
+libnvfatbin-dev-12.6.77 (nvidia/linux-aarch64)
+libnvjitlink-12.6.85 (nvidia/linux-aarch64)
+libnvjitlink-dev-12.6.85 (nvidia/linux-aarch64)
+libnvjpeg-12.3.1.117 (nvidia/linux-aarch64)
+libnvjpeg-dev-12.3.1.117 (nvidia/linux-aarch64)
+libpng-1.6.39 (defaults/linux-aarch64)
+libstdcxx-devel_linux-aarch64-11.2.0 (defaults/linux-aarch64)
+libstdcxx-ng-11.2.0 (defaults/linux-aarch64)
+libuuid-1.41.5 (defaults/linux-aarch64)
+libxcb-1.15 (defaults/linux-aarch64)
+libxkbcommon-1.0.1 (defaults/linux-aarch64)
+libxml2-2.13.5 (defaults/linux-aarch64)
+ncurses-6.4 (defaults/linux-aarch64)
+nsight-compute-2024.3.2.3 (nvidia/linux-aarch64)
+nspr-4.35 (defaults/linux-aarch64)
+nss-3.89.1 (defaults/linux-aarch64)
```



```
+openssl-3.5.2 (conda-forge/linux-aarch64)
+pcre2-10.42 (defaults/linux-aarch64)
+python-3.10.15 (defaults/linux-aarch64)
+readline-8.2 (defaults/linux-aarch64)
+sqlite-3.45.3 (defaults/linux-aarch64)
+sysroot_linux-aarch64-2.26 (defaults/noarch)
+tk-8.6.14 (defaults/linux-aarch64)
+tzdata-2024b (defaults/noarch)
+xz-5.4.6 (defaults/linux-aarch64)
+zlib-1.2.13 (defaults/linux-aarch64)
```

2025-10-16 12:57:53 (rev 1)

```
ca-certificates {2025.8.3 (conda-forge/noarch) -> 2025.10.5 (conda-forge/noarch)}
openssl {3.5.2 (conda-forge/linux-aarch64) -> 3.5.4 (conda-forge/linux-aarch64)}
+conda-pack-0.8.1 (conda-forge/noarch)
+setuptools-80.9.0 (conda-forge/noarch)
```

2025-10-17 11:03:36 (rev 2)

```
ncurses {6.4 (defaults/linux-aarch64) -> 6.5 (conda-forge/linux-aarch64)}
+libstdcxx-14.2.0 (conda-forge/linux-aarch64)
+nodejs-11.14.0 (conda-forge/linux-aarch64)
+nvidia-ml-py-12.575.51 (conda-forge/noarch)
+nvitop-1.5.3 (conda-forge/noarch)
+psutil-5.9.0 (conda-forge/linux-aarch64)
+python_abi-3.10 (conda-forge/linux-aarch64)
+uv-0.9.3 (conda-forge/linux-aarch64)
```

2025-10-17 11:12:07 (rev 3)

```
-nodejs-11.14.0 (conda-forge/linux-aarch64)
```

2025-10-17 11:12:49 (rev 4)

```
+nodejs-11.14.0 (conda-forge/linux-aarch64)
```

2025-10-17 12:16:07 (rev 5)

```
-nodejs-11.14.0 (conda-forge/linux-aarch64)
```

2025-10-17 12:17:25 (rev 6)

```
+nodejs-11.14.0 (conda-forge/linux-aarch64)
```

2025-10-27 12:29:37 (rev 7)

```
+asttokens-3.0.0 (conda-forge/noarch)
+decorator-5.2.1 (conda-forge/noarch)
+exceptiongroup-1.3.0 (conda-forge/noarch)
+executing-2.2.1 (conda-forge/noarch)
+ipdb-0.13.13 (conda-forge/noarch)
+ipython-8.37.0 (conda-forge/noarch)
+jedi-0.19.2 (conda-forge/noarch)
+matplotlib-inline-0.2.1 (conda-forge/noarch)
+parso-0.8.5 (conda-forge/noarch)
+pexpect-4.9.0 (conda-forge/noarch)
+pickleshare-0.7.5 (conda-forge/noarch)
+prompt-toolkit-3.0.52 (conda-forge/noarch)
+ptyprocess-0.7.0 (conda-forge/noarch)
+pure_eval-0.2.3 (conda-forge/noarch)
+pygments-2.19.2 (conda-forge/noarch)
+stack_data-0.6.3 (conda-forge/noarch)
```

```
+toml-0.10.2 (conda-forge/noarch)
+traitlets-5.14.3 (conda-forge/noarch)
+typing_extensions-4.15.0 (conda-forge/noarch)
+wcwidth-0.2.14 (conda-forge/noarch)

2025-10-28 09:17:42 (rev 8)
+pip-25.2 (conda-forge/noarch)
+wheel-0.45.1 (conda-forge/noarch)
```

```
2025-10-28 11:52:39 (rev 9)
+htop-3.4.1 (conda-forge/linux-aarch64)
+libnl-3.11.0 (conda-forge/linux-aarch64)
```

I thinkg things still worked on the 17th but did not work since the 28th. So I am trying

```
(deltaai-env) → pass_at_k git:(57-code-baseline) conda install --revision 6
```



and then I will run again.



GnarlyMshtep 20 hours ago

Author

...

Same issue:

```
Training Progress: 0% | 0/990 [00:00<?, ?it/s]
(WorkerDict pid=1723277) INFO:2025-11-04 15:19:45,632:[Rank 0] Saved model to
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/model_world_size_4_rank_0.pt
(WorkerDict pid=1723277) /u/ailyas/miniconda3/envs/deltaai-env/lib/python3.10/site-
packages/torch/distributed/fsdp/fully_sharded_data_parallel.py:680: FutureWarning:
FSDP.state_dict_type() and FSDP.set_state_dict_type() are being deprecated. Please use
APIs, get_state_dict() and set_state_dict(), which can support different parallelisms,
FSDP1, FSDP2, DDP. API doc:
https://pytorch.org/docs/stable/distributed.checkpoint.html#torch.distributed.checkpoint.sta
.Tutorial: https://pytorch.org/tutorials/recipes/distributed_checkpoint_recipe.html .
(WorkerDict pid=1723277) warnings.warn(
(WorkerDict pid=1723277) INFO:2025-11-04 15:20:55,474:[Rank 0] Saved optim to
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/optim_world_size_4_rank_0.pt
(WorkerDict pid=1723277) INFO:2025-11-04 15:20:55,496:[Rank 0] Saved extra_state to
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/extrastate_world_size_4_rank_0.pt
(WorkerDict pid=1723277) INFO:2025-11-04 15:20:55,682:[Rank 0] Saved model config and
tokenizer class to /work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/huggingface
(WorkerDict pid=1723279) INFO:2025-11-04 15:21:14,963:[Rank 2] Saved model to
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/model_world_size_4_rank_2.pt
(WorkerDict pid=1723280) INFO:2025-11-04 15:21:37,206:[Rank 3] Saved model to
```



```
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/model_world_size_4_rank_3.pt
(WorkerDict pid=1723278) INFO:2025-11-04 15:22:30,638: [Rank 1] Saved model to
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/model_world_size_4_rank_1.pt
(WorkerDict pid=1723277) /u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-
packages/pydevd_plugins/__init__.py:2: UserWarning: pkg_resources is deprecated as an
API. See https://setuptools.pypa.io/en/latest/pkg_resources.html. The pkg_resources
package is slated for removal as early as 2025-11-30. Refrain from using this package or
pin to Setuptools<81.
(WorkerDict pid=1723277) __import__("pkg_resources").declare_namespace(__name__)
(WorkerDict pid=1723277) Ray debugger is listening on 141.142.249.16:35289
(WorkerDict pid=1723277) Waiting for debugger to attach (see
https://docs.ray.io/en/latest/ray-observability/ray-distributed-debugger.html)...
```

[notify] ! Detected trigger: Ray debugger is listening

```
(WorkerDict pid=1723280) INFO:2025-11-04 15:42:00,479: [Rank 3] Saved optim to
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/optim_world_size_4_rank_3.pt
(WorkerDict pid=1723280) INFO:2025-11-04 15:42:00,483: [Rank 3] Saved extra_state to
/work/nvme/bfov/models/checkpoints/dapo/Qwen2_5-7B_dapo-
math_baseline_4096_matan_reward_fn/global_step_1/actor/extra_state_world_size_4_rank_3.pt
(WorkerDict pid=1723280) /u/aileyas/miniconda3/envs/deltaai-env/lib/python3.10/site-
packages/pydevd_plugins/__init__.py:2: UserWarning: pkg_resources is deprecated as an
API. See https://setuptools.pypa.io/en/latest/pkg_resources.html. The pkg_resources
package is slated for removal as early as 2025-11-30. Refrain from using this package or
pin to Setuptools<81.
(WorkerDict pid=1723280) __import__("pkg_resources").declare_namespace(__name__)
(WorkerDict pid=1723280) Ray debugger is listening on 141.142.249.16:50929
(WorkerDict pid=1723280) Waiting for debugger to attach (see
https://docs.ray.io/en/latest/ray-observability/ray-distributed-debugger.html)...
```

Here is the full script in case it helps.

```
if [ "$SKIP_VALIDATION" = false ]; then
    python3 validate_env.py \
        --train-path "$train_path" \
        --test-path "$test_path" \
        --model-path "$model_path" \
        --reward-path "$reward_path" \
        --reward-name "$reward_name" \
        --n-gpu "$n_gpu" \
        --batch-size "$batch_size" \
        --mini-batch-size "$mini_batch_size" \
        --n-rollout "$n_rollout" \
        --micro-batch-size-per-gpu "$micro_batch_size_per_gpu_prob_ignored" \
        --adv-estimator "$adv_est" \
        --norm-by-std "$norm_by_std" \
        --proj-name "$proj_name" \
        --exp-name "$exp_name"
else
    echo "Skipping validation (passed -y flag)"
fi
# False gradient checkpointing -- revert if weird err
```



```
python3 -m verl.trainer.main_ppo \
    algorithm.adv_estimator=$adv_est \
    algorithm.norm_adv_by_std_in_group=$norm_by_std \
    data.train_files="$train_files" \
    data.val_files="$test_files" \
    data.train_batch_size=512 \
    data.max_prompt_length=512 \
    data.max_response_length=$max_response_length \
    data.filter_overlong_prompts=True \
    data.shuffle=True \
    data.truncation='error' \
    actor_rollout_ref.model.path=$model_path \
    actor_rollout_ref.actor.optim.lr=1e-6 \
    actor_rollout_ref.actor.optim.lr_warmup_steps=15 \
    actor_rollout_ref.model.use_remove_padding=True \
    actor_rollout_ref.actor.ppo_mini_batch_size=256 \
    \
    actor_rollout_ref.actor.ppo_micro_batch_size_per_gpu=$micro_batch_size_per_gpu_prob_ignored \
    \
    actor_rollout_ref.actor.use_dynamic_bsz=True \
    actor_rollout_ref.actor.ppo_max_token_len_per_gpu=$max_token_len_per_gpu \
    actor_rollout_ref.ref.log_prob_use_dynamic_bsz=True \
    actor_rollout_ref.ref.log_prob_max_token_len_per_gpu=$max_token_len_per_gpu \
    actor_rollout_ref.rollout.log_prob_use_dynamic_bsz=True \
    actor_rollout_ref.rollout.log_prob_max_token_len_per_gpu=$max_token_len_per_gpu \
    actor_rollout_ref.actor.use_kl_loss=True \
    actor_rollout_ref.actor.kl_loss_coef=0.001 \
    actor_rollout_ref.actor.kl_loss_type=low_var_kl \
    actor_rollout_ref.actor.entropy_coeff=0 \
    actor_rollout_ref.model.enable_gradient_checkpointing=True \
    actor_rollout_ref.actor.fsdp_config.param_offload=False \
    actor_rollout_ref.actor.fsdp_config.optimizer_offload=False \
    \
    actor_rollout_ref.rollout.log_prob_micro_batch_size_per_gpu=$micro_batch_size_per_gpu_prob_i \
    \
    actor_rollout_ref.rollout.tensor_model_parallel_size=2 \
    actor_rollout_ref.rollout.name=vllm \
    actor_rollout_ref.rollout.gpu_memory_utilization=0.55 \
    actor_rollout_ref.rollout.n=4 \
    \
    actor_rollout_ref.ref.log_prob_micro_batch_size_per_gpu=$micro_batch_size_per_gpu_prob_ignor \
    \
    actor_rollout_ref.ref.fsdp_config.param_offload=True \
    algorithm.use_kl_in_reward=False \
    custom_reward_function.name=$reward_name \
    custom_reward_function.path=$reward_path \
    trainer.val_before_train=False \
    trainer.critic_warmup=0 \
    trainer.logger='["console", "wandb"]' \
    trainer.project_name=$proj_name \
    trainer.experiment_name=$exp_name \
    trainer.n_gpus_per_node=$n_gpu \
    trainer.nnodes=1 \
    trainer.save_freq=1 \
    +trainer.remove_previous_ckpt_in_save=True \
```

```
trainer.resume_mode=auto\
trainer.test_freq=30 \
"custom/run_scripts/10/27/dapo_newreward.sh" 126L, 4534B
```



GnarlyMshtep 20 hours ago · edited by GnarlyMshtep

Edits ▾

Author

...

I'll do one more experiment -- does 1 GPU 1B model work?

Yes it works. Commit below confirms and this [wandb](#). Here is the working script:

```
set -eux

# Check for -y flag to skip validation
SKIP_VALIDATION=false
if [[ "$*" == *"-y"* ]]; then
    SKIP_VALIDATION=true
    # Remove -y from arguments to pass to main_ppo
    set -- "${@//-/y/}"
fi

if [ -e "core" ]; then
    rm core
fi
datasetname=dapo-math
train_path=$HF_HOME/data/$datasetname/train.parquet
test_path=$HF_HOME/data/$datasetname/test.parquet

train_files="['$train_path']"
test_files="['$test_path']"

# reward_name=matan_reward_func
# reward_path=deepmath_utils/reward_utils/reward_func_2.py
reward_name=compute_score_math_boxed
reward_path=custom/reward/reward_utils.py

modelname=Qwen2_5-1_5B
model_path=$HF_HOME/models/$modelname

max_token_len_per_gpu=28000
max_response_length=4096

n_gpu=1

proj_name='dapo'
exp_name="${modelname}_${datasetname}_baseline_${max_response_length}_matan_reward_fn"
```

```

batch_size=512
mini_batch_size=256
n_rollout=4
micro_batch_size_per_gpu_prob_ignored=32

adv_est=grpo
norm_by_std=False

if [ "$SKIP_VALIDATION" = false ]; then
    python3 validate_env.py \
        --train-path "$train_path" \
        --test-path "$test_path" \
        --model-path "$model_path" \
        --reward-path "$reward_path" \
        --reward-name "$reward_name" \
        --n-gpu "$n_gpu" \
        --batch-size "$batch_size" \
        --mini-batch-size "$mini_batch_size" \
        --n-rollout "$n_rollout" \
        --micro-batch-size-per-gpu "$micro_batch_size_per_gpu_prob_ignored" \
        --adv-estimator "$adv_est" \
        --norm-by-std "$norm_by_std" \
        --proj-name "$proj_name" \
        --exp-name "$exp_name"
else
    echo "Skipping validation (passed -y flag)"
fi
# False gradient checkpointing -- revert if weird err

python3 -m verl.trainer.main_ppo \
    algorithm.adv_estimator=$adv_est \
    algorithm.norm_adv_by_std_in_grpo=$norm_by_std \
    data.train_files="$train_files" \
    data.val_files="$test_files" \
    data.train_batch_size=512 \
    data.max_prompt_length=512 \
    data.max_response_length=$max_response_length \
    data.filter_overlong_prompts=True \
    data.shuffle=True \
    data.truncation='error' \
    actor_rollout_ref.model.path=$model_path \
    actor_rollout_ref.actor.optim.lr=1e-6 \
    actor_rollout_ref.actor.optim.lr_warmup_steps=15 \
    actor_rollout_ref.model.use_remove_padding=True \
    actor_rollout_ref.actor.ppo_mini_batch_size=256 \
    actor_rollout_ref.actor.ppo_micro_batch_size_per_gpu=$micro_batch_size_per_gpu_prob_igno
actor_rollout_ref.actor.use_dynamic_bsz=True \
    actor_rollout_ref.actor.ppo_max_token_len_per_gpu=$max_token_len_per_gpu \
    actor_rollout_ref.ref.log_prob_use_dynamic_bsz=True \
    actor_rollout_ref.ref.log_prob_max_token_len_per_gpu=$max_token_len_per_gpu \
    actor_rollout_ref.rollout.log_prob_use_dynamic_bsz=True \
    actor_rollout_ref.rollout.log_prob_max_token_len_per_gpu=$max_token_len_per_gpu \
    actor_rollout_ref.actor.use_kl_loss=True \
    actor_rollout_ref.actor.kl_loss_coef=0.001 \

```

```
actor_rollout_ref.actor.kl_loss_type=low_var_kl \
actor_rollout_ref.actor.entropy_coeff=0 \
actor_rollout_ref.model.enable_gradient_checkpointing=True \
actor_rollout_ref.actor.fsdp_config.param_offload=False \
actor_rollout_ref.actor.fsdp_config.optimizer_offload=False \
actor_rollout_ref.rollout.log_prob_micro_batch_size_per_gpu=$micro_batch_size_per_gpu_pr \
actor_rollout_ref.rollout.tensor_model_parallel_size=1 \
actor_rollout_ref.rollout.name=vllm \
actor_rollout_ref.rollout.gpu_memory_utilization=0.55 \
actor_rollout_ref.rollout.n=4 \
actor_rollout_ref.ref.log_prob_micro_batch_size_per_gpu=$micro_batch_size_per_gpu_prob_i \
actor_rollout_ref.ref.fsdp_config.param_offload=True \
algorithm.use_kl_in_reward=False \
custom_reward_function.name=$reward_name \
custom_reward_function.path=$reward_path \
trainer.val_before_train=False \
trainer.critic_warmup=0 \
trainer.logger='["console","wandb"]' \
trainer.project_name=$proj_name \
trainer.experiment_name=$exp_name \
trainer.n_gpus_per_node=$n_gpu \
trainer.nnodes=1 \
trainer.save_freq=1 \
+trainer.remove_previous_ckpt_in_save=True \
trainer.resume_mode=auto \
trainer.test_freq=30 \
+trainer.rollout_dump_freq=1 \
trainer.rollout_data_dir="$HFH/rollouts/${proj_name}/${exp_name}/train" \
trainer.validation_data_dir="$HFH/rollouts/${proj_name}/${exp_name}/val" \
trainer.default_local_dir="$HFH/models/checkpoints/${proj_name}/${exp_name}" \
trainer.total_epochs=15 $@
```



GnarlyMshtep added a commit that references this issue [2 minutes ago](#)

[#67](#) 1B DAPo script trains and model saving works

9af7e66



Add a comment

Write

Preview

H B I | ≡ <> ⌂ | ≡ ½ ≈ ⌂ | @ ⌂ ↶ ↷

Use Markdown to format your comment

 Paste, drop, or click to add files Close issue

Comment

 Remember, contributions to this repository should follow its [contributing guidelines](#).

Metadata

Assignees

No one - Assign yourself

 Assign to Copilot

Labels

No labels



Projects

 matan-all-boardStatus  Experiments-active

Milestone

No milestone



Relationships

None yet



Development

 Code with agent mode

[Create a branch](#) for this issue or link a pull request.

Notifications

Customize

 Unsubscribe

You're receiving notifications because you're subscribed to this thread.

Participants

 Transfer issue Duplicate issue

11/5/25, 12:54 PM

Get verl to work on deltaAI · Issue #67 · GnarlyMshtep/pass_at_k

 Lock conversation

 Pin issue

 Delete issue