SFT 학습 개선: 학습-평가 Objective 일치

개요

기존 SFT는 생성(decoding)으로 학습하지만, 평가(mathQA)는 **multiple choice**로 정답 옵션의 log-likelihood가 가장 높은지 확인합니다.

이 노트북은 **학습과 평가의 objective를 일치**시키기 위해:

1. grade-school-math에서 **1500개 샘플** 사용
2. Rule 기반 파이썬으로 각 문제당 **5개 옵션** 생성 (정답 1개 + 오답 4개)
3. 모델이 각 옵션 continuation에 부여하는 **log-likelihood** 계산
4. **Cross-entropy (softmax over options)**로 정답 옵션이 가장 높아지도록 학습

1. 환경 설정

In [2]:
```
# 핵심 라이브러리 설치 (버전 명시)

# 기본 라이브러리 설치
!pip install transformers>=4.45.0 bitsandbytes>=0.44.0
!pip install --upgrade triton  # torch 2.9 호환 (2.2.0 고정 시 triton.backends 오류)
!pip install datasets==2.21.0
!pip install peft==0.12.0
!pip install trl==0.9.6
!pip install scipy==1.13.1
# !pip install numpy pandas
!pip install numpy --no-cache-dir
!pip install wandb
!pip install --upgrade "accelerate>=1.7.0"
!pip install --upgrade triton
```

```
Requirement already satisfied: triton in /usr/local/lib/python3.12/dist-packages (3.5.0)
Collecting triton
  Downloading triton-3.6.0-cp312-cp312-manylinux_2_27_x86_64.manylinux_2_28_x86_64.whl.metadata (1.7 kB)
Downloading triton-3.6.0-cp312-cp312-manylinux_2_27_x86_64.manylinux_2_28_x86_64.whl (188.3 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 188.3/188.3 MB 11.9 MB/s eta 0:00:0000:0100:01
Installing collected packages: triton
  Attempting uninstall: triton
    Found existing installation: triton 3.5.0
    Uninstalling triton-3.5.0:
      Successfully uninstalled triton-3.5.0
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behav
iour is the source of the following dependency conflicts.
torch 2.9.0+cu126 requires triton==3.5.0; platform_system == "Linux", but you have triton 3.6.0 which is incompatibl
e.
Successfully installed triton-3.6.0
```

```
Collecting datasets==2.21.0
  Downloading datasets-2.21.0-py3-none-any.whl.metadata (21 kB)
Requirement already satisfied: filelock in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0) (3.20.3)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0) (2.0.2)
Requirement already satisfied: pyarrow>=15.0.0 in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0) (1
8.1.0)
Requirement already satisfied: dill<0.3.9,>=0.3.0 in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0)
(0.3.8)
Requirement already satisfied: pandas in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0) (2.2.2)
Requirement already satisfied: requests>=2.32.2 in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0)
(2.32.4)
Requirement already satisfied: tqdm>=4.66.3 in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0) (4.67.
1)
Requirement already satisfied: xxhash in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0) (3.6.0)
Requirement already satisfied: multiprocess in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0) (0.70.
16)
Collecting fsspec<=2024.6.1,>=2023.1.0 (from fsspec[http]<=2024.6.1,>=2023.1.0->datasets==2.21.0)
  Downloading fsspec-2024.6.1-py3-none-any.whl.metadata (11 kB)
Requirement already satisfied: aiohttp in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0) (3.13.3)
Requirement already satisfied: huggingface-hub>=0.21.2 in /usr/local/lib/python3.12/dist-packages (from datasets==2.2
1.0) (0.36.0)
Requirement already satisfied: packaging in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0) (25.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.12/dist-packages (from datasets==2.21.0) (6.0.3)
Requirement already satisfied: aiohappyeyeballs>=2.5.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp->data
sets==2.21.0) (2.6.1)
Requirement already satisfied: aiosignal>=1.4.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets==
2.21.0) (1.4.0)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets==2.2
1.0) (25.4.0)
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets==
2.21.0) (1.8.0)
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets
==2.21.0) (6.7.1)
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets==
2.21.0) (0.4.1)
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets==
2.21.0) (1.22.0)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.12/dist-packages (from huggingfac
e-hub>=0.21.2->datasets==2.21.0) (4.15.0)
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub>
=0.21.2->datasets==2.21.0) (1.2.0)
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests>=2.
32.2->datasets==2.21.0) (3.4.4)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests>=2.32.2->datase
ts==2.21.0) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests>=2.32.2->
datasets==2.21.0) (2.5.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests>=2.32.2->
datasets==2.21.0) (2026.1.4)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.12/dist-packages (from pandas->datase
ts==2.21.0) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.12/dist-packages (from pandas->datasets==2.21.
0) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.12/dist-packages (from pandas->datasets==2.2
1.0) (2025.3)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.12/dist-packages (from python-dateutil>=2.8.2->pand
as->datasets==2.21.0) (1.17.0)
Downloading datasets-2.21.0-py3-none-any.whl (527 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 527.3/527.3 kB 4.2 MB/s eta 0:00:00a 0:00:01
Downloading fsspec-2024.6.1-py3-none-any.whl (177 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 177.6/177.6 kB 4.0 MB/s eta 0:00:00a 0:00:01
Installing collected packages: fsspec, datasets
  Attempting uninstall: fsspec
    Found existing installation: fsspec 2025.3.0
    Uninstalling fsspec-2025.3.0:
      Successfully uninstalled fsspec-2025.3.0
  Attempting uninstall: datasets
    Found existing installation: datasets 4.0.0
    Uninstalling datasets-4.0.0:
      Successfully uninstalled datasets-4.0.0
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behav
iour is the source of the following dependency conflicts.
torch 2.9.0+cu126 requires triton==3.5.0; platform_system == "Linux", but you have triton 3.6.0 which is incompatibl
e.
gcsfs 2025.3.0 requires fsspec==2025.3.0, but you have fsspec 2024.6.1 which is incompatible.
Successfully installed datasets-2.21.0 fsspec-2024.6.1
```

```
Collecting peft==0.12.0
  Downloading peft-0.12.0-py3-none-any.whl.metadata (13 kB)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.12/dist-packages (from peft==0.12.0) (2.0.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.12/dist-packages (from peft==0.12.0) (25.0)
Requirement already satisfied: psutil in /usr/local/lib/python3.12/dist-packages (from peft==0.12.0) (5.9.5)
Requirement already satisfied: pyyaml in /usr/local/lib/python3.12/dist-packages (from peft==0.12.0) (6.0.3)
Requirement already satisfied: torch>=1.13.0 in /usr/local/lib/python3.12/dist-packages (from peft==0.12.0) (2.9.0+cu
126)
Requirement already satisfied: transformers in /usr/local/lib/python3.12/dist-packages (from peft==0.12.0) (4.57.6)
Requirement already satisfied: tqdm in /usr/local/lib/python3.12/dist-packages (from peft==0.12.0) (4.67.1)
Requirement already satisfied: accelerate>=0.21.0 in /usr/local/lib/python3.12/dist-packages (from peft==0.12.0) (1.1
2.0)
Requirement already satisfied: safetensors in /usr/local/lib/python3.12/dist-packages (from peft==0.12.0) (0.7.0)
Requirement already satisfied: huggingface-hub>=0.17.0 in /usr/local/lib/python3.12/dist-packages (from peft==0.12.0)
(0.36.0)
Requirement already satisfied: filelock in /usr/local/lib/python3.12/dist-packages (from huggingface-hub>=0.17.0->pef
t==0.12.0) (3.20.3)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub>=0.1
7.0->peft==0.12.0) (2024.6.1)
Requirement already satisfied: requests in /usr/local/lib/python3.12/dist-packages (from huggingface-hub>=0.17.0->pef
t==0.12.0) (2.32.4)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.12/dist-packages (from huggingfac
e-hub>=0.17.0->peft==0.12.0) (4.15.0)
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub>
=0.17.0->peft==0.12.0) (1.2.0)
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-packages (from torch>=1.13.0->peft==0.12.
0) (75.2.0)
Requirement already satisfied: sympy>=1.13.3 in /usr/local/lib/python3.12/dist-packages (from torch>=1.13.0->peft==0.
12.0) (1.14.0)
Requirement already satisfied: networkx>=2.5.1 in /usr/local/lib/python3.12/dist-packages (from torch>=1.13.0->peft==
0.12.0) (3.6.1)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-packages (from torch>=1.13.0->peft==0.12.0)
(3.1.6)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch
>=1.13.0->peft==0.12.0) (12.6.77)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from tor
ch>=1.13.0->peft==0.12.0) (12.6.77)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.6.80 in /usr/local/lib/python3.12/dist-packages (from torch
>=1.13.0->peft==0.12.0) (12.6.80)
Requirement already satisfied: nvidia-cudnn-cu12==9.10.2.21 in /usr/local/lib/python3.12/dist-packages (from torch>=
1.13.0->peft==0.12.0) (9.10.2.21)
Requirement already satisfied: nvidia-cublas-cu12==12.6.4.1 in /usr/local/lib/python3.12/dist-packages (from torch>=
1.13.0->peft==0.12.0) (12.6.4.1)
Requirement already satisfied: nvidia-cufft-cu12==11.3.0.4 in /usr/local/lib/python3.12/dist-packages (from torch>=1.
13.0->peft==0.12.0) (11.3.0.4)
Requirement already satisfied: nvidia-curand-cu12==10.3.7.77 in /usr/local/lib/python3.12/dist-packages (from torch>=
1.13.0->peft==0.12.0) (10.3.7.77)
Requirement already satisfied: nvidia-cusolver-cu12==11.7.1.2 in /usr/local/lib/python3.12/dist-packages (from torch>
=1.13.0->peft==0.12.0) (11.7.1.2)
Requirement already satisfied: nvidia-cusparse-cu12==12.5.4.2 in /usr/local/lib/python3.12/dist-packages (from torch>
=1.13.0->peft==0.12.0) (12.5.4.2)
Requirement already satisfied: nvidia-cusparselt-cu12==0.7.1 in /usr/local/lib/python3.12/dist-packages (from torch>=
1.13.0->peft==0.12.0) (0.7.1)
Requirement already satisfied: nvidia-nccl-cu12==2.27.5 in /usr/local/lib/python3.12/dist-packages (from torch>=1.13.
0->peft==0.12.0) (2.27.5)
Requirement already satisfied: nvidia-nvshmem-cu12==3.3.20 in /usr/local/lib/python3.12/dist-packages (from torch>=1.
13.0->peft==0.12.0) (3.3.20)
Requirement already satisfied: nvidia-nvtx-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch>=1.1
3.0->peft==0.12.0) (12.6.77)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.6.85 in /usr/local/lib/python3.12/dist-packages (from torch>
=1.13.0->peft==0.12.0) (12.6.85)
Requirement already satisfied: nvidia-cufile-cu12==1.11.1.6 in /usr/local/lib/python3.12/dist-packages (from torch>=
1.13.0->peft==0.12.0) (1.11.1.6)
Collecting triton==3.5.0 (from torch>=1.13.0->peft==0.12.0)
  Downloading triton-3.5.0-cp312-cp312-manylinux_2_27_x86_64.manylinux_2_28_x86_64.whl.metadata (1.7 kB)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.12/dist-packages (from transformers->peft=
=0.12.0) (2025.11.3)
Requirement already satisfied: tokenizers<=0.23.0,>=0.22.0 in /usr/local/lib/python3.12/dist-packages (from transform
ers->peft==0.12.0) (0.22.2)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.12/dist-packages (from sympy>=1.13.3->tor
ch>=1.13.0->peft==0.12.0) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dist-packages (from jinja2->torch>=1.13.0
->peft==0.12.0) (3.0.3)
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests->hu
ggingface-hub>=0.17.0->peft==0.12.0) (3.4.4)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests->huggingface-hu
b>=0.17.0->peft==0.12.0) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests->huggingf
ace-hub>=0.17.0->peft==0.12.0) (2.5.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests->huggingf
ace-hub>=0.17.0->peft==0.12.0) (2026.1.4)
Downloading peft-0.12.0-py3-none-any.whl (296 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 296.4/296.4 kB 6.4 MB/s eta 0:00:00a 0:00:01
Downloading triton-3.5.0-cp312-cp312-manylinux_2_27_x86_64.manylinux_2_28_x86_64.whl (170.5 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 170.5/170.5 MB 12.6 MB/s eta 0:00:0000:0100:01
Installing collected packages: triton, peft
  Attempting uninstall: triton
    Found existing installation: triton 3.6.0
    Uninstalling triton-3.6.0:
      Successfully uninstalled triton-3.6.0
```

```
  Attempting uninstall: peft
    Found existing installation: peft 0.18.1
    Uninstalling peft-0.18.1:
      Successfully uninstalled peft-0.18.1
Successfully installed peft-0.12.0 triton-3.5.0
```

```
Collecting trl==0.9.6
  Downloading trl-0.9.6-py3-none-any.whl.metadata (12 kB)
Requirement already satisfied: torch>=1.4.0 in /usr/local/lib/python3.12/dist-packages (from trl==0.9.6) (2.9.0+cu12
6)
Requirement already satisfied: transformers>=4.31.0 in /usr/local/lib/python3.12/dist-packages (from trl==0.9.6) (4.5
7.6)
Collecting numpy<2.0.0,>=1.18.2 (from trl==0.9.6)
  Downloading numpy-1.26.4-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (61 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 61.0/61.0 kB 45.7 kB/s eta 0:00:00 0:00:01
Requirement already satisfied: accelerate in /usr/local/lib/python3.12/dist-packages (from trl==0.9.6) (1.12.0)
Requirement already satisfied: datasets in /usr/local/lib/python3.12/dist-packages (from trl==0.9.6) (2.21.0)
Collecting tyro>=0.5.11 (from trl==0.9.6)
  Downloading tyro-1.0.5-py3-none-any.whl.metadata (12 kB)
Requirement already satisfied: filelock in /usr/local/lib/python3.12/dist-packages (from torch>=1.4.0->trl==0.9.6)
(3.20.3)
Requirement already satisfied: typing-extensions>=4.10.0 in /usr/local/lib/python3.12/dist-packages (from torch>=1.4.
0->trl==0.9.6) (4.15.0)
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-packages (from torch>=1.4.0->trl==0.9.6)
(75.2.0)
Requirement already satisfied: sympy>=1.13.3 in /usr/local/lib/python3.12/dist-packages (from torch>=1.4.0->trl==0.9.
6) (1.14.0)
Requirement already satisfied: networkx>=2.5.1 in /usr/local/lib/python3.12/dist-packages (from torch>=1.4.0->trl==0.
9.6) (3.6.1)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-packages (from torch>=1.4.0->trl==0.9.6) (3.
1.6)
Requirement already satisfied: fsspec>=0.8.5 in /usr/local/lib/python3.12/dist-packages (from torch>=1.4.0->trl==0.9.
6) (2024.6.1)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch
>=1.4.0->trl==0.9.6) (12.6.77)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from tor
ch>=1.4.0->trl==0.9.6) (12.6.77)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.6.80 in /usr/local/lib/python3.12/dist-packages (from torch
>=1.4.0->trl==0.9.6) (12.6.80)
Requirement already satisfied: nvidia-cudnn-cu12==9.10.2.21 in /usr/local/lib/python3.12/dist-packages (from torch>=
1.4.0->trl==0.9.6) (9.10.2.21)
Requirement already satisfied: nvidia-cublas-cu12==12.6.4.1 in /usr/local/lib/python3.12/dist-packages (from torch>=
1.4.0->trl==0.9.6) (12.6.4.1)
Requirement already satisfied: nvidia-cufft-cu12==11.3.0.4 in /usr/local/lib/python3.12/dist-packages (from torch>=1.
4.0->trl==0.9.6) (11.3.0.4)
Requirement already satisfied: nvidia-curand-cu12==10.3.7.77 in /usr/local/lib/python3.12/dist-packages (from torch>=
1.4.0->trl==0.9.6) (10.3.7.77)
Requirement already satisfied: nvidia-cusolver-cu12==11.7.1.2 in /usr/local/lib/python3.12/dist-packages (from torch>
=1.4.0->trl==0.9.6) (11.7.1.2)
Requirement already satisfied: nvidia-cusparse-cu12==12.5.4.2 in /usr/local/lib/python3.12/dist-packages (from torch>
=1.4.0->trl==0.9.6) (12.5.4.2)
Requirement already satisfied: nvidia-cusparselt-cu12==0.7.1 in /usr/local/lib/python3.12/dist-packages (from torch>=
1.4.0->trl==0.9.6) (0.7.1)
Requirement already satisfied: nvidia-nccl-cu12==2.27.5 in /usr/local/lib/python3.12/dist-packages (from torch>=1.4.0
->trl==0.9.6) (2.27.5)
Requirement already satisfied: nvidia-nvshmem-cu12==3.3.20 in /usr/local/lib/python3.12/dist-packages (from torch>=1.
4.0->trl==0.9.6) (3.3.20)
Requirement already satisfied: nvidia-nvtx-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch>=1.4.
0->trl==0.9.6) (12.6.77)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.6.85 in /usr/local/lib/python3.12/dist-packages (from torch>
=1.4.0->trl==0.9.6) (12.6.85)
Requirement already satisfied: nvidia-cufile-cu12==1.11.1.6 in /usr/local/lib/python3.12/dist-packages (from torch>=
1.4.0->trl==0.9.6) (1.11.1.6)
Requirement already satisfied: triton==3.5.0 in /usr/local/lib/python3.12/dist-packages (from torch>=1.4.0->trl==0.9.
6) (3.5.0)
Requirement already satisfied: huggingface-hub<1.0,>=0.34.0 in /usr/local/lib/python3.12/dist-packages (from transfor
mers>=4.31.0->trl==0.9.6) (0.36.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.12/dist-packages (from transformers>=4.31.0-
>trl==0.9.6) (25.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.12/dist-packages (from transformers>=4.31.0->trl
==0.9.6) (6.0.3)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.12/dist-packages (from transformers>=4.31.
0->trl==0.9.6) (2025.11.3)
Requirement already satisfied: requests in /usr/local/lib/python3.12/dist-packages (from transformers>=4.31.0->trl==
0.9.6) (2.32.4)
Requirement already satisfied: tokenizers<=0.23.0,>=0.22.0 in /usr/local/lib/python3.12/dist-packages (from transform
ers>=4.31.0->trl==0.9.6) (0.22.2)
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.12/dist-packages (from transformers>=4.3
1.0->trl==0.9.6) (0.7.0)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.12/dist-packages (from transformers>=4.31.0->trl=
=0.9.6) (4.67.1)
Requirement already satisfied: docstring-parser>=0.15 in /usr/local/lib/python3.12/dist-packages (from tyro>=0.5.11->
trl==0.9.6) (0.17.0)
Requirement already satisfied: typeguard>=4.0.0 in /usr/local/lib/python3.12/dist-packages (from tyro>=0.5.11->trl==
0.9.6) (4.4.4)
Requirement already satisfied: psutil in /usr/local/lib/python3.12/dist-packages (from accelerate->trl==0.9.6) (5.9.
5)
Requirement already satisfied: pyarrow>=15.0.0 in /usr/local/lib/python3.12/dist-packages (from datasets->trl==0.9.6)
(18.1.0)
Requirement already satisfied: dill<0.3.9,>=0.3.0 in /usr/local/lib/python3.12/dist-packages (from datasets->trl==0.
9.6) (0.3.8)
Requirement already satisfied: pandas in /usr/local/lib/python3.12/dist-packages (from datasets->trl==0.9.6) (2.2.2)
Requirement already satisfied: xxhash in /usr/local/lib/python3.12/dist-packages (from datasets->trl==0.9.6) (3.6.0)
Requirement already satisfied: multiprocess in /usr/local/lib/python3.12/dist-packages (from datasets->trl==0.9.6)
(0.70.16)
Requirement already satisfied: aiohttp in /usr/local/lib/python3.12/dist-packages (from datasets->trl==0.9.6) (3.13.
```

```
3)
Requirement already satisfied: aiohappyeyeballs>=2.5.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp->data
sets->trl==0.9.6) (2.6.1)
Requirement already satisfied: aiosignal>=1.4.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets->t
rl==0.9.6) (1.4.0)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets->trl=
=0.9.6) (25.4.0)
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets->
trl==0.9.6) (1.8.0)
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets
->trl==0.9.6) (6.7.1)
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets->t
rl==0.9.6) (0.4.1)
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp->datasets->
trl==0.9.6) (1.22.0)
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub<
1.0,>=0.34.0->transformers>=4.31.0->trl==0.9.6) (1.2.0)
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests->tr
ansformers>=4.31.0->trl==0.9.6) (3.4.4)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests->transformers>=
4.31.0->trl==0.9.6) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests->transfor
mers>=4.31.0->trl==0.9.6) (2.5.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests->transfor
mers>=4.31.0->trl==0.9.6) (2026.1.4)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.12/dist-packages (from sympy>=1.13.3->tor
ch>=1.4.0->trl==0.9.6) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dist-packages (from jinja2->torch>=1.4.0-
>trl==0.9.6) (3.0.3)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.12/dist-packages (from pandas->dataset
ts->trl==0.9.6) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.12/dist-packages (from pandas->datasets->trl==
0.9.6) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.12/dist-packages (from pandas->datasets->trl=
=0.9.6) (2025.3)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.12/dist-packages (from python-dateutil>=2.8.2->pand
as->datasets->trl==0.9.6) (1.17.0)
Downloading trl-0.9.6-py3-none-any.whl (245 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 245.8/245.8 kB 10.7 MB/s eta 0:00:00
Downloading numpy-1.26.4-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (18.0 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 18.0/18.0 MB 125.1 MB/s eta 0:00:0000:0100:01
Downloading tyro-1.0.5-py3-none-any.whl (181 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 181.2/181.2 kB 23.2 MB/s eta 0:00:00
Installing collected packages: numpy, tyro, trl
  Attempting uninstall: numpy
    Found existing installation: numpy 2.0.2
    Uninstalling numpy-2.0.2:
      Successfully uninstalled numpy-2.0.2
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behav
iour is the source of the following dependency conflicts.
jaxlib 0.7.2 requires numpy>=2.0, but you have numpy 1.26.4 which is incompatible.
jax 0.7.2 requires numpy>=2.0, but you have numpy 1.26.4 which is incompatible.
tobler 0.13.0 requires numpy>=2.0, but you have numpy 1.26.4 which is incompatible.
rasterio 1.5.0 requires numpy>=2, but you have numpy 1.26.4 which is incompatible.
pytensor 2.37.0 requires numpy>=2.0, but you have numpy 1.26.4 which is incompatible.
opencv-contrib-python 4.13.0.90 requires numpy>=2; python_version >= "3.9", but you have numpy 1.26.4 which is incomp
atible.
opencv-python 4.13.0.90 requires numpy>=2; python_version >= "3.9", but you have numpy 1.26.4 which is incompatible.
shap 0.50.0 requires numpy>=2, but you have numpy 1.26.4 which is incompatible.
opencv-python-headless 4.13.0.90 requires numpy>=2; python_version >= "3.9", but you have numpy 1.26.4 which is incom
patible.
Successfully installed numpy-1.26.4 trl-0.9.6 tyro-1.0.5
```

```
Collecting scipy==1.13.1
  Downloading scipy-1.13.1-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (60 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 60.6/60.6 kB 2.8 MB/s eta 0:00:00
Requirement already satisfied: numpy<2.3,>=1.22.4 in /usr/local/lib/python3.12/dist-packages (from scipy==1.13.1) (1.
26.4)
Downloading scipy-1.13.1-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (38.2 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 38.2/38.2 MB 37.3 MB/s eta 0:00:0000:0100:01
Installing collected packages: scipy
  Attempting uninstall: scipy
    Found existing installation: scipy 1.16.3
    Uninstalling scipy-1.16.3:
      Successfully uninstalled scipy-1.16.3
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behav
iour is the source of the following dependency conflicts.
access 1.1.10.post3 requires scipy>=1.14.1, but you have scipy 1.13.1 which is incompatible.
tsfresh 0.21.1 requires scipy>=1.14.0; python_version >= "3.10", but you have scipy 1.13.1 which is incompatible.
jaxlib 0.7.2 requires numpy>=2.0, but you have numpy 1.26.4 which is incompatible.
jax 0.7.2 requires numpy>=2.0, but you have numpy 1.26.4 which is incompatible.
tobler 0.13.0 requires numpy>=2.0, but you have numpy 1.26.4 which is incompatible.
pytensor 2.37.0 requires numpy>=2.0, but you have numpy 1.26.4 which is incompatible.
shap 0.50.0 requires numpy>=2, but you have numpy 1.26.4 which is incompatible.
Successfully installed scipy-1.13.1
```

```
Requirement already satisfied: numpy in /usr/local/lib/python3.12/dist-packages (1.26.4)
Requirement already satisfied: wandb in /usr/local/lib/python3.12/dist-packages (0.24.0)
Requirement already satisfied: click>=8.0.1 in /usr/local/lib/python3.12/dist-packages (from wandb) (8.3.1)
Requirement already satisfied: gitpython!=3.1.29,>=1.0.0 in /usr/local/lib/python3.12/dist-packages (from wandb) (3.
1.46)
Requirement already satisfied: packaging in /usr/local/lib/python3.12/dist-packages (from wandb) (25.0)
Requirement already satisfied: platformdirs in /usr/local/lib/python3.12/dist-packages (from wandb) (4.5.1)
Requirement already satisfied: protobuf!=4.21.0,!=5.28.0,<7,>=3.19.0 in /usr/local/lib/python3.12/dist-packages (from
wandb) (5.29.5)
Requirement already satisfied: pydantic<3 in /usr/local/lib/python3.12/dist-packages (from wandb) (2.12.3)
Requirement already satisfied: pyyaml in /usr/local/lib/python3.12/dist-packages (from wandb) (6.0.3)
Requirement already satisfied: requests<3,>=2.0.0 in /usr/local/lib/python3.12/dist-packages (from wandb) (2.32.4)
Requirement already satisfied: sentry-sdk>=2.0.0 in /usr/local/lib/python3.12/dist-packages (from wandb) (2.51.0)
Requirement already satisfied: typing-extensions<5,>=4.8 in /usr/local/lib/python3.12/dist-packages (from wandb) (4.1
5.0)
Requirement already satisfied: gitdb<5,>=4.0.1 in /usr/local/lib/python3.12/dist-packages (from gitpython!=3.1.29,>=
1.0.0->wandb) (4.0.12)
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.12/dist-packages (from pydantic<3->wa
ndb) (0.7.0)
Requirement already satisfied: pydantic-core==2.41.4 in /usr/local/lib/python3.12/dist-packages (from pydantic<3->wan
db) (2.41.4)
Requirement already satisfied: typing-inspection>=0.4.2 in /usr/local/lib/python3.12/dist-packages (from pydantic<3->
wandb) (0.4.2)
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests<3,>
=2.0.0->wandb) (3.4.4)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests<3,>=2.0.0->wand
b) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests<3,>=2.0.0
->wandb) (2.5.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests<3,>=2.0.0
->wandb) (2026.1.4)
Requirement already satisfied: smmap<6,>=3.0.1 in /usr/local/lib/python3.12/dist-packages (from gitdb<5,>=4.0.1->gitp
ython!=3.1.29,>=1.0.0->wandb) (5.0.2)
Requirement already satisfied: accelerate>=1.7.0 in /usr/local/lib/python3.12/dist-packages (1.12.0)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.12/dist-packages (from accelerate>=1.7.0) (1.26.
4)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.12/dist-packages (from accelerate>=1.7.0) (2
5.0)
Requirement already satisfied: psutil in /usr/local/lib/python3.12/dist-packages (from accelerate>=1.7.0) (5.9.5)
Requirement already satisfied: pyyaml in /usr/local/lib/python3.12/dist-packages (from accelerate>=1.7.0) (6.0.3)
Requirement already satisfied: torch>=2.0.0 in /usr/local/lib/python3.12/dist-packages (from accelerate>=1.7.0) (2.9.
0+cu126)
Requirement already satisfied: huggingface_hub>=0.21.0 in /usr/local/lib/python3.12/dist-packages (from accelerate>=
1.7.0) (0.36.0)
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.12/dist-packages (from accelerate>=1.7.0)
(0.7.0)
Requirement already satisfied: filelock in /usr/local/lib/python3.12/dist-packages (from huggingface_hub>=0.21.0->acc
elerate>=1.7.0) (3.20.3)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.12/dist-packages (from huggingface_hub>=0.2
1.0->accelerate>=1.7.0) (2024.6.1)
Requirement already satisfied: requests in /usr/local/lib/python3.12/dist-packages (from huggingface_hub>=0.21.0->acc
elerate>=1.7.0) (2.32.4)
Requirement already satisfied: tqdm>=4.42.1 in /usr/local/lib/python3.12/dist-packages (from huggingface_hub>=0.21.0-
>accelerate>=1.7.0) (4.67.1)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.12/dist-packages (from huggingfac
e_hub>=0.21.0->accelerate>=1.7.0) (4.15.0)
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.12/dist-packages (from huggingface_hub>
=0.21.0->accelerate>=1.7.0) (1.2.0)
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-packages (from torch>=2.0.0->accelerate>=
1.7.0) (75.2.0)
Requirement already satisfied: sympy>=1.13.3 in /usr/local/lib/python3.12/dist-packages (from torch>=2.0.0->accelerat
e>=1.7.0) (1.14.0)
Requirement already satisfied: networkx>=2.5.1 in /usr/local/lib/python3.12/dist-packages (from torch>=2.0.0->acceler
ate>=1.7.0) (3.6.1)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-packages (from torch>=2.0.0->accelerate>=1.7.
0) (3.1.6)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch
>=2.0.0->accelerate>=1.7.0) (12.6.77)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from tor
ch>=2.0.0->accelerate>=1.7.0) (12.6.77)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.6.80 in /usr/local/lib/python3.12/dist-packages (from torch
>=2.0.0->accelerate>=1.7.0) (12.6.80)
Requirement already satisfied: nvidia-cudnn-cu12==9.10.2.21 in /usr/local/lib/python3.12/dist-packages (from torch>=
2.0.0->accelerate>=1.7.0) (9.10.2.21)
Requirement already satisfied: nvidia-cublas-cu12==12.6.4.1 in /usr/local/lib/python3.12/dist-packages (from torch>=
2.0.0->accelerate>=1.7.0) (12.6.4.1)
Requirement already satisfied: nvidia-cufft-cu12==11.3.0.4 in /usr/local/lib/python3.12/dist-packages (from torch>=2.
0.0->accelerate>=1.7.0) (11.3.0.4)
Requirement already satisfied: nvidia-curand-cu12==10.3.7.77 in /usr/local/lib/python3.12/dist-packages (from torch>=
2.0.0->accelerate>=1.7.0) (10.3.7.77)
Requirement already satisfied: nvidia-cusolver-cu12==11.7.1.2 in /usr/local/lib/python3.12/dist-packages (from torch>
=2.0.0->accelerate>=1.7.0) (11.7.1.2)
Requirement already satisfied: nvidia-cusparse-cu12==12.5.4.2 in /usr/local/lib/python3.12/dist-packages (from torch>
=2.0.0->accelerate>=1.7.0) (12.5.4.2)
Requirement already satisfied: nvidia-cusparselt-cu12==0.7.1 in /usr/local/lib/python3.12/dist-packages (from torch>=
2.0.0->accelerate>=1.7.0) (0.7.1)
Requirement already satisfied: nvidia-nccl-cu12==2.27.5 in /usr/local/lib/python3.12/dist-packages (from torch>=2.0.0
->accelerate>=1.7.0) (2.27.5)
Requirement already satisfied: nvidia-nvshmem-cu12==3.3.20 in /usr/local/lib/python3.12/dist-packages (from torch>=2.
0.0->accelerate>=1.7.0) (3.3.20)
```

```
Requirement already satisfied: nvidia-nvtx-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch>=2.0.
0->accelerate>=1.7.0) (12.6.77)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.6.85 in /usr/local/lib/python3.12/dist-packages (from torch>
=2.0.0->accelerate>=1.7.0) (12.6.85)
Requirement already satisfied: nvidia-cufile-cu12==1.11.1.6 in /usr/local/lib/python3.12/dist-packages (from torch>=
2.0.0->accelerate>=1.7.0) (1.11.1.6)
Requirement already satisfied: triton==3.5.0 in /usr/local/lib/python3.12/dist-packages (from torch>=2.0.0->accelerat
e>=1.7.0) (3.5.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.12/dist-packages (from sympy>=1.13.3->tor
ch>=2.0.0->accelerate>=1.7.0) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dist-packages (from jinja2->torch>=2.0.0-
>accelerate>=1.7.0) (3.0.3)
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests->hu
ggingface_hub>=0.21.0->accelerate>=1.7.0) (3.4.4)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests->huggingface_hu
b>=0.21.0->accelerate>=1.7.0) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests->huggingf
ace_hub>=0.21.0->accelerate>=1.7.0) (2.5.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests->huggingf
ace_hub>=0.21.0->accelerate>=1.7.0) (2026.1.4)
Requirement already satisfied: triton in /usr/local/lib/python3.12/dist-packages (3.5.0)
Collecting triton
  Using cached triton-3.6.0-cp312-cp312-manylinux_2_27_x86_64.manylinux_2_28_x86_64.whl.metadata (1.7 kB)
Using cached triton-3.6.0-cp312-cp312-manylinux_2_27_x86_64.manylinux_2_28_x86_64.whl (188.3 MB)
Installing collected packages: triton
  Attempting uninstall: triton
    Found existing installation: triton 3.5.0
    Uninstalling triton-3.5.0:
      Successfully uninstalled triton-3.5.0
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behav
iour is the source of the following dependency conflicts.
torch 2.9.0+cu126 requires triton==3.5.0; platform_system == "Linux", but you have triton 3.6.0 which is incompatibl
e.
Successfully installed triton-3.6.0
```

의존성 설치

학습에 필요한 라이브러리들을 설치합니다.

In [26]:
```python
import re
import random
import torch
import torch.nn.functional as F
from torch.utils.data import Dataset, DataLoader
from datasets import load_dataset
from transformers import (
    AutoModelForCausalLM,
    AutoTokenizer,
    get_linear_schedule_with_warmup,
)
from peft import LoraConfig, get_peft_model, TaskType
from tqdm import tqdm

# PyTorch 속도 최적화 (Ampere+ GPU)
torch.backends.cuda.matmul.allow_tf32 = True
torch.backends.cudnn.allow_tf32 = True
if hasattr(torch, "set_float32_matmul_precision"):
    torch.set_float32_matmul_precision("high")
```

라이브러리 임포트 및 PyTorch 최적화

필요한 라이브러리를 임포트하고 Ampere+ GPU를 위한 TF32 최적화를 활성화합니다.

In [27]:
```python
# Flash Attention 2 설치 (선택)
# pip install flash-attn은 소스 빌드로 30분+ 소요, CUDA 11.8 환경에서 빌드 실패 자주 발생.
# → SDPA fallback 사용 시에도 학습 정상 동작 (PyTorch 내장 최적화).
#
# [CUDA 12 + PyTorch 2.4 사용 시] pre-built wheel로 설치 가능:
# !pip install https://github.com/Dao-AILab/flash-attention/releases/download/v2.8.3/flash_attn-2.8.3+cu12torch2.4cx
x11abiFALSE-cp312-cp312-linux_x86_64.whl


def get_attn_implementation():
    """Flash Attention 2 사용, 실패 시 SDPA로 fallback (PyTorch 내장, 충분히 빠름)"""
    try:
        from flash_attn import flash_attn_func
        return "flash_attention_2"
    except ImportError:
        return "sdpa"

ATTN_IMPL = get_attn_implementation()
USE_TORCH_COMPILE = False  # True: JIT 최적화 (첫 epoch 느림, 저장 시 주의)
print(f"Attention implementation: {ATTN_IMPL}")
```

```
Attention implementation: sdpa
```

In [28]:
```
# GPU 확인
print(f"CUDA available: {torch.cuda.is_available()}")
```

```
CUDA available: True
```

2. 데이터 준비

2.1 grade-school-math에서 1500개 샘플링

In [29]:
```
def load_and_sample_gsm(n_samples=None, seed=42):
    """grade-school-math-instructions 로드. n_samples=None이면 전체 데이터셋 반환"""
    dataset = load_dataset("qwedsacf/grade-school-math-instructions")
    train_data = dataset["train"]

    if n_samples is None:
        samples = [train_data[i] for i in range(len(train_data))]
    else:
        random.seed(seed)
        indices = random.sample(range(len(train_data)), min(n_samples, len(train_data)))
        samples = [train_data[i] for i in indices]

    return samples

raw_samples = load_and_sample_gsm(n_samples=1500)
print(f"Loaded {len(raw_samples)} problems (full dataset)")
print("\nExample:")
print(raw_samples[0])
```

```
Loaded 1500 problems (full dataset)

Example:
{'INSTRUCTION': 'Five food companies sponsored a local food bank. Foster Farms donated 45 dressed chickens; American
Summits donated twice the number of bottled water than the number of dressed chicken donated by Foster Farms; Hormel
donated three times the number of dressed chickens that Foster Farms donated; Boudin Butchers donated one-third of th
e number of dressed chickens that Hormel donated; Del Monte Foods donated 30 fewer bottles of water than American Sum
mits. How many food items did the companies donate in total?\nGive me a solution to this problem', 'RESPONSE': 'Ameri
can Summits donated 45 x 2 = 90 bottled waters.\nHormel donated 45 x 3 = 135 spams.\nBoudin Bakery donated 135 x 1/3
= 45 sourdoughs.\nDel Monte Foods donated 90 - 30 = 60 canned fruits.\nTherefore, a total of 45 + 90 + 135 + 45 + 60
= 375 different food.', 'SOURCE': 'grade-school-math'}
```

데이터 로드 함수

grade-school-math-instructions 데이터셋을 로드하고 n_samples개를 샘플링합니다. None이면 전체 데이터셋을 반환합니다.

2.2 Rule 기반: RESPONSE에서 최종 답 추출

grade-school-math RESPONSE는 단계별 풀이이며, 마지막 숫자가 보통 최종 답입니다.

In [30]:
```
def extract_final_answer(response: str):
    """
    RESPONSE에서 최종 수치 답을 rule 기반으로 추출.
    - 마지막 줄/문장의 '= 숫자' 패턴 우선
    - 없으면 마지막 등장 숫자 사용
    """
    if not response or not response.strip():
        return None

    # '= 숫자' 패턴 (정수 또는 소수)
    eq_matches = list(re.finditer(r"=\s*(-?\d+(?:\.\d+)?)\b", response))
    if eq_matches:
        last_eq = eq_matches[-1].group(1)
        try:
            val = float(last_eq)
            if val == int(val):
                return str(int(val))
            return str(val)
        except ValueError:
            pass

    # 일반 숫자 (마지막 것)
    num_matches = list(re.finditer(r"(-?\d+(?:\.\d+)?)\b", response))
    if num_matches:
        last_num = num_matches[-1].group(1)
        try:
            val = float(last_num)
            if val == int(val):
                return str(int(val))
            return str(val)
        except ValueError:
            pass

    return None
```

최종 답 추출 함수

RESPONSE 텍스트에서 "= 숫자" 패턴 또는 마지막 숫자를 추출하여 최종 답을 반환합니다.

In [31]:
```
# 추출 테스트
test_responses = [
    "Natalia sold 48/2 = 24 clips in May.\nNatalia sold 48+24 = 72 clips altogether in April and May.",
    "Weng earns 12/60 = $0.2 per minute.\nWorking 50 minutes, she earned 0.2 x 50 = $10.",
    "He eats 32 from the largest pizzas because 2 x 16 = 32\nHe eats 16 from the small pizza because 2 x 8 = 16\nHe
eats 48 pieces because 32 + 16 = 48",
]
for r in test_responses:
    print(f"Response: {r[:80]}...")
    print(f"Extracted: {extract_final_answer(r)}")
    print()
```

```
Response: Natalia sold 48/2 = 24 clips in May.
Natalia sold 48+24 = 72 clips altogether in...
Extracted: 72

Response: Weng earns 12/60 = $0.2 per minute.
Working 50 minutes, she earned 0.2 x 50 = $1...
Extracted: 10

Response: He eats 32 from the largest pizzas because 2 x 16 = 32
He eats 16 from the small...
Extracted: 48
```

2.3 MathQA-style: 5개 옵션 생성 (정답 1 + diverse distractors 4)

In [32]:
```
def generate_options(correct_answer: str, n_options=5, seed=None):
    """
    MathQA-style diverse distractors: wider deltas, multiplicative, order-of-magnitude.
    Produces options like [24, 120, 625, 720, 1024] instead of [71, 73, 74, 70, 72].
    """
    if seed is not None:
        random.seed(seed)

    try:
        val = float(correct_answer)
        is_int = val == int(val)
        ival = int(val) if is_int else val
    except (ValueError, TypeError):
        return None

    def fmt(x):
        if isinstance(x, float) and x == int(x):
            return str(int(x))
        return str(x)

    # 오답 후보 (rule 기반)
    candidates = []
    if is_int:
        for delta in [1, 2, -1, -2, 5, -5, 10]:
            candidates.append(ival + delta)
        candidates.extend([ival * 2, ival // 2 if ival != 0 else 1, ival + 3, ival - 3])
    else:
        for delta in [1.0, 2.0, -1.0, 0.5, -0.5]:
            candidates.append(val + delta)
        candidates.extend([val * 2, val / 2])

    wrong = []
    for c in candidates:
        try:
            fc = float(c)
            if fc != val and fc > 0 and fc < 1e6:
                wrong.append(fmt(fc))
        except (ValueError, TypeError):
            pass

    wrong = list(dict.fromkeys(wrong))

    if len(wrong) < n_options - 1:
        extra = [ival + 7, ival - 7, ival * 4, ival + 15, ival - 15] if is_int else [val + 3, val - 2]
        for e in extra:
            try:
                fe = float(e)
                if fe != val and fe > 0 and fe < 1e6 and fmt(e) not in wrong:
                    wrong.append(fmt(e))
            except (ValueError, TypeError):
                pass
            if len(wrong) >= n_options - 1:
                break

    wrong = wrong[: n_options - 1]
    options = [correct_answer] + wrong
    random.shuffle(options)
    correct_idx = options.index(correct_answer)
    return options, correct_idx
```

MathQA 스타일 옵션 생성 함수

정답에 대해 MathQA와 유사한 다양한 오답 옵션(distractors)을 생성합니다. 덧셈/뺄셈, 곱셈/나눗셈 기반 오답을 포함합니다.

In [33]:
```
def to_mathqa_question(instruction: str) -> str:
    """
    Strip instruction suffix to match MathQA Problem format (pure word problem).
    MathQA uses Problem text without 'Give me a solution' etc.
    """
    suffixes = [
        "\nGive me a solution to this problem",
        "\nCan you show me the way?",
        "\nSolve this step by step.",
    ]
    q = instruction.strip()
    for suf in suffixes:
        if q.endswith(suf):
            q = q[: -len(suf)].strip()
            break
    return q
```

In [34]:
```
# 옵션 생성 테스트
for ans in ["72", "10", "48", "5", "0.2"]:
    opts, idx = generate_options(ans, seed=42)
    print(f"Correct: {ans} -> options: {opts}, correct_idx: {idx}")
```

```
Correct: 72 -> options: ['71', '73', '74', '70', '72'], correct_idx: 4
Correct: 10 -> options: ['9', '11', '12', '8', '10'], correct_idx: 4
Correct: 48 -> options: ['47', '49', '50', '46', '48'], correct_idx: 4
Correct: 5 -> options: ['4', '6', '7', '3', '5'], correct_idx: 4
Correct: 0.2 -> options: ['0.7', '1.2', '2.2', '0.4', '0.2'], correct_idx: 4
```

2.4 전체 데이터셋 구성

In [35]:
```
def build_mc_dataset_aligned(raw_samples):
    """
    lm-eval mathqa 평가와 완전히 일치하는 데이터셋.
    - question: to_mathqa_question(INSTRUCTION) - MathQA Problem 형식 (suffix 제거)
    - options: MathQA-style diverse distractors (실제 숫자 값들)
    - correct_idx: 정답 인덱스
    """
    data = []
    for i, s in enumerate(raw_samples):
        instruction = s.get("INSTRUCTION", "")
        response = s.get("RESPONSE", "")

        correct = extract_final_answer(response)
        if correct is None:
            continue

        result = generate_options(correct, seed=i)
        if result is None:
            continue

        options, correct_idx = result
        if len(options) != 5:
            continue

        data.append({
            "question": to_mathqa_question(instruction),  # MathQA Problem format (no suffix)
            "options": options,  # 실제 숫자 값들
            "correct_idx": correct_idx,
        })

    return data

mc_data = build_mc_dataset_aligned(raw_samples)
print(f"Valid samples: {len(mc_data)} / {len(raw_samples)}")
print("\nExample:")
ex = mc_data[0]
print(f"Q: {ex['question'][:100]}...")
print(f"Options: {ex['options']}")
print(f"Correct index: {ex['correct_idx']}")
```

```
Valid samples: 1497 / 1500

Example:
Q: Five food companies sponsored a local food bank. Foster Farms donated 45 dressed chickens; American ...
Options: ['377', '376', '375', '373', '374']
Correct index: 2
```

MC 데이터셋 생성 함수

원본 샘플들을 lm-eval mathqa 평가와 동일한 Multiple Choice 포맷으로 변환합니다.

3. 프롬프트 형식 및 Dataset 클래스 (lm-eval mathqa와 동일)

In [36]:
```
def format_prompt_eval_aligned(question):
    """lm-eval mathqa와 동일한 포맷: Question: ... Answer:"""
    return f"Question: {question}\nAnswer:"
```

In [37]:
```
# lm-eval mathqa와 동일한 포맷 사용 (format_prompt_eval_aligned 참조)
# 옵션 없이 "Question: ... Answer:" 형태, continuation은 실제 숫자 값
```

In [38]:
```
class MCDatasetAligned(Dataset):
    """lm-eval mathqa와 완전 일치하는 Multiple choice Dataset"""

    def __init__(self, data, tokenizer):
        self.data = data
        self.tokenizer = tokenizer

    def __len__(self):
        return len(self.data)

    def __getitem__(self, idx):
        item = self.data[idx]
        # lm-eval과 동일: 옵션 없이 Question + Answer
        prefix = f"Question: {item['question']}\nAnswer:"
        # continuation: 실제 숫자 값 (공백 prefix로 토큰 분리)
        continuations = [f" {opt}" for opt in item["options"]]
        return {
            "prefix": prefix,
            "continuations": continuations,
            "correct_idx": item["correct_idx"],
        }


# 시퀀스 길이 (짧게 = 속도 향상)
MAX_PREFIX_LEN = 256
MAX_FULL_LEN = 320


def collate_single(batch):
    """DataLoader용: batch_size=1일 때 단일 샘플 반환"""
    return batch[0]


def collate_batch(batch):
    """배치 collate: per-sample continuations (숫자 옵션은 샘플마다 다름)"""
    return {
        "prefix": [b["prefix"] for b in batch],
        "continuations": [b["continuations"] for b in batch],
        "correct_idx": torch.tensor([b["correct_idx"] for b in batch], dtype=torch.long),
    }
```

Dataset 클래스 및 Collate 함수

PyTorch Dataset 클래스와 배치 처리를 위한 collate 함수를 정의합니다. lm-eval mathqa와 동일한 "Question: ... Answer:" 형식을 사용합니다.

4. Log-likelihood 계산 및 Cross-Entropy 학습

각 옵션 continuation(실제 숫자 값, lm-eval mathqa와 동일)에 대해 모델이 부여하는 **log-likelihood**를 구한 뒤, **softmax over options**로 확률 분포를 만들고, 정답 인덱스에 대한 **cross-entropy**로 학습합니다.

In [39]:
```python
def compute_option_log_likelihoods(model, tokenizer, prefix, continuations, device):
    """
    prefix가 주어졌을 때 각 continuation의 (평균) log-likelihood 계산.
    반환: (batch_size, n_options) 형태의 log-likelihood 텐서

    log P(continuation | prefix) = sum over tokens in continuation of log P(token | context)
    """
    prefix_ids = tokenizer(
        prefix,
        return_tensors="pt",
        truncation=True,
        max_length=MAX_PREFIX_LEN,
        add_special_tokens=True,
    ).input_ids.to(device)

    log_likelihoods = []
    for cont in continuations:
        # prefix + continuation 전체로 forward, continuation 토큰들에 대한 log prob 합
        full_text = prefix + cont
        full_ids = tokenizer(
            full_text,
            return_tensors="pt",
            truncation=True,
            max_length=MAX_FULL_LEN,
            add_special_tokens=True,
        ).input_ids.to(device)

        cont_ids = tokenizer(
            cont,
            return_tensors="pt",
            add_special_tokens=False,
        ).input_ids.to(device)

        n_prefix = prefix_ids.shape[1]
        n_cont = cont_ids.shape[1]

        outputs = model(full_ids)
        logits = outputs.logits  # (1, seq_len, vocab)

        # continuation 토큰들의 log prob: logits[:-1]으로 다음 토큰 예측
        # continuation은 prefix 다음부터이므로, positions [n_prefix-1 : n_prefix-1+n_cont]
        # 에서의 log prob 합
        cont_log_probs = []
        for j in range(n_cont):
            pos = n_prefix - 1 + j
            if pos < 0:
                continue
            next_token_id = full_ids[0, pos + 1].item()
            log_prob = F.log_softmax(logits[0, pos], dim=-1)[next_token_id]
            cont_log_probs.append(log_prob)

        if cont_log_probs:
            ll = sum(cont_log_probs)
        else:
            ll = logits[0, 0, 0] * 0.0 - 1e9  # grad 연결 유지
        log_likelihoods.append(ll)

    return torch.stack(log_likelihoods)
```

Log-likelihood 계산 함수 (단일 샘플)

주어진 prefix에 대해 각 continuation 옵션의 log-likelihood를 계산합니다.

In [40]:
```
def compute_option_log_likelihoods_batched(model, tokenizer, prefix, continuations, device):
    """
    단일 또는 배치 샘플에 대해 5개 옵션의 log-likelihood 계산.
    prefix: str 또는 list[str]
    continuations: list[str] (공유) 또는 list[list[str]] (per-sample, lm-eval aligned)
    반환: (n_options,) 또는 (batch, n_options)
    """
    per_sample = continuations and isinstance(continuations[0], (list, tuple))
    if per_sample:
        batch_size = len(prefix)
        log_likelihoods_per_option = []
        for k in range(5):
            full_texts = [prefix[i] + continuations[i][k] for i in range(batch_size)]
            full_enc = tokenizer(full_texts, return_tensors="pt", truncation=True, max_length=MAX_FULL_LEN, padding=
True, add_special_tokens=True)
            full_ids = full_enc.input_ids.to(device)
            attn_mask = full_enc.attention_mask.to(device)
            prefix_lengths = tokenizer(prefix, return_tensors="pt", truncation=True, max_length=MAX_PREFIX_LEN, padd
ing=True, add_special_tokens=True).attention_mask.sum(dim=1)
            outputs = model(full_ids, attention_mask=attn_mask)
            logits = outputs.logits
            batch_lls = []
            for b in range(batch_size):
                n_prefix = prefix_lengths[b].item()
                n_full = attn_mask[b].sum().item()
                n_cont = n_full - n_prefix
                if n_cont <= 0:
                    batch_lls.append(logits[b, 0, 0] * 0.0 - 1e9)
                    continue
                ll_sum = logits[b, 0, 0] * 0.0
                for j in range(n_cont):
                    pos = n_prefix - 1 + j
                    next_id = full_ids[b, pos + 1].item()
                    ll_sum = ll_sum + F.log_softmax(logits[b, pos], dim=-1)[next_id]
                batch_lls.append(ll_sum)
            log_likelihoods_per_option.append(torch.stack(batch_lls))
        out = torch.stack(log_likelihoods_per_option, dim=1)
        return out.squeeze(0) if batch_size == 1 else out
    single = isinstance(prefix, str)
    if single:
        prefix = [prefix]

    prefix_enc = tokenizer(
        prefix,
        return_tensors="pt",
        truncation=True,
        max_length=MAX_PREFIX_LEN,
        padding=True,
        add_special_tokens=True,
    )
    prefix_lengths = prefix_enc.attention_mask.sum(dim=1)

    log_likelihoods_per_option = []
    for cont in continuations:
        full_texts = [p + cont for p in prefix]
        full_enc = tokenizer(
            full_texts,
            return_tensors="pt",
            truncation=True,
            max_length=MAX_FULL_LEN,
            padding=True,
            add_special_tokens=True,
        )
        full_ids = full_enc.input_ids.to(device)
        attn_mask = full_enc.attention_mask.to(device)

        outputs = model(full_ids, attention_mask=attn_mask)
        logits = outputs.logits

        batch_lls = []
        for b in range(len(prefix)):
            n_prefix = prefix_lengths[b].item()
            n_full = attn_mask[b].sum().item()
            n_cont = n_full - n_prefix

            if n_cont <= 0:
                ll = logits[b, 0, 0] * 0.0 - 1e9
                batch_lls.append(ll)
                continue

            ll_sum = logits[b, 0, 0] * 0.0
            for j in range(n_cont):
                pos = n_prefix - 1 + j
                next_id = full_ids[b, pos + 1].item()
                ll_sum = ll_sum + F.log_softmax(logits[b, pos], dim=-1)[next_id]
            batch_lls.append(ll_sum)

        log_likelihoods_per_option.append(torch.stack(batch_lls))
```

```
    out = torch.stack(log_likelihoods_per_option, dim=1)
    return out.squeeze(0) if single else out
```

Log-likelihood 계산 함수 (배치)

배치 단위로 5개 옵션의 log-likelihood를 효율적으로 계산합니다.

In [41]:
```
def mc_cross_entropy_loss(log_likelihoods, correct_idx):
    """
    Softmax over options + Cross-entropy loss.
    log_likelihoods: (n_options,) 또는 (batch, n_options)
    correct_idx: int 또는 (batch,) tensor
    """
    if log_likelihoods.dim() == 1:
        log_likelihoods = log_likelihoods.unsqueeze(0)
    if not isinstance(correct_idx, torch.Tensor):
        correct_idx = torch.tensor([correct_idx], device=log_likelihoods.device, dtype=torch.long)
    elif correct_idx.dim() == 0:
        correct_idx = correct_idx.unsqueeze(0)
    log_probs = F.log_softmax(log_likelihoods, dim=-1)
    return F.nll_loss(log_probs, correct_idx)
```

Multiple Choice Cross-Entropy Loss 함수

옵션들의 log-likelihood에 softmax를 적용하고 정답 인덱스에 대한 NLL loss를 계산합니다.

5. 모델 로드 및 학습 루프

In [42]:
```
MODEL_ID = "Qwen/Qwen2.5-0.5B"
OUTPUT_DIR = "./outputs/03_sft_improved_mc"
```

In [43]:
```
tokenizer = AutoTokenizer.from_pretrained(MODEL_ID, trust_remote_code=True)
if tokenizer.pad_token is None:
    tokenizer.pad_token = tokenizer.eos_token
    tokenizer.pad_token_id = tokenizer.eos_token_id

model = AutoModelForCausalLM.from_pretrained(
    MODEL_ID,
    torch_dtype=torch.float16,
    device_map="auto",
    trust_remote_code=True,
    attn_implementation=ATTN_IMPL,
)

# LoRA 적용
lora_config = LoraConfig(
    r=16,
    lora_alpha=32,
    lora_dropout=0.05,
    bias="none",
    task_type=TaskType.CAUSAL_LM,
    target_modules=["q_proj", "k_proj", "v_proj", "o_proj", "gate_proj", "up_proj", "down_proj"],
)
model = get_peft_model(model, lora_config)
# model.gradient_checkpointing_enable()  # 메모리 절약, 대형 모델/긴 시퀀스에 유리
model.print_trainable_parameters()

# torch.compile: PyTorch 2.0+ JIT 최적화 (선택)
if USE_TORCH_COMPILE and hasattr(torch, "compile"):
    model = torch.compile(model, mode="reduce-overhead")
```

```
trainable params: 8,798,208 || all params: 502,830,976 || trainable%: 1.7497
```

모델 및 토크나이저 로드, LoRA 적용

Qwen2.5-0.5B 모델을 bfloat16으로 로드하고 LoRA 어댑터를 적용합니다.

In [44]:
```
# LoRA 사용 시 prepare_model_for_kbit_training은 full model용. float16 모델에는 get_peft_model만 사용
# 위 셀에서 prepare_model_for_kbit_training 제거 (float16 모델용)
lora_config = LoraConfig(
    r=16,
    lora_alpha=32,
    lora_dropout=0.05,
    bias="none",
    task_type=TaskType.CAUSAL_LM,
    target_modules=["q_proj", "k_proj", "v_proj", "o_proj", "gate_proj", "up_proj", "down_proj"],
)
model = get_peft_model(model, lora_config)
model.print_trainable_parameters()
```

```
trainable params: 8,798,208 || all params: 502,830,976 || trainable%: 1.7497
```

In [45]:
```
# 학습/검증 분할
split_idx = int(len(mc_data) * 0.9)
train_data = mc_data[:split_idx]
eval_data = mc_data[split_idx:]

train_dataset = MCDatasetAligned(train_data, tokenizer)
eval_dataset = MCDatasetAligned(eval_data, tokenizer)

# DataLoader: OOM 시 BATCH_SIZE 4→2로 감소, 메모리 여유 시 8로 증가
BATCH_SIZE = 4
NUM_WORKERS = 0
train_loader = DataLoader(
    train_dataset, batch_size=BATCH_SIZE, shuffle=True,
    num_workers=NUM_WORKERS, pin_memory=True, collate_fn=collate_batch,
)
eval_loader = DataLoader(
    eval_dataset, batch_size=BATCH_SIZE, shuffle=False,
    num_workers=NUM_WORKERS, pin_memory=True, collate_fn=collate_batch,
)

print(f"Train: {len(train_dataset)}, Eval: {len(eval_dataset)}")
```

```
Train: 1347, Eval: 150
```

데이터 분할 및 DataLoader 설정

데이터를 90% 학습, 10% 검증으로 분할하고 DataLoader를 생성합니다.

In [46]:
```
def train_epoch(model, tokenizer, dataloader, optimizer, scheduler, device, epoch):
    model.train()
    total_loss = 0.0
    n = 0

    pbar = tqdm(dataloader, desc=f"Epoch {epoch}")
    for item in pbar:
        prefix = item["prefix"]
        continuations = item["continuations"]
        correct_idx = item["correct_idx"].to(device)

        log_likelihoods = compute_option_log_likelihoods_batched(
            model, tokenizer, prefix, continuations, device
        )

        loss = mc_cross_entropy_loss(log_likelihoods, correct_idx)

        optimizer.zero_grad()
        loss.backward()
        torch.nn.utils.clip_grad_norm_(model.parameters(), 1.0)
        optimizer.step()
        if scheduler is not None:
            scheduler.step()

        total_loss += loss.item()
        n += len(prefix) if isinstance(prefix, list) else 1
        pbar.set_postfix({"loss": f"{loss.item():.4f}"})

    return total_loss / len(dataloader)
```

In [47]:
```
def evaluate(model, tokenizer, dataloader, device):
    model.eval()
    total_loss = 0.0
    correct = 0
    n = 0

    with torch.no_grad():
        for item in tqdm(dataloader, desc="Eval"):
            prefix = item["prefix"]
            continuations = item["continuations"]
            correct_idx = item["correct_idx"].to(device)

            log_likelihoods = compute_option_log_likelihoods_batched(
                model, tokenizer, prefix, continuations, device
            )

            loss = mc_cross_entropy_loss(log_likelihoods, correct_idx)
            total_loss += loss.item()

            pred = log_likelihoods.argmax(dim=1)
            correct += (pred == correct_idx).sum().item()
            n += len(prefix)

    return total_loss / len(dataloader), correct / n
```

5. 학습 루프

In [51]:
```
device = next(model.parameters()).device
optimizer = torch.optim.AdamW(model.parameters(), lr=2e-5, weight_decay=0.01)
num_epochs = 3
num_training_steps = num_epochs * len(train_loader)
scheduler = get_linear_schedule_with_warmup(optimizer, num_warmup_steps=int(0.03 * num_training_steps), num_training
_steps=num_training_steps)
```

In [52]:
```
import os
os.makedirs(OUTPUT_DIR, exist_ok=True)

for epoch in range(num_epochs):
    train_loss = train_epoch(model, tokenizer, train_loader, optimizer, scheduler, device, epoch + 1)
    eval_loss, eval_acc = evaluate(model, tokenizer, eval_loader, device)
    print(f"Epoch {epoch+1} | Train Loss: {train_loss:.4f} | Eval Loss: {eval_loss:.4f} | Eval Acc: {eval_acc:.4f}")

    model.save_pretrained(os.path.join(OUTPUT_DIR, f"checkpoint-epoch{epoch+1}"))
    tokenizer.save_pretrained(os.path.join(OUTPUT_DIR, f"checkpoint-epoch{epoch+1}"))
```

```
Epoch 1: 100%|██████████| 337/337 [06:01<00:00,  1.07s/it, loss=0.1924]
Eval: 100%|██████████| 38/38 [00:15<00:00,  2.48it/s]

Epoch 1 | Train Loss: 1.0176 | Eval Loss: 1.1584 | Eval Acc: 0.6000

Epoch 2: 100%|██████████| 337/337 [06:01<00:00,  1.07s/it, loss=1.7812]
Eval: 100%|██████████| 38/38 [00:14<00:00,  2.56it/s]

Epoch 2 | Train Loss: 0.8030 | Eval Loss: 1.1369 | Eval Acc: 0.6000

Epoch 3: 100%|██████████| 337/337 [06:02<00:00,  1.08s/it, loss=0.0107]
Eval: 100%|██████████| 38/38 [00:14<00:00,  2.53it/s]

Epoch 3 | Train Loss: 0.5566 | Eval Loss: 1.4658 | Eval Acc: 0.6133
```

In [53]:
```
# 최종 모델 저장
model.save_pretrained(OUTPUT_DIR)
tokenizer.save_pretrained(OUTPUT_DIR)
print(f"Model saved to {OUTPUT_DIR}")
```

```
Model saved to ./outputs/03_sft_improved_mc
```

6. Google Drive 업로드 (평가용)

학습된 모델을 Google Drive에 업로드하여 02_evaluation.ipynb에서 평가할 수 있도록 합니다.

In [54]:
```
# LoRA merge + Google Drive 업로드 (학습 완료 후 실행)
from peft import PeftModel
import shutil

# 1. LoRA adapter를 base 모델에 merge (평가 시 full model 필요)
model = model.merge_and_unload()
model.save_pretrained(OUTPUT_DIR)
tokenizer.save_pretrained(OUTPUT_DIR)
print(f"Merged model saved to {OUTPUT_DIR}")

# 2. Google Drive 마운트
from google.colab import drive
drive.mount('/content/drive')

# 3. Drive에 복사 (02_evaluation.ipynb에서 이 경로로 로드)
DRIVE_MODEL_DIR = "/content/drive/MyDrive/llm-math-models/qwen2.5-0.5b-math-sft-improved-mc"
os.makedirs(DRIVE_MODEL_DIR, exist_ok=True)
for f in os.listdir(OUTPUT_DIR):
    src = os.path.join(OUTPUT_DIR, f)
    dst = os.path.join(DRIVE_MODEL_DIR, f)
    if os.path.isfile(src):
        shutil.copy2(src, dst)
    elif os.path.isdir(src):
        shutil.copytree(src, dst, dirs_exist_ok=True)
print(f"Model uploaded to Google Drive: {DRIVE_MODEL_DIR}")
print("02_evaluation.ipynb에서 SFT_IMPROVED_MODEL_05B_PATH로 이 경로를 사용하세요.")
```

```
Mounted at /content/drive
Model uploaded to Google Drive: /content/drive/MyDrive/llm-math-models/qwen2.5-0.5b-math-sft-improved-mc
02_evaluation.ipynb에서 SFT_IMPROVED_MODEL_05B_PATH로 이 경로를 사용하세요.
```

6.1 1.5B SFT Improved 모델 학습 및 Drive 업로드

0.5B와 동일한 MC objective로 1.5B 모델을 학습하고 Google Drive에 저장합니다.

In [ ]:
```
# 1.5B 모델 로드 및 LoRA 적용
MODEL_ID_15B = "Qwen/Qwen2.5-1.5B"
OUTPUT_DIR_15B = "./outputs/03_sft_improved_mc_1.5b"

tokenizer_15b = AutoTokenizer.from_pretrained(MODEL_ID_15B, trust_remote_code=True)
if tokenizer_15b.pad_token is None:
    tokenizer_15b.pad_token = tokenizer_15b.eos_token
    tokenizer_15b.pad_token_id = tokenizer_15b.eos_token_id

model_15b = AutoModelForCausalLM.from_pretrained(
    MODEL_ID_15B,
    torch_dtype=torch.bfloat16,
    device_map="auto",
    trust_remote_code=True,
    attn_implementation=ATTN_IMPL,
)

lora_config_15b = LoraConfig(
    r=16,
    lora_alpha=32,
    lora_dropout=0.05,
    bias="none",
    task_type=TaskType.CAUSAL_LM,
    target_modules=["q_proj", "k_proj", "v_proj", "o_proj", "gate_proj", "up_proj", "down_proj"],
)
model_15b = get_peft_model(model_15b, lora_config_15b)
# model_15b.gradient_checkpointing_enable()  # LoRA+checkpoint 시 grad_fn 오류 가능 → 비활성화
model_15b.print_trainable_parameters()

if USE_TORCH_COMPILE and hasattr(torch, "compile"):
    model_15b = torch.compile(model_15b, mode="reduce-overhead")

# 1.5B용 데이터셋 및 DataLoader
train_dataset_15b = MCDatasetAligned(train_data, tokenizer_15b)
eval_dataset_15b = MCDatasetAligned(eval_data, tokenizer_15b)
train_loader_15b = DataLoader(
    train_dataset_15b, batch_size=BATCH_SIZE, shuffle=True,
    num_workers=NUM_WORKERS, pin_memory=True, collate_fn=collate_batch,
)
eval_loader_15b = DataLoader(
    eval_dataset_15b, batch_size=BATCH_SIZE, shuffle=False,
    num_workers=NUM_WORKERS, pin_memory=True, collate_fn=collate_batch,
)
```

```
trainable params: 18,464,768 || all params: 1,562,179,072 || trainable%: 1.1820
```

7. 요약

- **학습 objective**: 정답 옵션 continuation의 log-likelihood가 softmax over options에서 최대가 되도록 cross-entropy로 학습
- **평가 objective (mathQA)**: 동일하게 multiple choice에서 정답 옵션의 log-likelihood가 최대인지 확인
- **결과**: 학습과 평가의 objective가 일치하여, mathQA 등 multiple choice 평가에서 더 나은 성능을 기대할 수 있습니다.