

EasyR1 Prompt

```
<|im_start|>system
You are Qwen, created by Alibaba Cloud. You are a helpful assistant. You FIRST
think about the reasoning process as an internal monologue and then provide the
final answer. The reasoning process MUST BE enclosed within <think> </think>
tags. The final answer MUST BE put in \boxed{ }.<|im_end|>
<|im_start|>user
<|vision_start|>{image_token}<|vision_end|>
{question}<|im_end|>
<|im_start|>assistant
```

Figure 27: Prompt for EasyR1 training and evaluation.

VeRL Training and Evaluation Prompt

A conversation between User and Assistant. The user asks a question, and the Assistant solves it. The assistant first thinks about the reasoning process in the mind and then provides the user with the answer. The reasoning process and answer are enclosed within <think> </think> and <answer> </answer> tags, respectively, i.e., <think> reasoning process here </think> <answer> answer here </answer>. User: {}
Assistant:

Figure 28: Prompt for VeRL training on Omni-math-train and evaluation on Omni-math-eval and MATH500.

Mistral & Magistral Prompt

A conversation between User and Assistant. The user asks a question, and the Assistant solves it. The assistant first thinks about the reasoning process in the mind and then provides the user with the answer. The reasoning process and answer are enclosed within <think> </think> and <answer> </answer> tags, respectively, i.e., <think> reasoning process here </think> <answer> answer here </answer>. User: {question}.
Assistant:

Figure 29: Prompt for Mistral and Magistral model evaluation.

E. Broader Impacts

The potential negative social impacts of our method align with those typically associated with general LLM reasoning technologies. We emphasize the importance of adhering to the principles of fair and safe deployment in LLM systems.