A valid explanation, then an invalid move

GM guess_feedback: c<red> r<green> a<red> n<green> e<red>

Model explanation: … I need to find a word with 'r' and 'n' in the correct positions, and avoid the letters 'c', 'a', and 'e'.
guess: **ringer**

6 letters | uses e | r, n misplaced

GM Error: The length of the guessed word is not 5. **ABORT**

Real transcript, broad-SFT checkpoint, Qwen3.5-2B.

✗ Whole-dialogue DPO

Successful vs. failed game

→ prose before its moves · clemscore 43.85 → **12.35**

↓

✓ Turn-local DPO

Same history h : original reply vs. one mechanical error

guess: brink ✓

guess: brin ✗

Hold the history fixed; supervise only the next move.

Acquire, Repair, Preserve

A Acquire

SFT on success-only games, ≤ 700 per game

B Repair

Turn-local DPO on mechanical Wordle errors

C Refine

Branch DPO on the model's own play

D Preserve

Keep 85% of the LoRA update

PUBLIC CLEMSCORE

10.67 → 38.92

CLOSED IN-DOMAIN

13.41 → 41.17†

CLOSED OUT-OF-DOMAIN

3.72 → 7.88

Official final evaluation, Qwen3.5-2B base → ours.

Above the official 4B and 9B baselines on public clemscore; most gain comes from broad SFT. OOD gains concentrate in Wordle variants.

Nan Li
Utrecht University
n.li@uu.nl · nanli.nl

† Paper numbers (snapshot `2dd5a533`). After the 18 Sep re-scoring: closed ID 6.70 → 40.24; public and OOD unchanged.