Can an Easy-to-Hard Curriculum Make Reasoning Emerge in Small Language Models? Evidence from a Four-Stage Curriculum on GPT-2

Xiang Fu

Faculty of Computing and Data Sciences, Boston University   xfu@bu.edu

TL;DR: We present Cognivolve, a four-stage curriculum that teaches GPT-2 small to reason—from word-level tasks up to symbolic inference—unlocking 7.8× more specialized heads, doubling convergence speed, and producing richer attention patterns without adding parameters or compute.

The Task

Teach a small transformer to solve progressively harder reasoning problems — and to reveal its intermediate thought process.

The Method

- **Curriculum principle:** present tasks in a strictly easy → hard order to mimic human skill acquisition
- **Specialization tracking:** checkpoint every 500 steps and inspect attention heads with saliency probes to reveal emerging circuits
- **Stage-adaptive learning rates:** lower the peak LR at each stage so the model stabilizes on harder tasks while retaining earlier gains

Cognivolve Curriculum

The model meets tasks in strictly ascending difficulty while its optimizer, embeddings and weights evolve continuously. This mirrors human learning, preserves early gains and lets us watch specialised circuits emerge in real time.

"Is 7 greater than 3?"

"If Tom has 3 apples and buys 2 more, how many apples does he have?"

"A recipe calls for ¾ cup of sugar. You are tripling the recipe. How many cups of sugar are needed?"

"Alice is twice as old as Bob. In 4 years their combined age will be 50. How old is Alice now?"

Simple $\xrightarrow{\text{G2L}}$ **Basic** $\xrightarrow{\text{L2R}}$ **Intermediate** $\xrightarrow{\text{R2X}}$ **Complex**

Simple	Basic	Intermediate	Complex
Teaches token copy & grounding	Introduces two-step logic & short chains	Requires latent state across clauses	Demands open-domain multi-fact reasoning with world knowledge

Experiments

Strategies

- **Baseline:** Conventional single-phase pre-training over the entire curriculum corpus with fixed learning rate.
- **Curriculum:** Continuous training: weights and Adam optimizer moments carry across stages. No task-specific heads or adapters are introduced.
- **Ablations / controls:**
  - **Stage-shuffle:** Same four splits, but presented in random order (breaks easy-to-hard progression).
  - **Mixed-stage fine-tune:** half an epoch over the full corpus after Stage 4 to probe final-answer recovery.

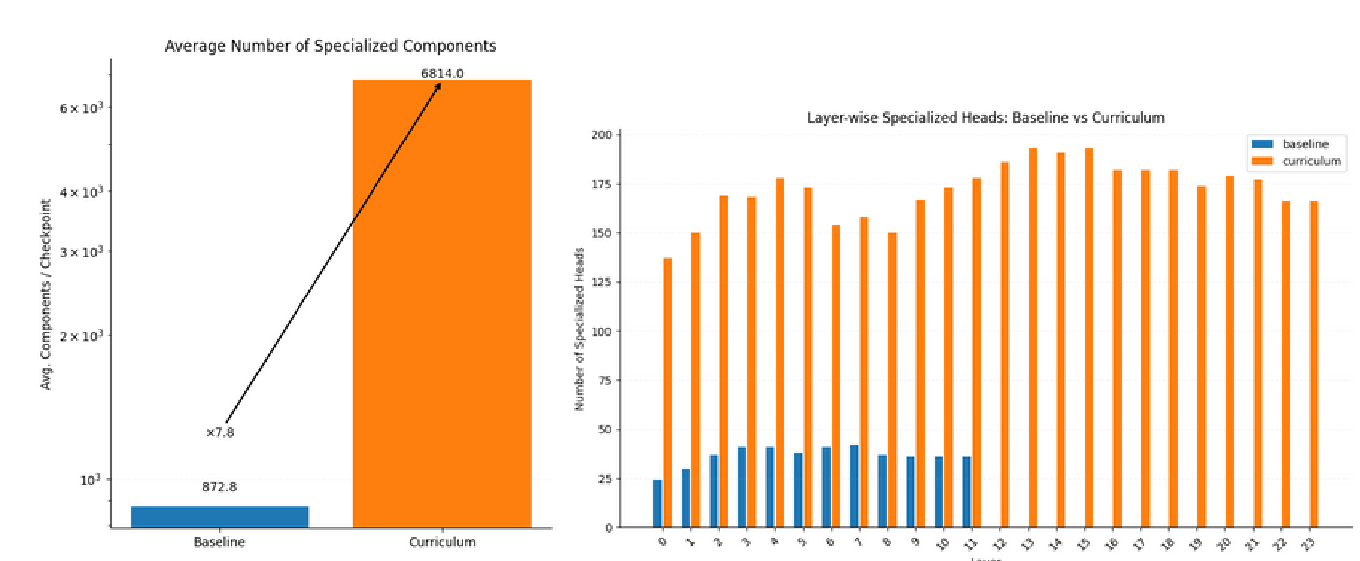
Model Architecture

GPT-2 small with the original byte-pair tokenizer and sinusoidal positional encodings.

Datasets

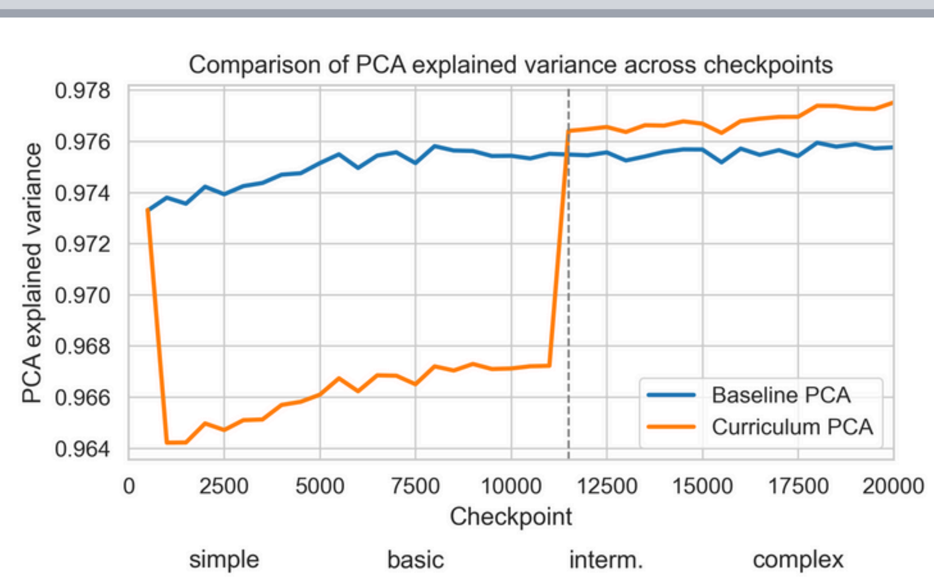
The curriculum is carved from the **Facebook Natural Reasoning** corpus and grouped by heuristic difficulty into four splits of 5 k, 4.7 k, 3.8 k, and 2.9 k instances (Stages 1-4).

Results -Specialized Heads



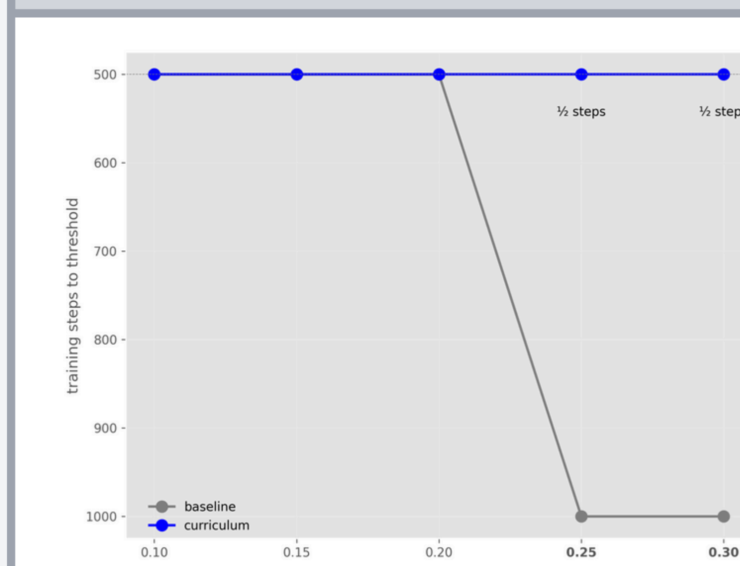
Curriculum unlocks nearly eight times more specialized attention heads than baseline training and redistributes them throughout the depth of the model, with pronounced gains in the late (reasoning-associated) layers.

Results - Global Representation Geometry



- Early syllabus stages scatter variance: curriculum's PCA score starts ≈ 0.9 pp below baseline, signaling broader, less-structured representations.
- Stage-4 hand-off flips the picture: curriculum jumps +1.1 pp, overtakes baseline, and keeps a ~0.13 pp edge through 20 k steps.
- Curriculum first explores many directions, then compacts them into a tighter low-rank manifold—evidence of progressive abstraction rather than brute memorisation.

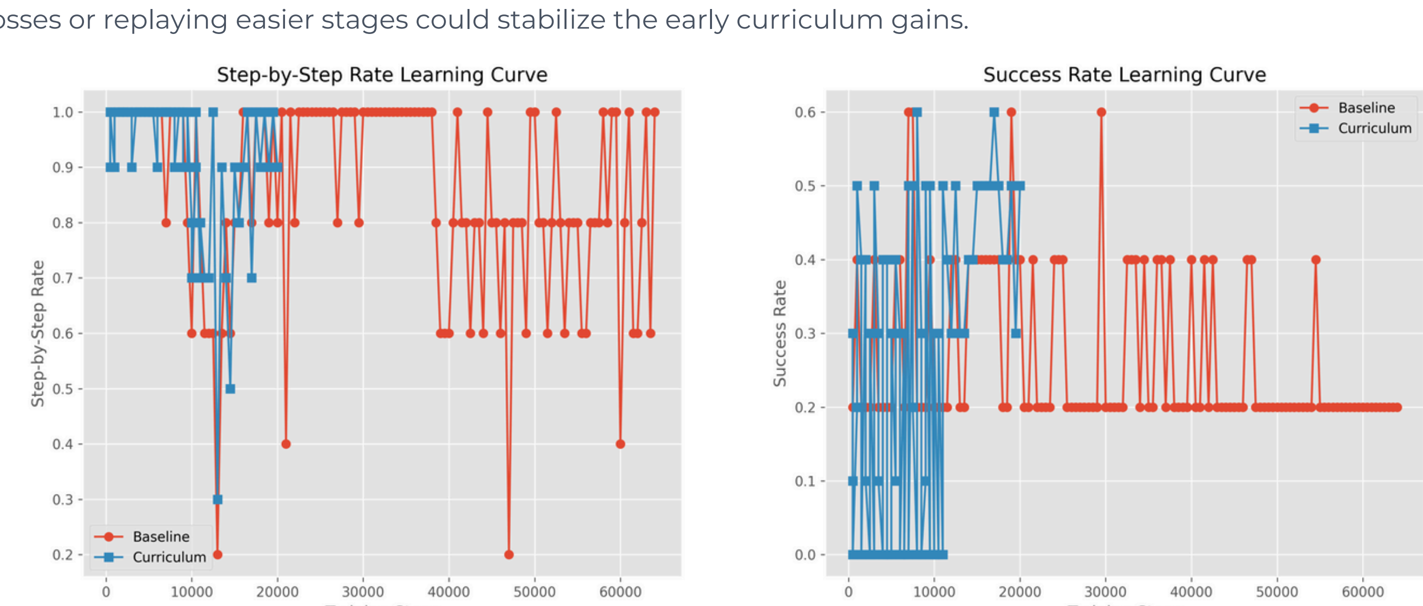
Results - Sample Efficiency



- No gain on easy targets: both models hit 0.10–0.20 success in the first 500 updates.
- Curriculum halves the training steps needed to reach 0.25 / 0.30 success (500 vs 1 000 → 2 × faster, ~2 × compute savings).
- Baseline can surpass curriculum only with > 50 k extra updates—evidence that structured progression yields quicker wins, even if longer fine-tuning can narrow the gap later.

Results - Learning Curves – Fast but Fragile Gains

- **Format mastery ≠ task mastery.** Curriculum models hit near-perfect step-by-step rates almost immediately, but their final-answer success stays modest—evidence that "showing the work" can outpace actually solving the problem.
- Given enough steps the vanilla baseline also learns the reasoning template while success plateaus at 0.20–0.40, suggesting the current objective weighs explanation style more than end accuracy; mixing in answer-focused losses or replaying easier stages could stabilize the early curriculum gains.



Results - Component Emergence

- Low-level "copy" circuitry is saturated already: induction heads appear in equal numbers of baseline and curriculum
- Curriculum unlocks >2x more reasoning heads and 30% pattern matches, expanding the model's higher-order toolkit
- Net gain of +158 specialized components shows that staged exposure creates representational niches a single-phase baseline never discovers