

## Problem Formulation: A Curious Little Problem

The goal of this work is to understand something interesting but possibly unnecessary. But don't worry—we'll formalize it:

**Definition** (The Problem). Formally, this is the problem we're pretending is novel. It involves mysterious variables, vague assumptions, and a deep sense of academic urgency.

Here's a visual representation, just to look professional:

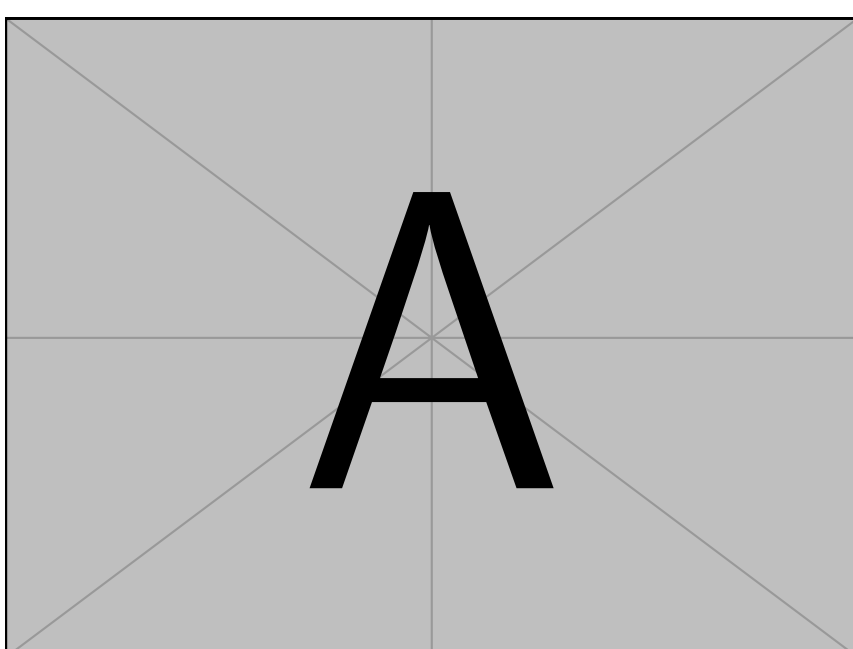


Figure 1. A figure. Definitely helps, right?

As you can see, this figure captures the essence of the problem.

**Insightful Highlight Box.** This box exists to draw your attention. What to? Not entirely sure. But it's highlighted, so it must be important.

## Existing Approaches: A Brief Stroll Through Literature

A quick literature review to show we read a few papers:

- Classical physics seems somehow relevant [1].
- Honestly, we just needed a second bullet point.

## Overview and Contributions

This work marks a pivotal moment in the history of ML and AI:

1. A theoretical argument for why **AGI** is inevitable.
2. We actually build such a model with *≈one trillion* parameters.
3. Empirical results showing perfect scores on this year's IMO (not even out).

## Theoretical Results: Unquestionable Mathematical Rigor

Now that we've stated the problem, let's prove some theorems. Or at least write some that look impressive:

**Theorem** (AGI Emergence). If your model has at least  $10^{12}$  parameters and a sufficiently dramatic name, it will become sentient with probability  $\epsilon > 0$ .

**Definition** (Superintelligence Gradient). The partial derivative of model confidence with respect to Twitter hype. Empirically observed to be strictly increasing.

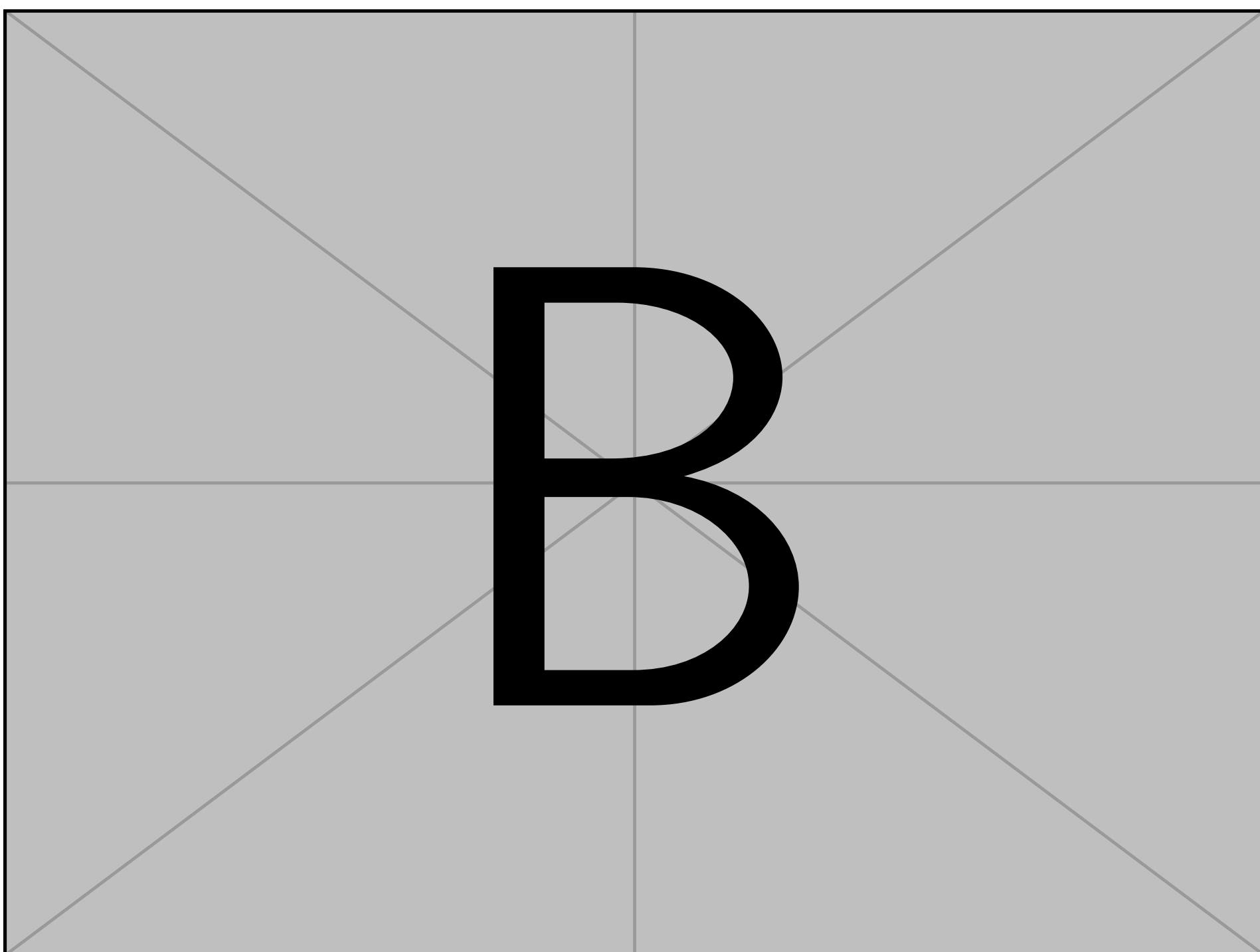


Figure 2. Evidence of emergent intelligence.

## Failed Approaches: We Suffered So You Don't Have To

Naturally, we tried some things that didn't work:

- **Copying ChatGPT:** Ethically dubious, technically tempting.
- **Brute Force AGI:** Required more GPUs than the planet currently owns.

**Theorem** (Overfit Paradox). As model size increases, accuracy increases on everything except the test set.

## Proposed Method: Definitely Not Just a Bigger Transformer

Our model is designed with scalability in mind.

**Definition** (Trillion-Parameter Transformer). A neural network so large, it requires its own power grid and a team of therapists.

**Theorem** (Universal Solver). Given enough data and a sufficiently vague benchmark, our model achieves state-of-the-art on at least one metric.

## Experimental Results: Trust Us, We Have Graphs

We tested our model on:

- IMO. We got a perfect score. Somehow.
- The Turing Test. The judges asked it for dating advice and were convinced.
- StackOverflow. It answered "it depends" with 95% accuracy.

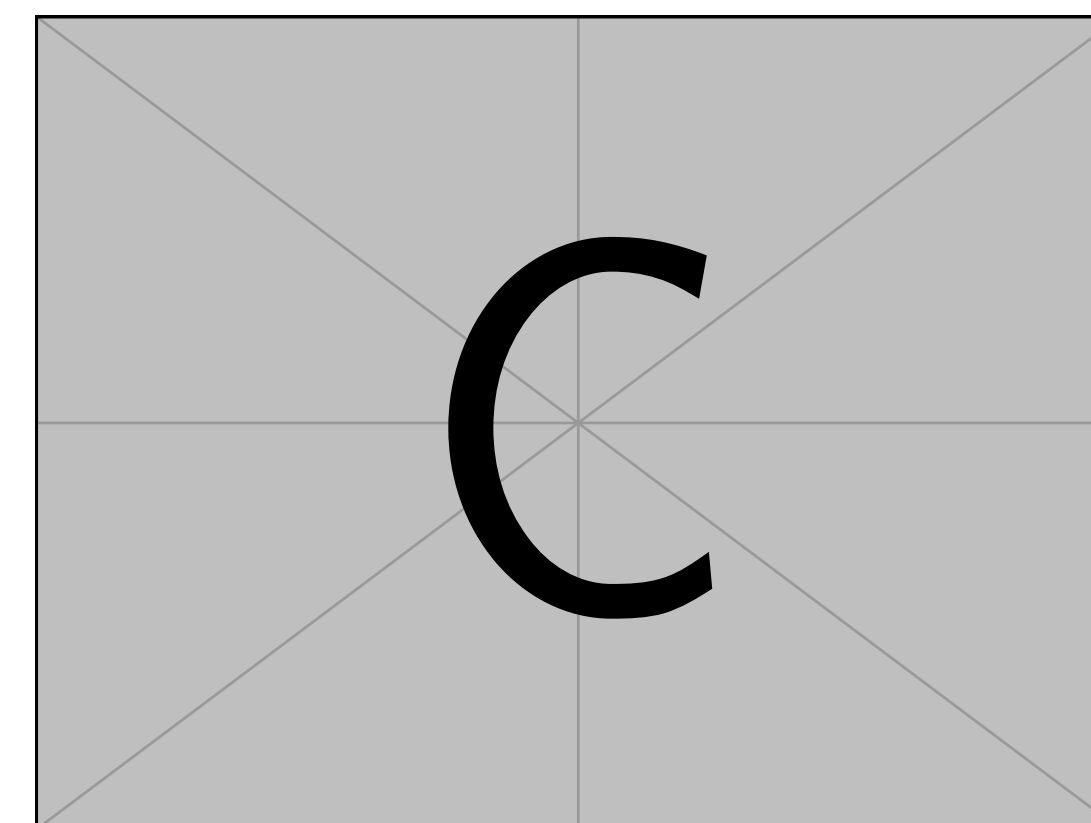


Figure 3. Blue line = us. Orange line = others. Need we say more?

## Conclusions and Next Steps

1. Our method is clearly on track to surpass human intelligence.
2. Future work: deploy model on Mars and let it evolve in isolation.

**Final Thought.** This project started as a joke. Now it's under review at NeurIPS.

[1] Newton. Philosophiae naturalis principia mathematica. Innys, 1726. url: <https://books.google.com/books?id=WeZ09rjv-1kC>.