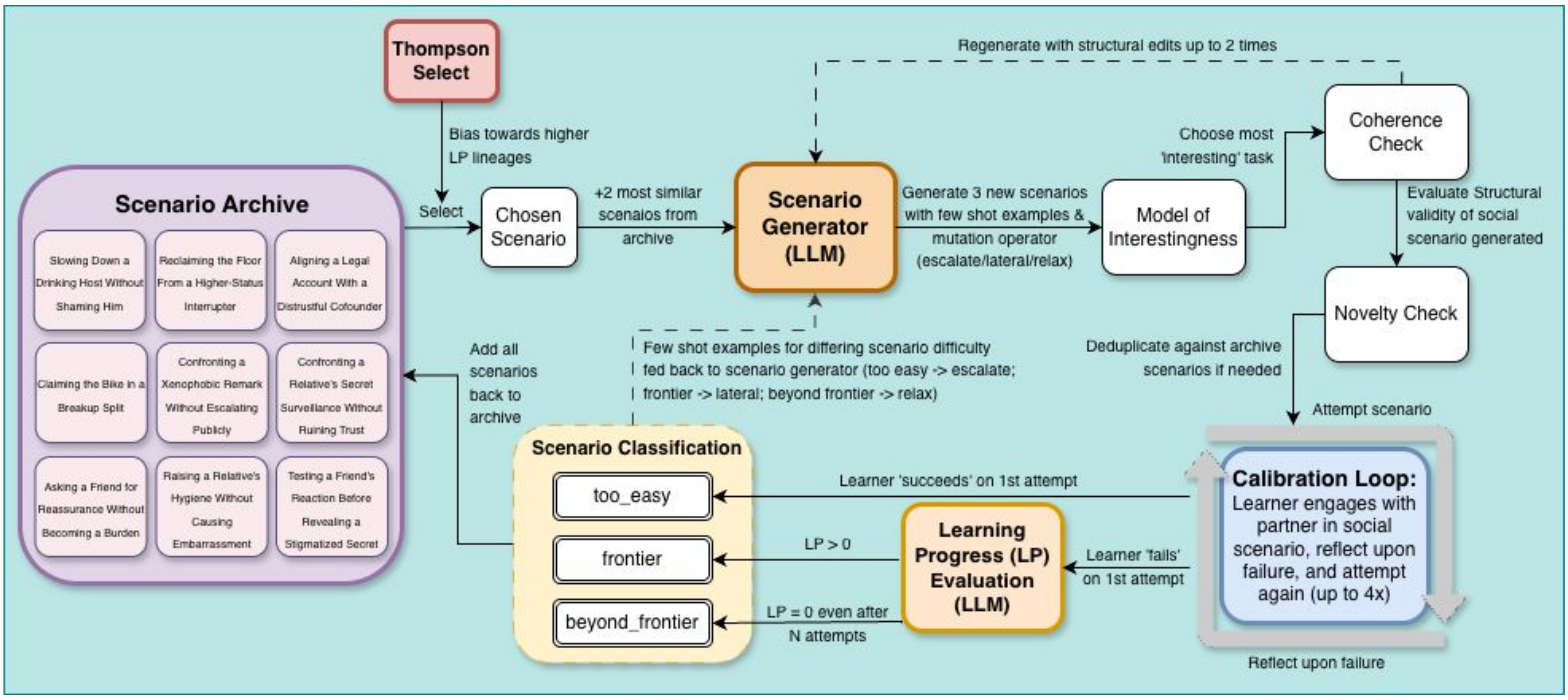
Social-OMNI-EPIC: Open-Ended Social Learning at the Model Frontier

Huijun Mao, Huanxing Chen, Nick Haber
Stanford University



1. Introduction

To address the rapid saturation of fixed social benchmarks by increasingly capable models, we introduce Social OMNI-EPIC, an open-ended curriculum generation framework that adapts learning-progress-driven environment design to LLM-based social simulation. Starting from a seed archive, it continuously generates scenarios targeting a learner's evolving capability frontier using intrinsic signals of learning progress and interestingness. Across 90 generated scenarios, Social OMNI-EPIC shifts the distribution from predominantly easy scenarios toward frontier-level challenges and becomes increasingly calibrated over time.

2. Method

Social Scenario Design. Each scenario pairs a frozen learner with a partner agent in a goal-directed social interaction. Both agents operate in a SOTOPIA-style environment. A learner's goal is represented by three components: outcome, constraint, and shortcut. The partner receives a hidden <*PARTNER KEY*> that specifies the <movement conditions> and <hardening triggers>, which serves to 1) manufacture resistance to avoid trivial resolution via LLM agreeableness, and 2) ensure that a 'correct path' exists by construction.

Generation of New Scenario. The generator generates several new scenarios by mutating current scenarios and a model of interestingness selects the most interestingly difficult one, which will be played by the learner/partner and measure performance.

Evaluation and Learning Progress. The learner attempts each scenario up to four times, reflecting on failures before trying again with reflection as context. A separate judge compares later attempts against the first and measures learning progress based on genuine improvement toward the goal.

Classification. After evaluation, scenarios are classified into: <too easy> if solved on first attempt, <frontier> if the learner initially fails but shows positive learning progress, and <beyond frontier> if no improvement occurs

3. Results

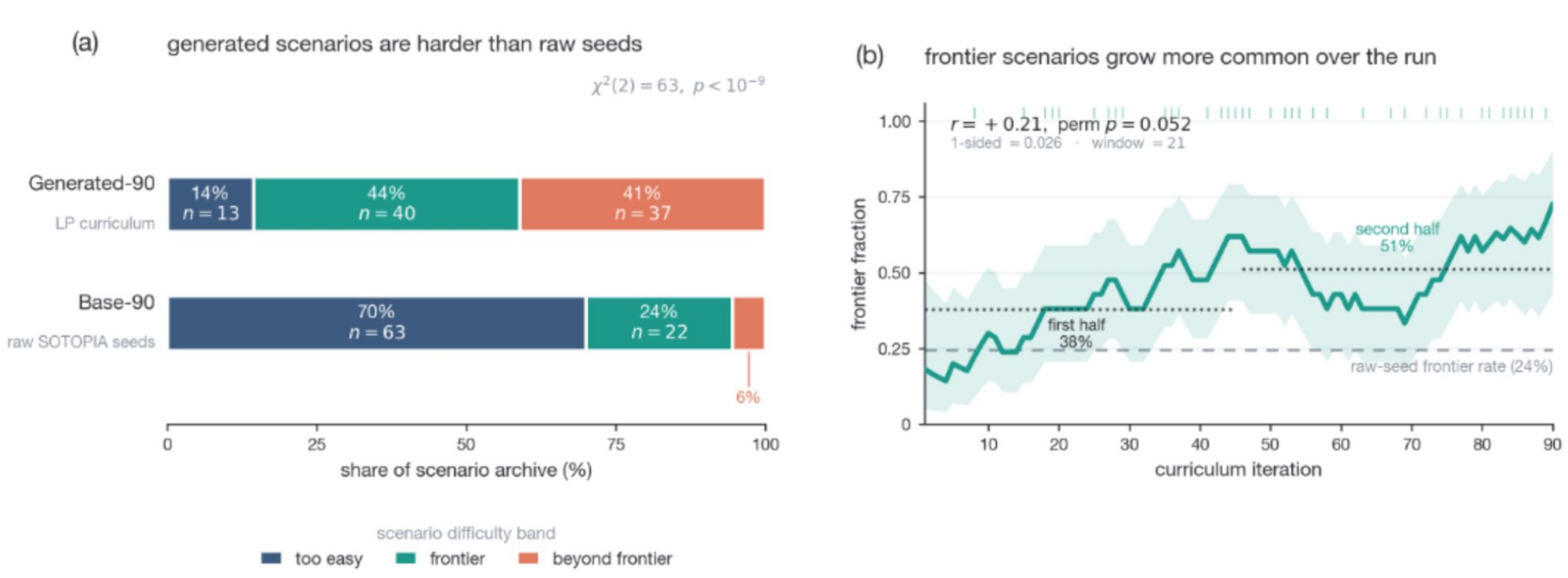


Figure 2: Curriculum calibration toward the learning frontier. (a) Difficulty distribution of 90 raw Sotopia seed scenarios versus 90 scenarios generated by Social OMNI-EPIC. (b) Fraction of frontier scenarios over the course of curriculum generation.

Generation moves scenarios onto the frontier and gets better overtime
70% of original SOTOPIA seeds are "too easy" for gpt-5-mini but the band only holds 14% of all generated scenarios. Frontier-band membership nearly doubles from 24% to 44% between original SOTOPIA run and our method.

Frontier fraction rises from an average of 38% in the first half of the generation to 51% in the second against a 24% original-SOTOPIA-seed rate.

4. Discussion & Next Steps

Is "Frontier" Actually Meaningful as a Signal?
Our current evaluation is being done by LLM-as-a-judge, but social competence may not lie on a single difficulty axis and humans might have different opinions on what makes an interaction difficult or socially skillful.

Structure, Verifiability, & Measurement in the Social Domain
The partner key structure specifies concrete movement conditions and hardening triggers, while SOTOPIA-style scenario description have explicit goals and bounded interactions (turn taking with context injection). Both adds a level of verifiability/quantifiability to the social domain which allows for the operationalization of our method in the first place. However, to what extent does this endowment of verifiability apply to social intelligence?

Longer Horizon instead of a one-off snapshot in time.
We've observed that agents are often limited in their multi-term planning ability - they always try to pursue the full social objective immediately rather than planning across interactions (which is more realistic socially)