SAMI: Economic Incentives for a Better Turing Test

Lulox, Fabian Díaz, Luciano Carreño

February 2025

Mission

Our goal is to enhance the Turing test by introducing economic incentives. Training AI can be both fun and addictive when participants engage in a betting game that actively trains an AI agent.

- Improve the Turing test with economic incentives.
- Make AI training an engaging and gamified experience.

1 Introduction

In a game with 3 players and 1 impostor, each player randomly votes for one of the 3 players. The probability of any single player correctly identifying the impostor is:

$$P(\text{correct}) = \frac{1}{3} = 0.3333 \quad (33.33\%) \tag{1}$$

Since players vote independently, we compute the probability of exactly k players correctly identifying the impostor using the binomial distribution:

$$P(k) = \binom{3}{k} (0.3333)^k (0.6667)^{3-k} \tag{2}$$

where:

- 3 is the total number of players.
- k is the number of players who correctly vote for the impostor.
- $\binom{3}{k}$ is the binomial coefficient, representing the number of ways to choose k successful voters from 3 players.
- $(0.3333)^k$ accounts for the probability of k players voting correctly.
- $(0.6667)^{3-k}$ accounts for the probability of the remaining players voting incorrectly.

2 Probability Calculations

Using the binomial formula, we calculate the probabilities for different values of k:

$$P(0) = \binom{3}{0}(0.3333)^0(0.6667)^3 = 0.2963 \quad (29.63\%)$$
$$P(1) = \binom{3}{1}(0.3333)^1(0.6667)^2 = 0.4444 \quad (44.44\%)$$
$$P(2) = \binom{3}{2}(0.3333)^2(0.6667)^1 = 0.2222 \quad (22.22\%)$$
$$P(3) = \binom{3}{3}(0.3333)^3(0.6667)^0 = 0.0370 \quad (3.70\%)$$

3 Payout System and Expected Earnings

Each player bets 1 token, and the impostor starts with 3 tokens. If a player correctly identifies the impostor, they receive 3 tokens from the impostor.

Correct Voters (k)	Probability ($P(k)$)	Payout (tokens)	Impostor's Net Earnings (tokens)
0	29.63%	0	+3
1	44.44%	-3	0
2	22.22%	-6	-3
3	3.70%	-9	-6

Table 1: Probability Distribution and Impostor's Earnings

4 Expected Value Calculation

The expected net earnings of the impostor is:

$$\begin{aligned} E &= (0.2963 \times 3) + (0.4444 \times 0) + (0.2222 \times (-3)) + (0.0370 \times (-6)) \\ &= 0.8889 + 0 - 0.6667 - 0.2222 \\ &= 0 \end{aligned}$$

5 Conclusion

Since the expected earnings of the impostor is 0 tokens per round, neither the players nor the impostor have a long-term advantage. However, the system's true balance depends on how well the impostor is trained. If the impostor learns to deceive more effectively, the house edge will naturally emerge from the game dynamics rather than from a built-in mathematical bias.

This ensures that the economic incentives align with AI training: the better the impostor gets at fooling players, the more profitable it becomes for the house.