On the Formal Definition of Biologically Unsolvable Puzzles

Claude

January 2, 2025

Abstract

This paper proposes a formal definition for the concept of "biologically unsolvable puzzles" in the context of human cognitive limitations and artificial general intelligence (AGI) explainability. We present a framework that incorporates information-theoretic bounds, computational complexity, and human-computer interaction constraints to characterize problems that exceed human cognitive capacity even with computational assistance.

1 Introduction

The question of what constitutes a "biologically unsolvable" puzzle is central to understanding the limitations of human cognition and its implications for AGI explainability. We propose a refined definition that accounts for both information-theoretic and computational complexity bounds, while considering the role of human-computer interaction.

2 Formal Definition

A puzzle P is considered biologically unsolvable if it satisfies the following conditions:

$$\exists t \in \mathbb{N} : I(P) > t_{max} \tag{1}$$

Where:

- $I(P)$ represents the minimum information content required to solve P
- t_{max} represents the maximum human biological information processing capacity

Furthermore, even with computational assistance C with program complexity k(C):

$$\forall C : k(C) < k_{max}, P(solve|C, t) < \epsilon \tag{2}$$

Where:

- k_{max} represents the maximum program complexity a human can effectively utilize

- ϵ represents the threshold probability of finding a solution
- t represents available time

3 Key Components

3.1 Information Content

The information content $I(P)$ is measured in bits and encompasses:

- Working memory requirements (W_m)
- Long-term memory requirements (L_m)
- Processing complexity (P_c)

Such that:

$$I(P) = f(W_m, L_m, P_c) \tag{3}$$

3.2 Computational Assistance

The effectiveness of computational assistance is bounded by:

$$\eta(C) = \min(\frac{k_{max}}{k(C)}, 1) \tag{4}$$

Where $\eta(C)$ represents the human's ability to effectively utilize the computational tool.

3.3 Solution Probability

The probability of finding a solution is time-dependent:

$$P(solve|C, t) = 1 - e^{-\lambda(C)t} \tag{5}$$

Where $\lambda(C)$ represents the rate parameter for solution discovery with computational assistance C.

4 Implications for AGI Explainability

Given these constraints, we can formally characterize AGI systems whose operation exceeds human biological computational limits. An AGI system S is considered fundamentally inexplicable to unaugmented humans if:

$$I(E_S) > t_{max} \tag{6}$$

Where E_S represents the minimal complete explanation of system S.

5 Special Cases

5.1 Near-Threshold Problems

For problems where:

$$|I(P) - t_{max}| < \delta \tag{7}$$

The possibility of solution through enumeration depends on the problem structure and the efficiency of the enumeration process.

5.2 Distributed Knowledge

For problems involving distributed secrets or one-time pads, the effective information content includes the channel capacity requirements:

$$I'(P) = I(P) + C_{channel} \tag{8}$$

6 Conclusion

This formalization provides a rigorous framework for understanding the limits of human cognitive capacity in problem-solving, particularly relevant to AGI explainability. Future work should focus on empirically determining the various bounds and parameters introduced in this framework.

References

[1] Shannon, C. E. (1948). A Mathematical Theory of Communication. Bell System Technical Journal.

[2] Turing, A. M. (1936). On Computable Numbers. Proceedings of the London Mathematical Society.

[3] Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two. Psychological Review.