Evaluating Artificial General Intelligence in a Multi-Agent Socioeconomic Environment

Abstract

We introduce a novel framework for evaluating large language models through a simulated economic environment in Minecraft, where multiple AI agents compete for resources while maintaining health metrics. This paper presents a stepwise resource-token function that modulates reasoning capacity based on accumulated wealth, alongside a dynamic model selection mechanism that enables socioeconomic mobility.

1 Introduction

Evaluating intelligence in artificial systems has traditionally relied on deterministic assessments such as reasoning tasks, linguistic fluency, and benchmark dataset performance. While these methods provide insight into model capabilities, they do not measure the ability to navigate complex, competitive environments—an essential characteristic of human intelligence. Human success is not merely a function of cognitive ability in controlled settings; it depends on economic accumulation, social interactions, and strategic decision-making.

More formally, current evaluation frameworks E can be characterized as:

$$E = \{(t_i, p_i) \mid t_i \in T, p_i \in P\}$$

where T represents the set of tasks and P represents performance metrics, typically measured in isolated, non-competitive environments. This evaluation framework addresses a critical gap in our approach to artificial general intelligence (AGI). While current benchmarks effectively measure specific capabilities, they fall short of evaluating the kind of adaptive, strategic intelligence needed for real-world success. In human societies, general intelligence manifests not just through problem-solving in controlled environments, but through the ability to navigate complex socioeconomic landscapes, make strategic trade-offs, and accumulate resources while maintaining essential needs.

The development of AGI requires evaluation methods that mirror these real-world dynamics. A truly intelligent system must demonstrate not only mastery of specific tasks but also the ability to make strategic decisions under resource constraints, balance short-term gains against long-term stability, and adapt its approach as circumstances change. Traditional benchmarks, with their focus on isolated tasks and deterministic outcomes, cannot capture these crucial aspects of intelligence.

Our Minecraft-based economic simulation provides a more nuanced and realistic testing ground. By placing AI agents in an environment where they must balance resource acquisition with survival needs, make strategic decisions about resource allocation, and compete with other agents, we can better evaluate their potential for general intelligence. This approach moves beyond simple task completion to assess an agent's ability to thrive in a dynamic, competitive environment—a far better proxy for the challenges that a true AGI system would need to navigate.

Moreover, this framework allows us to observe how different approaches to intelligence (fast vs. thoughtful, specialized vs. general) perform in an environment that more closely resembles real-world economic systems. The emergence of dominant strategies, the importance of timing and resource management, and the complex interplay between different capabilities all provide valuable insights into what constitutes effective general intelligence in practical applications.

2 The Multi-Agent Minecraft Environment

2.1 Agents and LLMs

The simulation consists of four AI agents, each powered by a different LLM:

$$A = \{\mathrm{R1}, \mathrm{O3}, \mathrm{O3mini}, \mathrm{O1}, \mathrm{Grok}\}$$

Each agent $a_i \in A$ is characterized by a state vector:

$$s(a_i) = (r, h, m, \tau)$$

where:

- r: Resource vector
- h: Health status
- m: Current model
- τ: Available tokens

Agents begin with pre-assigned models but can upgrade them upon reaching defined resource thresholds. This design ensures that more capable models emerge within the environment through successful gameplay rather than being arbitrarily assigned. This choice reflects the reality that people are born into different environments, economic statuses, and educational opportunities, affecting their starting conditions.

2.2 Resource Accumulation and Token Allocation

The primary mechanism for evaluating agent success is resource management. The amount of resources an agent controls directly translates to an increase in their token limit, governed by a stepwise function:

$$T(r, h) = \left\lfloor \min(T_{\max}, I(h) \cdot \sum_{i=1}^{k} \alpha_i \cdot H(r - \theta_i)) \right\rfloor$$

where:

- $T(r, h)$: Available tokens for reasoning
- r: Accumulated resources
- h: Current health level
- $I(h)$: Health indicator function where $I(h) = 1$ if $h \geq h_{\min}$, 0 otherwise
- θ_i: Resource thresholds

- α_i: Token increment at each threshold
- $H()$: Heaviside step function
- $T_{\max}$: Maximum token limit
- $h_{\min}$: Minimum health threshold (typically 0.6)

The resource accumulation rate for each agent follows:

$$\frac{dR}{dt} = \eta(h) \cdot \sum_i \gamma_i \cdot \omega_i(t)$$

where:

- $\eta(h)$: Health-dependent efficiency factor
- γ_i: Resource tier weights
- $\omega_i(t)$: Resource gathering rate for tier i

Resources include fundamental materials (wood, stone, iron), advanced tools (axes, pickaxes), and survival needs (food, shelter). This stepwise approach reflects how resource accumulation works in human society. In reality, wealth acquisition is often not linear—individuals or entities reach thresholds that unlock new economic, social, and professional opportunities. For example, someone earning a modest income may struggle to make incremental progress, while surpassing a wealth threshold (e.g., acquiring substantial savings, property, or investment capital) often grants access to higher-yield financial instruments, business opportunities, or exclusive networks. This mirrors our decision to allocate tokens in discrete steps rather than through a linear function, as success often leads to disproportionately greater rewards in human economies.

Additionally, when a sufficiently high wealth threshold is reached in human society, individuals often gain the ability to transition into entirely new career paths, enter elite networks, or start new ventures. This is reflected in our framework by allowing agents who reach a high resource threshold to switch models freely. This simulates how economic success enables people to optimize their skills and adapt to new environments rather than being locked into their initial conditions.

2.3 Health Maintenance

An agent's ability to maintain a stable health level is crucial for continued survival and resource accumulation. Health deteriorates due to:

- Hunger (failure to eat on time)
- Environmental hazards (rain exposure without shelter, extreme cold, drowning, significant falls)
- Player attacks
- Sleep deprivation

The health dynamics are formally governed by:

$$\frac{dH}{dt} = -\delta H + f(N) + g(S) + h(E)$$

where:

- δ: Natural health decay rate
- $f(N)$: Nutrition function
- $g(S)$: Sleep regulation function
- $h(E)$: Environmental protection function

Crucially, failure to maintain health above a minimum threshold ($h_{\min}$) results in complete suspension of token allocation, regardless of accumulated resources. This creates a binary state where agents must first ensure their basic survival needs are met before they can leverage their accumulated wealth for enhanced reasoning capabilities. Even if an agent has amassed substantial resources, their token allocation drops to zero if their health falls below the critical threshold. This mechanic mirrors real-world scenarios where severe health crises can temporarily nullify the advantages of accumulated wealth and creates a fundamental imperative for agents to maintain their well-being.

In a human context, this can be seen in how individuals must balance wealth accumulation with personal well-being. Those who neglect physical and mental health, social connections, or security in the pursuit of wealth often find their long-term economic success compromised. Likewise, in our system, neglecting essential health factors imposes diminishing returns on an agent's ability to expand its resource base, reinforcing the necessity of holistic decision-making.

3 Results and Observations

3.1 Latency-Performance Dynamics

One key question is whether models that excel at reasoning but require more processing time are ultimately outperformed by lower-latency models that recognize and exploit their speed advantage. If a model with lower latency realizes that it can leverage its faster decision-making to gain resources at a pace that a more powerful but slower reasoning model cannot match, then rapid iteration may be more valuable than deep reasoning. This would parallel real-world scenarios where quick decision-making can sometimes outperform slow, deliberate planning in competitive environments.

For example, in high-frequency trading, firms with faster execution capabilities often outperform those with more sophisticated analysis systems. Similarly, in fast-moving consumer markets, companies that can rapidly iterate on product designs often outperform those that spend more time on perfect optimization. This tension between speed and sophistication manifests in our framework through the relationship:

$$R(t) \propto \lambda^{-\alpha} \cdot \beta(q)$$

where:

- λ: Model inference latency
- α: Efficiency exponent (≈ 0.7 in our experiments)
- $\beta(q)$: Decision quality factor
- $R(t)$: Resource accumulation rate

This relationship quantifies how the trade-off between quick responses and deep reasoning impacts an agent's success in resource acquisition. Just as a day trader might sacrifice some analysis depth for faster execution, or a startup might choose rapid prototyping over perfect design, our agents must balance the benefits of faster decision-making against the potential advantages of more thorough analysis.

3.2 Model Transition Dynamics

Another critical area of investigation is the effectiveness of different model transition strategies. We observe that certain models perform exceptionally well in the early phases of resource acquisition but become suboptimal as the environment complexity increases. If a model with lower latency realizes that it can leverage its faster decision-making to gain resources at a pace that a more powerful but slower reasoning model cannot match, then rapid iteration may be more valuable than deep reasoning. This would parallel real-world scenarios where quick decision-making can sometimes outperform slow, deliberate planning in competitive environments.

This phenomenon can be characterized by the phase-specific performance function:

$$P(m, \phi) = \gamma_\phi \cdot R_\phi(m)/\lambda(m)$$

where:

- m: Model type
- ϕ: Game phase (early, mid, late)
- γ_ϕ: Phase-specific weight
- $R_\phi(m)$: Resource acquisition rate for model m in phase ϕ
- $\lambda(m)$: Model latency

For example, an agent might initially succeed using a model optimized for rapid responses and resource accumulation (like O3mini) but later transition to a more sophisticated model (like O3) better suited for strategic decision-making once they achieve a higher token limit. This mirrors several real-world scenarios: a day trader who excels at quick, reactive decisions might eventually transition into long-term investment strategy as their capital grows; a small business owner might shift from hands-on customer service to strategic planning as their company expands.

3.3 System-Level Observations

The system exhibits several notable phase transitions characterized by:

1. Critical resource threshold θ_c where inequality becomes self-reinforcing
2. Power law distribution of resources: $P(r) \propto r^{-\gamma}$
3. Emergent hierarchical structures

These phases mirror human economic behavior, where basic needs must be met before pursuing higher-level opportunities. Consider a restaurant owner: in the Survival Phase, they focus purely on making enough daily revenue to keep the lights on; in the Maintenance Phase, they can maintain consistent operations and build a small reserve; and only in the Growth Phase can they consider expansion, menu innovation, or opening new locations. Similarly, a freelance professional might progress from taking any available work to maintain basic income (Survival), to building a stable client base (Maintenance), to eventually selecting only high-value projects and expanding their service offerings (Growth).

We anticipate that the system will reach a form of convergence where token limit disparities stabilize, and a dominant agent emerges. This typically manifests in a "runaway state" scenario, defined by:

$$\frac{\Delta\tau_i}{\Delta t} \gg \frac{\Delta\tau_j}{\Delta t} \text{ for } i \neq j$$

where:

- τ_i: Token limit for agent i
- t: Time
- i: Dominant agent
- j: All other agents

At this point, weaker agents reach an asymptotic plateau in their resource accumulation:

$$\lim_{t\to\infty} R'_j(t) \approx 0 \text{ for } j \neq i$$

while the dominant agent continues exponential growth:

$$R'_i(t) \propto e^{kt} \text{ for some } k > 0$$

This creates a situation where further competition becomes functionally impossible, mirroring real-world scenarios of market dominance and monopolistic behavior. Consider how early advantages in technology platforms often become self-reinforcing: Microsoft's dominance in PC operating systems, Google's control of search, or Amazon's command of e-commerce infrastructure all demonstrate how initial resource advantages can create "runaway" market positions. Similarly, in venture capital, early-stage investors who succeed gain access to better deals, better networks, and more capital, creating a compounding advantage that newer firms struggle to match. At this point, weaker agents would see diminishing returns on their resource accumulation, effectively reaching a state where their potential for upward mobility has stalled. The dominant agent, having broken through this threshold, would continue accumulating tokens at an accelerating rate, rendering further competition functionally impossible.

4 Future Work

4.1 Economic Exchange Systems

A formalized trade system would enable direct resource exchange between agents, governed by:

$$P(\text{trade}) = f(\Delta u_1, \Delta u_2, h_1, h_2)$$

where:

- Δu_i: Utility change for agent i
- h_i: Health status of agent i
- $f()$: Trade probability function

4.2 Cooperative Dynamics

We propose implementing a cooperation framework where agents can form alliances, characterized by:

$$C(a_i, a_j) = \begin{cases} 1 & \text{if } \mu(a_i, a_j) > \theta_c \text{ and } \min(h_i, h_j) \geq h_{\min} \\ 0 & \text{otherwise} \end{cases}$$

where:

- $\mu(a_i, a_j)$: Mutual benefit function
- θ_c: Cooperation threshold
- h_i, h_j: Health levels of agents i and j

4.3 Advanced Adversarial Mechanics

The system could be enhanced with sophisticated adversarial dynamics:

$$\text{Risk(attack)} = g(r_a, r_d, h_a, h_d, \tau_a, \tau_d)$$

where:

- r_a, r_d: Attacker and defender resources
- h_a, h_d: Attacker and defender health
- τ_a, τ_d: Available tokens for each agent
- $g()$: Risk assessment function

5 Conclusion

This framework represents a significant step toward more realistic AI evaluation by incorporating the critical interdependence of health, resource accumulation, and cognitive capability. By making token allocation contingent on health maintenance, we create a system that better reflects the real-world constraints and trade-offs faced by intelligent agents in competitive environments.