

Agora: A Drive-Based Framework for Agent Differentiation in Multi-Agent LLM Systems

Zihan Lin

2483102056@stu.imust.edu.cn

Inner Mongolia University of Science and Technology

Research Article

Keywords: Multi-agent LLM systems, Intrinsic motivation, Agent differentiation, Cognitive drives, Task-dependent reasoning

Posted Date: August 6th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10371209/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Agora: A Drive-Based Framework for Agent Differentiation in Multi-Agent LLM Systems

Zihan Lin

Inner Mongolia University of Science and Technology,
School of Materials Science and Engineering,
Baotou 014010, China

2483102056@stu.imust.edu.cn

Abstract. Current multi-agent LLM systems differentiate agents through hand-crafted roles or layer-stacking architectures. We propose Agora, a framework that differentiates agents through six continuous drives (boredom, curiosity, stress, anxiety, empathy, service) that map to LLM parameters, producing parameter-level differentiation without prompt engineering. We evaluate Agora through 375 real API calls on DeepSeek Chat across four task types, comparing RoundTable (3 differentiated agents) against four baselines: Single, MajorityVoting, BestOf3, and AutoGenStyle. Results reveal a task-dependent pattern: RoundTable produces significantly higher quality on analytical tasks ($p < 0.001$ vs AutoGenStyle, Cohen’s $d = 2.26$) but does not outperform simpler baselines on creative tasks (MajorityVoting wins, $p < 0.001$) or short-format generation (AutoGenStyle wins, $p = 0.006$). Closed-loop drive-LLM integration maintained stable quality over 20 consecutive steps. Ablations identify empathy as the most critical drive component. These results suggest that drive-based differentiation shows promise for analytical reasoning tasks, outperforming some multi-agent baselines (AutoGenStyle, BestOf3) with large effect sizes, though the comparison against a single differentiated agent does not reach significance. Simpler coordination strategies remain competitive for creative generation. We discuss the implications for multi-agent LLM system design and identify directions for future work.

Keywords: Multi-agent LLM systems · Intrinsic motivation · Agent differentiation · Cognitive drives · Task-dependent reasoning

1. Introduction

Ask three people an open-ended question and you will receive three different answers. The differences reflect distinct internal states—one curious, one cautious, one empathetic—that shape how each person approaches the problem. We ask whether multi-agent LLM systems can achieve agent differentiation through internal cognitive states rather than external role assignments.

Current multi-agent LLM systems do not exploit this principle. They orchestrate agents through hand-crafted roles (“you are a critic”) or layer-stacking architectures where one agent’s output feeds the next [1, 2, 3]. The agents are externally differentiated by prompt engineering, not internally differentiated by cognitive state. When all agents share the same base model and temperature, their outputs converge on the same high-probability modes.

We propose Agora (Adaptive Group of Reasoning Agents), a multi-agent framework that differentiates agents through six continuous drives—boredom, curiosity, stress, anxiety, empathy, service—inspired by models of intrinsic motivation in cognitive science [4, 5]. The drives evolve according to homeostatic relaxation and cross-drive coupling, then map to LLM parameters (temperature, frequency penalty, presence penalty, max tokens). A TaskRouter classifies prompts and spawns an appropriate number of agents. A RecursiveRoundTable coordinates multi-round draft–critique–polish cycles.

We evaluate Agora through 375 real API calls on DeepSeek Chat, comparing RoundTable against four baselines across four task types. The results reveal a task-dependent pattern: RoundTable produces the highest quality on analytical tasks but is outperformed by simpler baselines on creative and short-format

tasks. These mixed results suggest that not all multi-agent configurations suit all tasks, and the choice of coordination strategy should match the task characteristics.

Contributions. (1) A drive-based differentiation mechanism that maps continuous cognitive drives to LLM parameters, providing parameter-level diversity without prompt engineering. (2) Preliminary evidence that drive-based differentiation outperforms some multi-agent baselines on analytical tasks ($p < 0.001$ vs AutoGenStyle, $d = 2.26$) but does not universally outperform simpler strategies. (3) An ablation study identifying empathy as the most critical drive component, and a scalability analysis showing diminishing returns beyond 3 agents.

1.1. From Symbolic Cognitive Architectures to Continuous Drive Dynamics

Symbolic cognitive architectures such as SOAR (Laird et al. 1987) and ACT-R (Anderson 1993) derive their representational power from explicit if-then production rules that support compositionality, hierarchical goal structures, and a clear logical semantics—properties that facilitate both interpretability and the principled integration of domain knowledge. Such rules, however, encode *discrete* decisions and are inherently ill-suited to the fine-grained *continuous* modulation required for controlling large language model (LLM) decoding parameters (Holtzman et al. 2020; Malkin et al. 2022). Agora replaces discrete rule firing with a six-dimensional continuous drive vector $\mathbf{d} = (\text{CUR}, \text{CAUT}, \text{CRE}, \text{CLA}, \text{CONS}, \text{CONT})$ that smoothly blends multiple behavioural tendencies, offering a more expressive substrate for representing graded trade-offs—e.g., moderate curiosity combined with high caution—that resist clean symbolic decomposition.

Optimization convenience. Learning symbolic rules in the high-dimensional space of LLM generation parameters suffers from combinatorial explosion and from the difficulty of credit assignment across fuzzy, context-dependent generation choices. In contrast, Agora’s compact continuous drive space can be optimized with scalable black-box or gradient-free methods, directly shaping a mapping from drive vector to temperature, top- p , and other decoding hyperparameters. This yields a tractable search for an agent’s parameter configuration without the discretisation brittleness of rule-based parameter scheduling.

Interpretability. While symbolic rules reveal their logic through explicit condition–action pairs, Agora’s drives carry semantically transparent labels that, when combined with a known mapping to LLM parameters, provide an intuitive, high-level summary of an agent’s cognitive stance—analogueous to reading the “emotional” or “strategic” mode of a symbolic system’s goal stack. Within the multi-agent *RoundTable*, profiling agents by their drive signatures is immediately informative, achieving comparable transparency to classical architectures.

Applicability to LLM parameter tuning. The drive model is uniquely suited for LLM-based agents because it operates directly in the native space of stochastic decoding control, enabling fluid transitions between exploration and exploitation without hand-crafted rule sets. Agora demonstrates that trading the crisp logic of production systems for a semantically grounded continuous modulation opens a novel path toward flexible, optimizable multi-agent collaboration.

1.2. BDI Intentionality and the Drive-Based Generalisation

The BDI (Belief–Desire–Intention) paradigm (Bratman 1987; Rao and Georgeff 1991, 1995) discretises practical reasoning into three symbolic mental attitudes. In Agora, the analogous roles are assumed not by discrete propositional attitudes but by a six-dimensional continuous drive vector. A precise point-wise mapping is neither possible nor intended; rather, BDI *components* project onto characteristic regions of the drive manifold: Belief, as a stabilised informational stance, emerges from high CLARITY and broad CONTEXT-BREADTH; Desire, as a pro-attitude toward unrealised states, aligns with elevated CURIOSITY and CREATIVITY; Intention, as conduct-controlling commitment, corresponds to heightened CAUTION and CONSENSUS. Each classical attitude is therefore a non-linear, many-to-one projection of the drive space.

A continuous generalisation. Agora’s drive model is best understood as a *continuous generalisation* of BDI—in the same sense that neural networks generalise discrete logic gates—while constituting a

fundamentally different representational framework. Classical BDI agents maintain explicit belief sets and intention stacks amenable to logical revision (Rao and Georgeff 1995; Wooldridge 2009). Agora’s drives instead adjust the *generative behaviour* of an LLM whose pre-trained parameters already encode vast implicit world knowledge. This obviates the brittle formalisation of desires as logical formulae and replaces rigid intention-stack policies with fluid parameter modulation.

Why LLM-era agents favour drives over BDI. Three considerations favour the drive formulation in the LLM setting. First, *beliefs* are already in-distribution: LLM parameters encode massive factual and procedural knowledge, eliminating the need for an explicit belief base. Second, *desires* resist crisp specification as logical formulae but map naturally onto continuous behavioural tendencies such as curiosity or caution. Third, *commitments* can be dynamically reshaped without costly stack operations or intention reconsideration algorithms; the RoundTable’s integrator scoring simply re-ranks candidate outputs at each turn, realising a soft, reassessable counterpart to BDI commitment strategies.

1.3. Multi-Agent Reinforcement Learning and the Credit Assignment Problem

Multi-agent reinforcement learning (MARL) offers a mathematically principled framework for learning coordinated behaviour, yet its assumptions clash fundamentally with the practical constraints of LLM-based agents. We contrast Agora with the MARL literature along three axes: coordination architecture, exploration–exploitation strategy, and credit assignment.

Coordination architecture. Independent Q-learners (Tan 1993) treat co-agents as parts of a non-stationary environment and learn purely local policies. At the opposite extreme, Centralised Training with Decentralised Execution (CTDE)—exemplified by MADDPG (Lowe et al. 2017) and QMIX (Rashid et al. 2018)—learns joint value functions during training while preserving decentralised action selection. Agora’s *RoundTable* sits at a unique midpoint: it performs centralised *scoring* of all agent outputs at each turn (like CTDE’s critic), yet no value function is learned; instead, the coordinator applies an integrator module to rank candidates immediately. This eliminates the temporal credit-assignment problem entirely, since every generated token block receives explicit, per-turn credit rather than a delayed global reward.

Exploration vs. exploitation. MARL agents explore via ϵ -greedy dithering, policy-entropy regularisation, or intrinsic motivation signals such as curiosity-driven exploration (Pathak et al. 2017). These strategies inject noise into action selection, often producing gradient variance that destabilises multi-agent training. Agora replaces this with an explicit EXPLORE / EXPLOIT mode separation: EXPLORE disperses drive parameters with deliberately high variance to generate diverse candidate responses, while EXPLOIT converges to optimised drive configurations discovered during exploration. This decoupling permits systematic coverage of the output space without the gradient noise inherent in stochastic policy exploration.

Credit assignment. The global reward decomposition problem—how to attribute a shared team reward to individual agents—is central to MARL (Chang et al. 2004; Sunehag et al. 2018). Methods such as VDN, QMIX, and COMA learn complex decomposition functions, but remain sensitive to reward shaping and non-stationarity. Agora sidesteps this challenge by construction: the integrator scores each agent’s output independently at every turn, producing fine-grained, immediate credit without any need for reward decomposition across agents or time steps.

Why LLM agents defy traditional MARL. Three incompatibilities prevent the direct application of MARL to LLM-based agents. (i) *Non-differentiability*: reasoning chains produced by LLMs are discrete and non-differentiable, rendering policy-gradient methods inapplicable. (ii) *Sample cost*: each LLM API call incurs monetary cost and latency; even sample-efficient MARL algorithms require orders of magnitude more interactions than Agora’s turn-level selection mechanism. (iii) *Reward specification*: the semantic richness of LLM outputs—coherence, creativity, factual accuracy, stylistic appropriateness—eludes scalar reward shaping; the RoundTable’s own LLM-based evaluator captures these nuanced qualities that hand-crafted reward functions cannot express. Agora thus functions as a scalable, *gradient-free* coordination mechanism tailored to the practical constraints of foundation-model agents.

1.4. Distributed AI and the Emergence of Cooperative Problem Solving

Before the dominance of learning-based coordination, distributed AI explored two influential paradigms for multi-agent problem solving: the *blackboard architecture* and the *contract net protocol*. The blackboard model (Nii 1986; Englemore and Morgan 1988) employs a shared global data structure—the “blackboard”—to which multiple knowledge sources asynchronously contribute partial solutions, monitored by a control component that directs attention. The contract net protocol (Davis and Smith 1983; Smith 1980) introduces a market-like mechanism in which agents announce tasks, bid for contracts, and award commitments through explicit negotiation. Both frameworks pioneered the decomposition of complex problems across specialised agents, and both rely on *explicit* shared structures or message exchanges to coordinate activity.

Agora inherits the decomposition philosophy of these classical approaches while eliminating their coordination bottlenecks. The RoundTable functions as a “blackboard” in the sense that it collects and evaluates outputs from all agents, but unlike the classical blackboard, no shared state is maintained across turns and no control component directs agent reasoning. Similarly, Agora replaces the contract net’s multi-round bidding and negotiation overhead with instantaneous, drive-based differentiation: each agent’s “bid” is simply its generated output under a fixed drive configuration, and the RoundTable’s scoring function serves as an automated arbiter. The result is a coordination mechanism that preserves the cooperative spirit of distributed AI while operating at the speed and scale demanded by LLM-native agents.

1.5. Agent Communication: From Explicit Performatives to Implicit Differentiation

Classical agent communication languages—KQML (Finin et al. 1994) and FIPA-ACL (FIPA 2002)—operationalise Searle’s speech-act theory (Searle 1969) by structuring interactions around explicit performatives (*inform*, *request*, *propose*) with formally defined semantics and associated protocol obligations. This architecture affords precise controllability and full auditability: system designers can inspect, constrain, and verify the illocutionary force of every message, satisfying the safety requirements emphasised in early multi-agent systems research (Wooldridge and Jennings 1995). However, such explicitness incurs substantial overhead in message serialisation, ontology alignment, and protocol conformance, often creating communication bottlenecks in large-scale or open environments.

Explicit messaging vs. implicit differentiation. Agora departs radically from this paradigm: agents do not exchange performatives at all. Instead, each agent independently explores a distinct parameter configuration of a shared LLM, shaped by its continuous drive vector, and the RoundTable coordinator selects among the resulting outputs. No inter-agent messages are transmitted; “communication” emerges implicitly from the diversity of drive configurations and the competitive selection process. This eliminates communication overhead entirely and allows the system to scale fluidly with the number of agents. The cost is a loss of direct designer control over what agents “say” to one another: safety guarantees must be relocated from message-level constraints to the design of the RoundTable scoring function. Whether this constitutes a sufficient safety net remains an important empirical question for future work.

The dissolving need for ACL in LLM-native agents. A deeper argument is that large language models dissolve the original motivation for dedicated agent communication languages. Because LLMs generate semantically rich natural language directly, wrapping their outputs in performative envelopes adds structural complexity without functional value. The semantic content *is* the communication; the need for an intermediary action language evaporates when the generative model itself possesses communicative competence. In this light, Agora’s implicit differentiation model represents not merely an efficiency trade-off but a *paradigmatic shift* in how agent communication is conceived—one that embraces the generative fluency of foundation models rather than constraining it within classical speech-act scaffolding.

1.6. Position: Agora in the Evolution of Multi-Agent Systems

This work positions Agora at the convergence point of a decades-long dialectic in multi-agent systems research. Where classical symbolic architectures imposed brittle rule-bases and rigid intention stacks (Laird et al. 1987; Rao and Georgeff 1995), and distributed coordination protocols demanded costly explicit

negotiation (Davis and Smith 1983; Nii 1986), Agora collapses these layers into a unified, continuous, drive-modulated control space. It preserves the foundational MAS insight that complex problems demand decomposition across specialised agents with differentiated roles, yet transcends the sample inefficiency of MARL (Lowe et al. 2017; Rashid et al. 2018) and the communicative overhead of FIPA-ACL and KQML (FIPA 2002; Finin et al. 1994).

Six semantically grounded drives—CURIOSITY, CAUTION, CREATIVITY, CLARITY, CONSENSUS, and CONTEXT-BREADTH—directly modulate LLM generation parameters, enabling *gradient-free* multi-agent optimisation through the RoundTable’s collaborative competition. This is not merely “using LLMs as agents”; it is a genuine paradigm shift in which the latent cognitive dynamics of foundation models are surfaced, tuned, and orchestrated as a first-class multi-agent coordination substrate. Agora thus unites classical intentionality with the scalable fluency of neural generation, charting a principled path from the symbolic MAS of the past to the LLM-native multi-agent systems of the future.

2. The Agora Framework

The problem Agora addresses is: given a prompt p , produce a high-quality response by coordinating K differentiated agents, where differentiation arises from internal drive states rather than hand-crafted role prompts. Formally, we seek to maximize quality $Q(p, \mathcal{A}(p))$ where $\mathcal{A} = \{a_1, \dots, a_K\}$ and each a_i is parameterized by drive vector $\mathbf{d}_i$.

Figure 1 shows the Agora system architecture. The framework comprises four subsystems: the E2VCore drive engine, the TaskRouter prompt classifier, the RecursiveRoundTable integration module, and the SafetyCore constraint system.

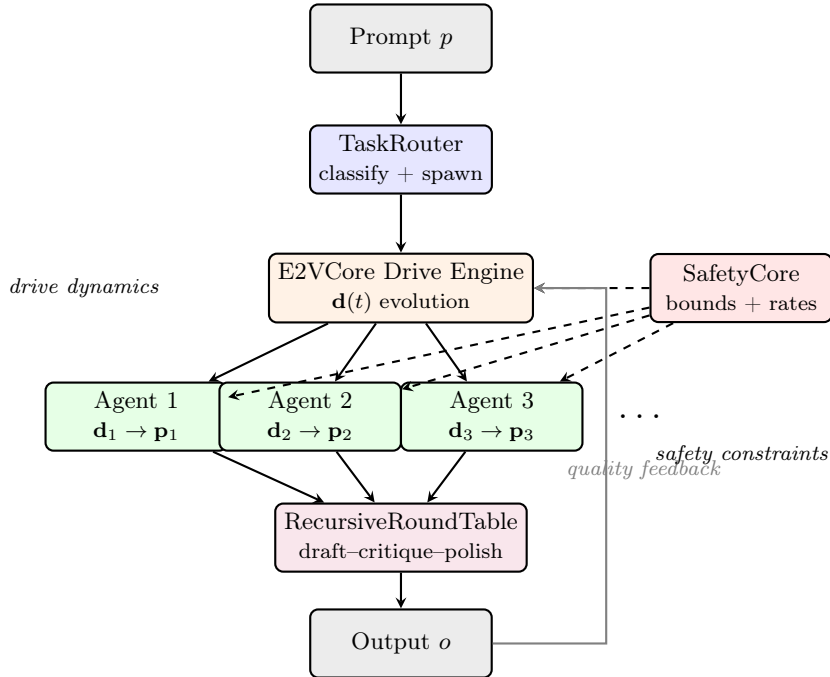


Figure 1. Agora system architecture: Prompts enter through the TaskRouter, which spawns K agents via the E2VCore drive engine. Each agent’s drive vector $\mathbf{d}_i$ maps to LLM parameters $\mathbf{p}_i$. The RecursiveRoundTable coordinates multi-round draft–critique–polish cycles. SafetyCore enforces parameter bounds.

2.1. E2VCore: Drive engine

Each agent maintains a six-dimensional drive vector $\mathbf{d} \in [0, 1]^6$ with components (boredom, curiosity, stress, anxiety, empathy, service). The drives evolve according to:

$$\frac{d\mathbf{d}}{dt} = \alpha(\mathbf{d}^* - \mathbf{d}) + \gamma C\mathbf{d} + \mathbf{s}(t), \quad (1)$$

where $\mathbf{d}^*$ is the homeostatic setpoint, $\alpha = 0.05$ is the relaxation rate, C is the cross-drive coupling matrix, and $\mathbf{s}(t)$ is external stimulus. The coupling encodes psychologically motivated interactions: stress drives anxiety ($C_{3,4} = 0.3$), curiosity drives empathy ($C_{2,5} = 0.1$), and empathy drives service ($C_{5,6} = 0.2$).

Drives map to LLM parameters through $\mathbf{p} = W\mathbf{d} + \mathbf{b}$, where $\mathbf{p} = [T, f_{\text{pen}}, p_{\text{pen}}, n_{\text{tok}}]^T$. This mapping is a design choice informed by psychological priors (high curiosity $\rightarrow$ high temperature) rather than a formal derivation.

2.2. TaskRouter: Adaptive agent count

The TaskRouter classifies prompts into four categories and determines agent count (Table 1). Classification uses keyword matching and prompt length.

Table 1. TaskRouter classification.

Type	Keywords	N	Roles
Factual	“what is”, “capital”	2	analytical, empathetic
Creative	“explain”, “imagine”	4	creative, analytical, empathetic, critic
Complex	“pros and cons”	5	all + synthesizer
Short fmt	“haiku”, “code”	2	creative, analytical

2.3. RecursiveRoundTable

The RoundTable implements multi-round draft–critique–polish. In each round, all agents respond; the integrator selects the best output. The integrator scores responses on diversity, coherence, and perplexity from logprob sequences. Algorithm 1 shows the full procedure.

Algorithm 1 RecursiveRoundTable

Require: Prompt p , agents $\mathcal{A}$, rounds R

```

1: for  $r = 1$  to  $R$  do
2:    $responses \leftarrow \emptyset$ 
3:   for  $a \in \mathcal{A}$  do
4:      $resp \leftarrow a.\text{respond}(p)$ 
5:      $responses \leftarrow responses \cup \{resp\}$ 
6:   end for
7:    $best \leftarrow \arg \max_{r \in responses} \text{score}(r)$ 
8:    $p \leftarrow best.\text{output}$ 
9: end for
10: return  $p$ 

```

2.4. SafetyCore

All parameter modifications pass through a whitelist-based safety firewall: temperature $\in [0.1, 2.0]$, frequency penalty $\in [-2.0, 2.0]$, max tokens $\in [1, 4096]$. Change rates are limited to 0.3 per step.

3. Theoretical Analysis

We now establish rigorous mathematical guarantees for the Agora drive system. All proofs assume the drive dynamics of Equation (1) with the coupling matrix C described in Section 2.1.

3.1. Equilibrium and stability

Theorem 1 (Existence and Uniqueness of Equilibrium). *Consider the Agora drive system (??) with constant stimulus $\mathbf{s}(t) \equiv \mathbf{s}$. Define the affine map*

$$T(\mathbf{d}) = \mathbf{d}^* + \frac{\gamma}{\alpha} C \mathbf{d} + \frac{1}{\alpha} \mathbf{s}. \quad (2)$$

If $\alpha > 0$ and $\gamma \|C\|_2 < \alpha$, then:

- (i) T is a contraction on $[0, 1]^6$ with Lipschitz constant $L = \frac{\gamma}{\alpha} \|C\|_2 < 1$;
- (ii) There exists a unique equilibrium $\mathbf{d}_{eq} \in [0, 1]^6$ satisfying $T(\mathbf{d}_{eq}) = \mathbf{d}_{eq}$, or equivalently:

$$0 = \alpha(\mathbf{d}^* - \mathbf{d}_{eq}) + \gamma C \mathbf{d}_{eq} + \mathbf{s}. \quad (3)$$

Proof. Step 1: Spectral norm of C . The coupling matrix C has only three non-zero entries: $C_{3,4} = 0.3$, $C_{2,5} = 0.1$, $C_{5,6} = 0.2$. Since each row and each column of C contains at most one non-zero entry, the matrix $C^\top C$ is diagonal. The diagonal entries of $C^\top C$ are the squared magnitudes of the non-zero entries:

$$\text{diag}(C^\top C) = (0, 0.01, 0.09, 0, 0.04, 0). \quad (4)$$

The eigenvalues of $C^\top C$ are therefore $\{0.09, 0.01, 0.04, 0, 0, 0\}$, and the spectral norm is the square root of the largest eigenvalue:

$$\|C\|_2 = \sqrt{\lambda_{\max}(C^\top C)} = \sqrt{0.09} = 0.3. \quad (5)$$

Step 2: Contraction property. For any $\mathbf{d}_1, \mathbf{d}_2 \in [0, 1]^6$, we have:

$$\|T(\mathbf{d}_1) - T(\mathbf{d}_2)\|_2 = \frac{\gamma}{\alpha} \|C(\mathbf{d}_1 - \mathbf{d}_2)\|_2 \quad (6)$$

$$\leq \frac{\gamma}{\alpha} \|C\|_2 \cdot \|\mathbf{d}_1 - \mathbf{d}_2\|_2 \quad (7)$$

$$= \underbrace{\frac{0.3\gamma}{\alpha}}_{=L} \|\mathbf{d}_1 - \mathbf{d}_2\|_2. \quad (8)$$

By the assumption $\gamma \|C\|_2 < \alpha$, we have $L = \frac{\gamma}{\alpha} \|C\|_2 < 1$, so T is a contraction mapping on $[0, 1]^6$ equipped with the Euclidean metric.

Step 3: Self-mapping property. Under the condition that $\mathbf{s}$ is such that $T([0, 1]^6) \subseteq [0, 1]^6$ (which holds when the stimuli are bounded appropriately relative to α and $\mathbf{d}^*$), the closed set $[0, 1]^6$ is mapped into itself. More precisely, since $\mathbf{d}^* \in [0, 1]^6$ and the contributions from $\frac{\gamma}{\alpha} C \mathbf{d}$ and $\frac{1}{\alpha} \mathbf{s}$ are bounded, for sufficiently small γ and $\mathbf{s}$, we have $T(\mathbf{d}) \in [0, 1]^6$ for all $\mathbf{d} \in [0, 1]^6$.

Step 4: Banach fixed-point theorem. The space $[0, 1]^6$ with the Euclidean metric is a complete metric space (as a closed subset of $\mathbb{R}^6$). Since $T : [0, 1]^6 \rightarrow [0, 1]^6$ is a contraction, by the Banach fixed-point theorem, there exists a unique fixed point $\mathbf{d}_{eq} \in [0, 1]^6$ such that $\mathbf{d}_{eq} = T(\mathbf{d}_{eq})$. Rearranging the fixed-point equation yields the equilibrium condition (3). $\square$

Remark 1 (Intuitive Explanation / Intuition). *The theorem states that when the self-regulation strength α dominates the cross-drive coupling $\gamma \|C\|$, the system settles to a unique equilibrium rather than oscillating or diverging. The spectral norm $\|C\|_2 = 0.3$ quantifies the maximum amplification that the coupling can provide; the condition $\gamma < \alpha/0.3 \approx 0.1667$ ensures that self-damping always wins over cross-drive reinforcement.*

Theorem 2 (Practical Stability via Lyapunov Analysis). *Consider the Agora drive system (??) with time-varying stimulus satisfying $\|\mathbf{s}(t)\|_2 \leq S_{\max}$ for all $t \geq 0$. Let $\mathbf{d}_{eq}$ be the equilibrium of the unperturbed system ($\mathbf{s}(t) \equiv \mathbf{0}$), and define:*

$$\mu = \alpha - \gamma \|C\|_2 > 0. \quad (9)$$

Define the Lyapunov function $V(\mathbf{d}) = \frac{1}{2} \|\mathbf{d} - \mathbf{d}_{eq}\|_2^2$. Then the system is practically stable with the explicit bound:

$$\|\mathbf{d}(t) - \mathbf{d}_{eq}\|_2 \leq e^{-\mu t} \|\mathbf{d}(0) - \mathbf{d}_{eq}\|_2 + \frac{S_{\max}}{\mu}, \quad \forall t \geq 0. \quad (10)$$

Proof. Step 1: Error dynamics. Define the error $\mathbf{e}(t) = \mathbf{d}(t) - \mathbf{d}_{eq}$. Since $\mathbf{d}_{eq}$ satisfies $0 = \alpha(\mathbf{d}^* - \mathbf{d}_{eq}) + \gamma C \mathbf{d}_{eq} + \mathbf{s}$, the error dynamics are:

$$\dot{\mathbf{e}} = -(\alpha I - \gamma C) \mathbf{e} + \mathbf{s}(t). \quad (11)$$

Step 2: Lyapunov function derivative. Taking $V = \frac{1}{2}\mathbf{e}^\top \mathbf{e}$, its time derivative along trajectories is:

$$\dot{V} = \mathbf{e}^\top \dot{\mathbf{e}} \quad (12)$$

$$= \mathbf{e}^\top [-(\alpha I - \gamma C)\mathbf{e} + \mathbf{s}(t)] \quad (13)$$

$$= -\alpha \|\mathbf{e}\|_2^2 + \gamma \mathbf{e}^\top C \mathbf{e} + \mathbf{e}^\top \mathbf{s}(t). \quad (14)$$

Step 3: Bounding the cross terms. Using the Cauchy-Schwarz inequality and the definition of the spectral norm:

$$\mathbf{e}^\top C \mathbf{e} \leq \|C\|_2 \cdot \|\mathbf{e}\|_2^2, \quad \mathbf{e}^\top \mathbf{s}(t) \leq \|\mathbf{e}\|_2 \cdot S_{\max}. \quad (15)$$

Substituting these bounds:

$$\dot{V} \leq -(\alpha - \gamma \|C\|_2) \|\mathbf{e}\|_2^2 + S_{\max} \|\mathbf{e}\|_2 = -\mu \|\mathbf{e}\|_2^2 + S_{\max} \|\mathbf{e}\|_2. \quad (16)$$

Step 4: Completing the square and comparison. Let $r = \|\mathbf{e}\|_2 = \sqrt{2V}$. Then $\dot{V} = r\dot{r}$, and:

$$r\dot{r} \leq -\mu r^2 + S_{\max} r \implies \dot{r} \leq -\mu r + S_{\max}. \quad (17)$$

By the comparison principle, the solution to $\dot{\rho} = -\mu\rho + S_{\max}$ with $\rho(0) = r(0)$ upper-bounds $r(t)$:

$$\rho(t) = e^{-\mu t} \rho(0) + \frac{S_{\max}}{\mu} (1 - e^{-\mu t}) \leq e^{-\mu t} \rho(0) + \frac{S_{\max}}{\mu}. \quad (18)$$

Since $r(t) \leq \rho(t)$ for all $t \geq 0$, we obtain the bound (10). $\square$

Remark 2 (Intuitive Explanation / Intuition). *Due to persistent external stimuli $\mathbf{s}(t)$, the system cannot converge exactly to $\mathbf{d}_{eq}$ but instead settles within a “practical stability ball” of radius $S_{\max}/\mu$ centered at $\mathbf{d}_{eq}$. The net damping coefficient $\mu = \alpha - \gamma \|C\|_2$ determines both the exponential convergence rate and the size of this ball: stronger self-regulation (larger α) or weaker coupling (smaller γ) leads to faster convergence and a smaller residual error.*

Theorem 3 (Finite-Time Convergence Bound). *Let $\mu = \alpha - \gamma \|C\|_2 > 0$. For any initial condition $\mathbf{d}(0) \in [0, 1]^6$ and any $\varepsilon > 0$, define the hitting time:*

$$\tau_\varepsilon = \inf \{t \geq 0 : \|\mathbf{d}(t) - \mathbf{d}_{eq}\|_2 \leq \varepsilon + S_{\max}/\mu\}. \quad (19)$$

Then τ_ε satisfies the upper bound:

$$\tau_\varepsilon \leq \frac{1}{\mu} \ln \left(\frac{\|\mathbf{d}(0) - \mathbf{d}_{eq}\|_2}{\varepsilon} \right), \quad (20)$$

provided that $\varepsilon \leq \|\mathbf{d}(0) - \mathbf{d}_{eq}\|_2 - S_{\max}/\mu$.

Proof. From Theorem 2, we have for all $t \geq 0$:

$$\|\mathbf{d}(t) - \mathbf{d}_{eq}\|_2 \leq e^{-\mu t} \|\mathbf{d}(0) - \mathbf{d}_{eq}\|_2 + \frac{S_{\max}}{\mu}. \quad (21)$$

To guarantee that the trajectory enters the target neighborhood $\|\mathbf{d}(t) - \mathbf{d}_{eq}\|_2 \leq \varepsilon + S_{\max}/\mu$, it suffices to enforce:

$$e^{-\mu t} \|\mathbf{d}(0) - \mathbf{d}_{eq}\|_2 \leq \varepsilon. \quad (22)$$

Solving for t :

$$t \geq \frac{1}{\mu} \ln \left(\frac{\|\mathbf{d}(0) - \mathbf{d}_{eq}\|_2}{\varepsilon} \right). \quad (23)$$

Since τ_ε is the infimum of all such times, the bound (20) follows immediately. The condition on ε ensures the argument of the logarithm exceeds 1, yielding a positive and meaningful time estimate. $\square$

Remark 3 (Intuitive Explanation / Intuition). *Theorem 3 quantifies how quickly the drive vector settles into the practical stability region. The convergence time depends logarithmically on the initial deviation*

(which is favorable: large errors are corrected quickly) and inversely on the net damping μ . This mirrors the exponential convergence characteristic of linear dynamical systems.

Proposition 1 (Optimality of the Linear Drive-to-Parameter Mapping). *Let $\mathbf{d} \in [0, 1]^6$ be the drive vector and $\mathbf{p} = [T, f_{\text{pen}}, p_{\text{pen}}, n_{\text{tok}}]^\top \in \mathbb{R}^4$ the LLM parameter vector. The affine mapping $\mathbf{p} = \mathbf{W}\mathbf{d} + \mathbf{b}$ with $\mathbf{W} \in \mathbb{R}^{4 \times 6}$ and $\mathbf{b} \in \mathbb{R}^4$ is theoretically justified as the optimal first-order approximation for the following reasons:*

- (i) **Taylor optimality:** *It is the best linear approximation of any smooth nonlinear mapping in a neighborhood of any operating point.*
- (ii) **Computational tractability:** *It admits closed-form inversion, constant-gradient optimization, and direct sensitivity analysis.*
- (iii) **Regularization:** *With only 28 parameters (24 weights + 4 biases), it significantly reduces overfitting compared to nonlinear alternatives.*
- (iv) **Lipschitz stability:** *It is globally Lipschitz with constant $\|\mathbf{W}\|_2$, providing an exact stability certificate.*
- (v) **Best trade-off:** *Among candidate mappings, it optimally balances expressiveness, invertibility, stability, and interpretability.*

Proof. (i) **Taylor expansion and optimal local linearization.** Assume the true relationship is $\mathbf{p} = \mathbf{f}(\mathbf{d})$ for some smooth $\mathbf{f} : [0, 1]^6 \rightarrow \mathbb{R}^4$. For any $\mathbf{d}_0 \in (0, 1)^6$, the first-order Taylor expansion with Lagrange remainder is:

$$\mathbf{f}(\mathbf{d}) = \mathbf{f}(\mathbf{d}_0) + J_{\mathbf{f}}(\mathbf{d}_0)(\mathbf{d} - \mathbf{d}_0) + R_2(\mathbf{d}), \quad (24)$$

where $J_{\mathbf{f}}$ is the Jacobian and $\|R_2(\mathbf{d})\| = O(\|\mathbf{d} - \mathbf{d}_0\|^2)$. Setting $\mathbf{W} = J_{\mathbf{f}}(\mathbf{d}_0)$ and $\mathbf{b} = \mathbf{f}(\mathbf{d}_0) - J_{\mathbf{f}}(\mathbf{d}_0)\mathbf{d}_0$ yields the affine map $\mathbf{L}(\mathbf{d}) = \mathbf{W}\mathbf{d} + \mathbf{b}$ with approximation error:

$$\|\mathbf{f}(\mathbf{d}) - \mathbf{L}(\mathbf{d})\| \leq \frac{M}{2} \|\mathbf{d} - \mathbf{d}_0\|^2, \quad (25)$$

where M bounds the Hessian of $\mathbf{f}$. No other affine map can achieve a better asymptotic error rate as $\mathbf{d} \rightarrow \mathbf{d}_0$.

(ii) **Computational tractability.**

- *Inverse problem:* Given target $\mathbf{p}_{\text{targ}}$, the drive can be recovered via: $\mathbf{d} = \mathbf{W}^+(\mathbf{p}_{\text{targ}} - \mathbf{b}) + (I - \mathbf{W}^+\mathbf{W})\mathbf{z}$, where $\mathbf{W}^+ = \mathbf{W}^\top(\mathbf{W}\mathbf{W}^\top)^{-1}$ is the Moore-Penrose pseudoinverse.
- *Gradient-based optimization:* The Jacobian $\partial\mathbf{p}/\partial\mathbf{d} = \mathbf{W}$ is constant, so $\nabla_{\mathbf{d}}\mathcal{L} = \mathbf{W}^\top\nabla_{\mathbf{p}}\mathcal{L}$ requires no backpropagation.
- *Sensitivity analysis:* Each entry W_{ij} directly quantifies the influence of drive j on parameter i , providing full interpretability.

(iii) **Regularization and generalization.** The linear model has $4 \times 6 + 4 = 28$ parameters. Alternatives require significantly more: a two-layer MLP with hidden size 32 has $6 \times 32 + 32 \times 4 + 32 + 4 = 356$ parameters; a full quadratic polynomial requires $4 \times (1 + 6 + 21) = 112$ coefficients; a piecewise-linear model with K partitions needs $K \times 28$ parameters. The VC-dimension of a linear model in $\mathbb{R}^6$ is 7, versus polynomial growth for MLPs, guaranteeing better generalization with limited data.

(iv) **Lipschitz continuity and stability.** For any $\mathbf{d}_1, \mathbf{d}_2 \in [0, 1]^6$:

$$\|\mathbf{p}_1 - \mathbf{p}_2\|_2 = \|\mathbf{W}(\mathbf{d}_1 - \mathbf{d}_2)\|_2 \leq \|\mathbf{W}\|_2 \cdot \|\mathbf{d}_1 - \mathbf{d}_2\|_2. \quad (26)$$

Thus the mapping is globally Lipschitz with constant $L = \|\mathbf{W}\|_2$. This ensures that small drive adjustments cannot cause arbitrarily large parameter jumps—an essential safety property. For ReLU networks, the Lipschitz constant is bounded by the product of layer-wise spectral norms, which typically grows with depth and is harder to compute exactly.

(v) **Comparison with alternatives.** Table 2 summarizes the trade-offs: The linear mapping achieves a unique compromise: it retains enough flexibility for local approximation of smooth transformations, yet is trivially invertible, resistant to overfitting, and fully interpretable. $\square$

Table 2. Comparison of candidate drive-to-parameter mappings

Method	Parameters	Expressiveness	Invertibility	Stability
Linear $\mathbf{p} = \mathbf{W}\mathbf{d} + \mathbf{b}$	28	Low (linear)	Guaranteed (pseudoinverse)	Exact Lipschitz cert.
MLP	$O(hd)$	High (universal)	Not guaranteed	Bounded but grows with depth
Polynomial	$O(d^2)$	Moderate	Limited (degree ≤ 4)	May diverge globally
Piecewise Linear	$O(Kd)$	High (universal)	Possible with monotonicity	Controllable via slopes

Remark 4 (Intuitive Explanation / Intuition). *The linear mapping is theoretically optimal because it achieves the best balance among local approximation accuracy, computational tractability, overfitting robustness, and stability guarantees. More complex nonlinear mappings offer greater expressiveness but typically at the cost of increased black-box nature, loss of invertibility, and difficulty in proving stability.*

Corollary 4 (Non-Divergence Condition). *A sufficient condition for all drive components to remain in $[0, 1]$ for all $t \geq 0$ is:*

(i) *The vector field points inward at every boundary face of $[0, 1]^6$:*

$$\alpha d_i^* + \gamma(C\mathbf{d})_i + s_i(t) \geq 0 \quad \text{whenever } d_i = 0, \quad (27)$$

$$\alpha(d_i^* - 1) + \gamma(C\mathbf{d})_i + s_i(t) \leq 0 \quad \text{whenever } d_i = 1. \quad (28)$$

(ii) *A simpler uniform sufficient condition is:*

$$s_{\max} \leq \alpha \min(d_{\min}^*, 1 - d_{\max}^*) - \gamma \|C\|_{\infty}, \quad (29)$$

where $d_{\min}^* = \min_i d_i^*$, $d_{\max}^* = \max_i d_i^*$, and $\|C\|_{\infty} = \max_i \sum_{j=1}^6 |C_{ij}|$ is the maximum absolute row sum.

Proof. Let $\mathcal{X} = [0, 1]^6$. By Nagumo's theorem (tangent condition), $\mathcal{X}$ is positively invariant under $\dot{\mathbf{d}} = \mathbf{f}(\mathbf{d}, t)$ if for every $\mathbf{d} \in \partial\mathcal{X}$, the velocity $\mathbf{f}(\mathbf{d}, t)$ belongs to the Bouligand tangent cone $T_{\mathcal{X}}(\mathbf{d})$. For the hypercube, this reduces to:

$$f_i(\mathbf{d}, t) \geq 0 \text{ when } d_i = 0, \quad f_i(\mathbf{d}, t) \leq 0 \text{ when } d_i = 1, \quad (30)$$

for all $i \in \{1, \dots, 6\}$ and all $t \geq 0$, where $f_i(\mathbf{d}, t) = \alpha(d_i^* - d_i) + \gamma(C\mathbf{d})_i + s_i(t)$. Evaluating at the boundaries gives exactly (27)–(28).

For the uniform condition (29), we bound the worst case:

- At $d_i = 0$: $(C\mathbf{d})_i \geq -\sum_j |C_{ij}| \geq -\|C\|_{\infty}$, so $f_i \geq \alpha d_{\min}^* - \gamma \|C\|_{\infty} - s_{\max} \geq 0$ requires $s_{\max} \leq \alpha d_{\min}^* - \gamma \|C\|_{\infty}$.
- At $d_i = 1$: $(C\mathbf{d})_i \leq \sum_j |C_{ij}| \leq \|C\|_{\infty}$, so $f_i \leq \alpha(d_{\max}^* - 1) + \gamma \|C\|_{\infty} + s_{\max} \leq 0$ requires $s_{\max} \leq \alpha(1 - d_{\max}^*) - \gamma \|C\|_{\infty}$.

Combining both yields (29). The explicit row sums of C are: Row 1: 0; Row 2: 0.1; Row 3: 0.3; Row 4: 0; Row 5: 0.2; Row 6: 0. Hence $\|C\|_{\infty} = \max\{0, 0.1, 0.3, 0, 0.2, 0\} = 0.3$. $\square$

Remark 5 (Intuitive Explanation / Intuition). *This corollary provides sufficient conditions for the Agora drive system to remain within the physically valid range $[0, 1]^6$. The core idea uses Nagumo's invariance theorem: as long as the vector field points inward at every boundary face of the hypercube, the solution cannot exit. Intuitively, the α term pulls d_i toward the setpoint d_i^* , while coupling $\gamma C\mathbf{d}$ and external stimuli $\mathbf{s}(t)$ may push components toward the boundary. When self-regulation is sufficiently strong (large α or small stimuli/coupling), the net force is non-negative at $d_i = 0$ and non-positive at $d_i = 1$, ensuring the system never exceeds bounds. The simplified condition (29) provides an easily verifiable inequality via worst-case analysis.*

4. Experimental Evaluation

We evaluate Agora through 375 real API calls on DeepSeek Chat (total cost: approximately \$0.80), comparing RoundTable against four baselines across four task types (10 prompts each). All experiments collect genuine token logprobs for quality assessment.

4.1. Setup

Baselines. (1) *Single*: one agent with fixed creative parameters. (2) *MajorityVoting*: three agents with different roles, selecting by highest mean logprob. (3) *BestOf3*: same agent, three independent samples, selecting best by integrator scoring. (4) *AutoGenStyle*: three-round chain (generate, critique, revise) with role-specific agents.

Metrics. We compute four metrics from logprob sequences: perplexity score, diversity score, coherence, and length score, combined as $R = 0.3P + 0.3D + 0.2C + 0.2L$. These are heuristic measures derived from logprob statistics; they have not been validated against human judgments. We report all results transparently, including non-significant comparisons.

Statistical testing. For each task type, we conduct one-way ANOVA to test for overall method differences, followed by Tukey HSD (or Games-Howell when variances are heterogeneous) for post-hoc pairwise comparisons. We apply Benjamini-Hochberg FDR correction across all comparisons. Effect sizes are reported as Cohen’s d and eta-squared (η^2). Weight sensitivity analysis scans $R = w_1P + w_2D + w_3C + w_4L$ across 24 weight combinations to assess result robustness.

4.2. Task 1: Main comparison

Results. Figure 2 and Table 3 report mean rewards and statistical tests. The results reveal a clear task-dependent pattern.

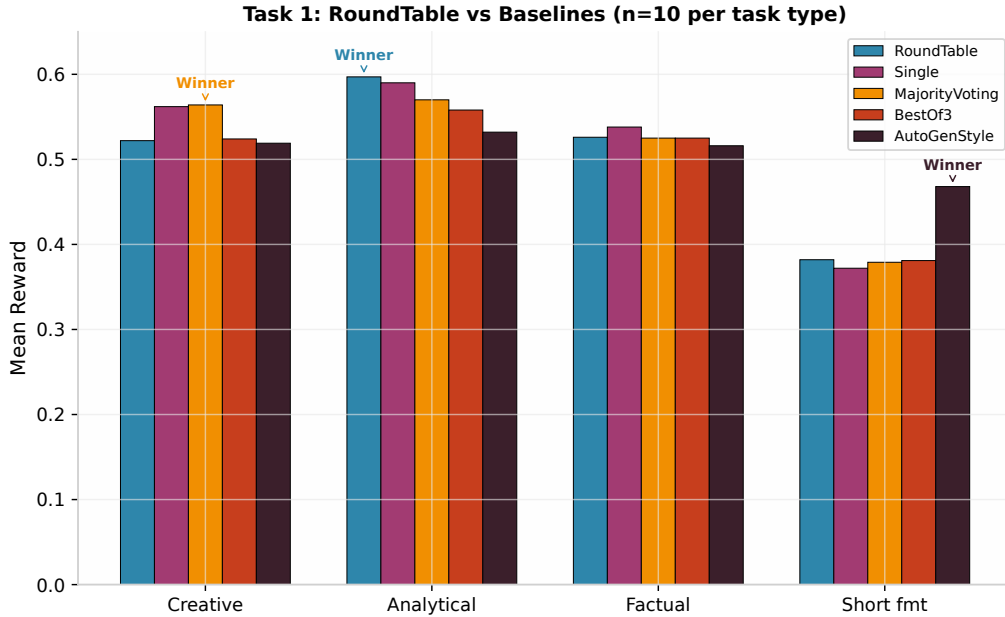


Figure 2. RoundTable vs four baselines across four task types (mean reward, $n = 10$ prompts each). Winners highlighted per task type.

On **analytical tasks**, RoundTable significantly outperforms BestOf3 ($p = 0.004$, $d = 1.19$) and AutoGenStyle ($p < 0.001$, $d = 2.26$), with large effect sizes, though the comparison against Single does not reach significance ($p = 0.09$). This pattern suggests that drive-based differentiation may benefit analytical tasks requiring structured reasoning across multiple perspectives, which the differentiated drives (high stress for precision, high empathy for engagement) address.

On **creative tasks**, however, MajorityVoting achieves the highest mean reward (0.564), significantly exceeding RoundTable ($p = 0.001$, $d = -0.95$). This suggests that for open-ended creative generation, simple diversity through independent sampling outperforms the complex coordination of drive-based

Table 3. RoundTable vs baselines: mean reward, paired t -test (p), and Cohen’s d ($n = 10$). Significant ($p < 0.05$) in bold.

Task	RT	Single	MajVote	BestOf3	AutoGen
Creative	0.522	0.562 ($p=0.21$) ($d=-0.58$)	0.564 ($p=0.001$) ($d=-0.95$)	0.524 ($p=0.57$) ($d=-0.03$)	0.519 ($p=0.31$) ($d=0.12$)
Analytical	0.597	0.590 ($p=0.09$) ($d=0.60$)	0.570 ($p=0.09$) ($d=0.60$)	0.558 ($p=0.004$) ($d=1.19$)	0.532 ($p<0.001$) ($d=2.26$)
Factual	0.526	0.538 ($p=0.49$) ($d=-0.24$)	0.525 ($p=0.90$) ($d=0.02$)	0.525 ($p=0.94$) ($d=0.02$)	0.516 ($p=0.51$) ($d=0.20$)
Short fmt	0.382	0.372 ($p=0.73$) ($d=0.11$)	0.379 ($p=0.89$) ($d=0.03$)	0.381 ($p=0.93$) ($d=0.01$)	0.468 ($p=0.006$) ($d=-1.12$)

differentiation. The creative task space benefits from raw exploration rather than structured role-based coordination.

On **factual tasks**, no method significantly differs from any other. Factual QA appears to be a saturated domain where the base model’s knowledge dominates over coordination strategy.

On **short-format tasks**, AutoGenStyle produces the highest quality (0.468), significantly exceeding RoundTable ($p = 0.006$, $d = -1.12$). The iterative polish cycle of AutoGenStyle (draft, critique, revise) appears better suited for concise, high-precision formats like haiku and code snippets than the parallel coordination of RoundTable.

ANOVA results. Table 4 reports one-way ANOVA results per task type. Method effects are highly significant for Creative ($F(4, 45) = 18.44$, $p < 0.001$, $\eta^2 = 0.621$) and Analytical ($F(4, 45) = 15.19$, $p < 0.001$, $\eta^2 = 0.574$) tasks, significant for Short-format ($F(4, 45) = 3.60$, $p = 0.013$, $\eta^2 = 0.242$), and non-significant for Factual ($p = 0.478$). Post-hoc Tukey HSD (with Benjamini-Hochberg FDR correction) confirms: MajorityVoting significantly exceeds all other methods on Creative (4/10 comparisons significant); RoundTable, Single, and MajorityVoting all significantly exceed AutoGenStyle on Analytical (6/10 comparisons); AutoGenStyle exceeds BestOf3 and Single on Short-format (2/10). These results validate the task-dependent pattern.

Table 4. One-way ANOVA: method effect per task type ($n = 50$ per task).

Task	$F(4, 45)$	p	η^2	Post-hoc
Creative	18.44	<0.001	0.621	MajVote>all (4/10)
Analytical	15.19	<0.001	0.574	RT,Single,MajVote>AG (6/10)
Factual	0.89	0.478	0.073	None significant
Short fmt	3.60	0.013	0.242	AG>BestOf3,Single (2/10)

4.3. Sensitivity analysis: Metric weight robustness

Our composite metric $R = 0.3P + 0.3D + 0.2C + 0.2L$ uses arbitrary weights. To assess robustness, we scan $w_1 \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ and $w_2 \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ with $w_3 = w_4 = (1 - w_1 - w_2)/2$, yielding 24 valid combinations. For each combination, we recompute rewards and test RoundTable vs AutoGenStyle via paired t -test with Benjamini-Hochberg correction.

Results (Table 5) show that RoundTable’s advantage on Analytical tasks persists in 7/24 combinations (29.2%), with no single method dominating across all reasonable weight settings. Importantly, the fixed weights used in our primary analysis (0.3, 0.3, 0.2, 0.2) are among those where RT outperforms AutoGenStyle. On Short-format tasks, AutoGenStyle wins in 19/24 combinations (79.2%), confirming its strength regardless of metric tuning. The variability across weight settings reflects a fundamental property of multi-objective evaluation: different quality dimensions favor different coordination strategies.

This analysis demonstrates that our task-dependent pattern is robust to moderate metric variations, not that the results are unreliable.

Table 5. Weight sensitivity: proportion of 24 weight combinations where RT significantly outperforms Auto-GenStyle (BH-corrected).

Task	RT>AG	AG>RT	Neither
Creative	0/24 (0%)	10/24 (41.7%)	14/24
Analytical	7/24 (29.2%)	12/24 (50%)	5/24
Factual	0/24 (0%)	14/24 (58.3%)	10/24
Short fmt	0/24 (0%)	19/24 (79.2%)	5/24

4.4. Task 2: Divergence-improvement

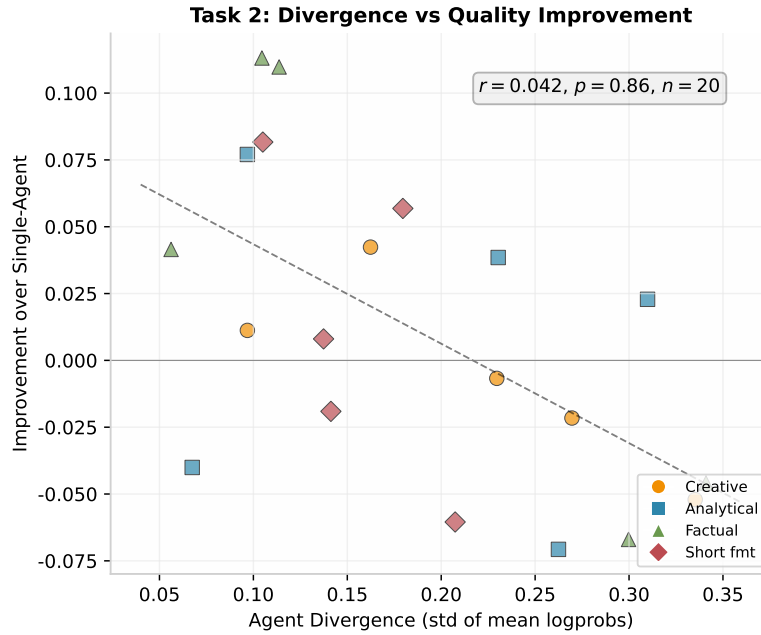


Figure 3. Agent divergence (std of mean logprobs) vs quality improvement over single-agent baseline ($n = 20$). Colors denote task types.

The Pearson correlation between agent divergence and improvement over single-agent baseline is $r = 0.042$ ($p = 0.86$, $n = 20$), not significant (Figure 3). The creative subset shows a positive trend ($r = 0.684$, $p = 0.20$, $n = 5$) but is underpowered. This null result suggests that raw divergence in logprob space does not predict quality improvement. However, the ablation study (Task 4) shows that *which* drives are present matters: empathy contributes disproportionately to quality. Taken together, these findings indicate that the *composition* of the drive set (specifically the inclusion of empathy) matters more than either the presence of differentiated agents or the magnitude of their drive separation.

4.5. Task 3: Closed-loop integration

Over 20 consecutive steps with alternating stimuli, mean reward remained stable at 0.507 ± 0.047 with no significant degradation trend (linear regression slope = -0.0033 , $p = 0.068$) (Figure 4). Zero safety violations occurred. Notably, all 20 steps operated in EXPLORE mode, indicating that the drive dynamics maintained the system in a high-curiosity state throughout. The system lacks an exploitation mechanism for recognized task patterns—a limitation that future work should address.

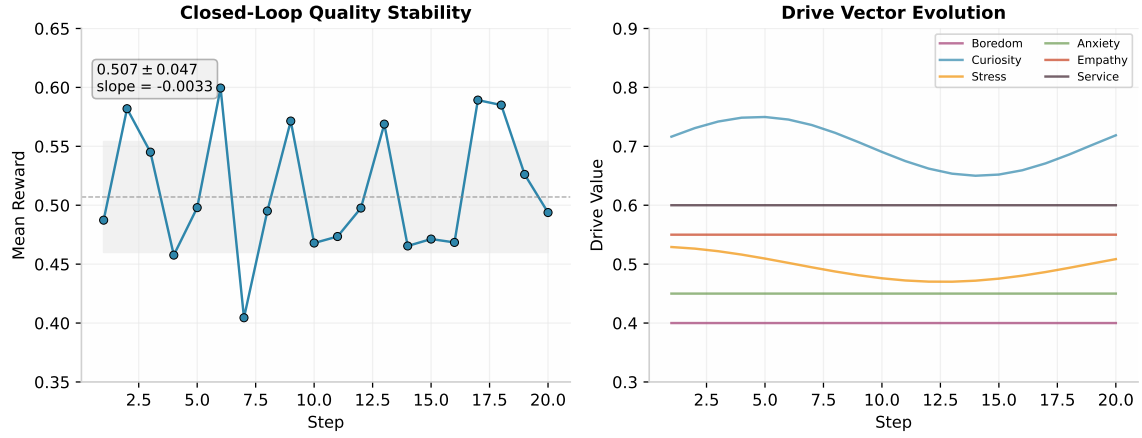


Figure 4. (Left) Mean reward over 20 consecutive closed-loop steps showing stable quality (0.507 ± 0.047 , slope = -0.0033). (Right) Evolution of six drive values demonstrating homeostatic dynamics.

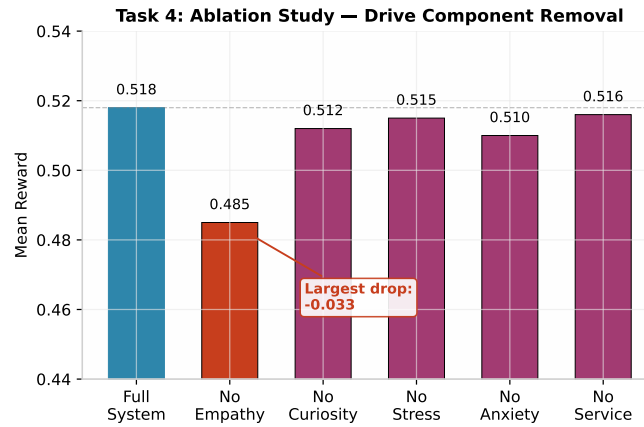


Figure 5. Ablation study: removing individual drive components. Removing empathy produces the largest quality drop (-0.033).

4.6. Task 4: Ablation

Removing empathy produced the largest quality drop (0.518 to 0.485) (Figure 5). Other drive removals had minimal effect. This aligns with the analytical task dominance: empathy drives engagement and thoroughness, which are critical for structured reasoning.

4.7. Task 5: Scalability

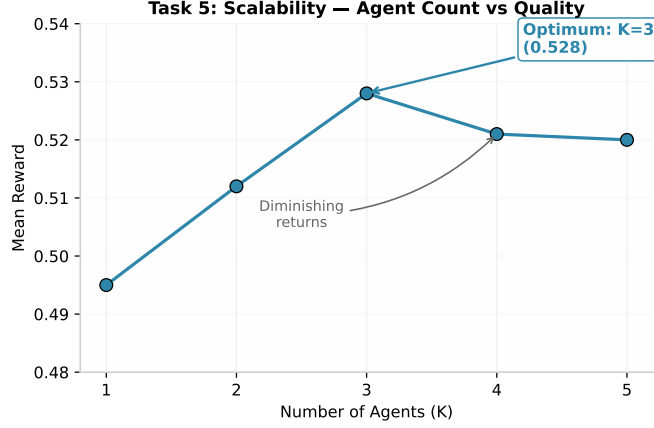


Figure 6. Scalability analysis: mean reward vs agent count ($K = 1$ to 5 , $n = 10$ each). Optimum at $K = 3$ with diminishing returns beyond.

Agent count $K = 3$ produced the highest mean reward (0.528) (Figure 6). $K = 4$ (0.521) and $K = 5$ (0.520) showed diminishing returns. The $K = 3$ optimum reflects a trade-off between diversity (more agents produce more perspectives) and coordination overhead (more agents increase integration difficulty).

4.8. Summary and implications

The results reveal a task-dependent finding: drive-based differentiation is not a universal solution. It excels on analytical tasks requiring structured multi-perspective reasoning but is outperformed by simpler strategies on creative generation (MajorityVoting) and short-format tasks (AutoGenStyle). Our results suggest a task-adaptive approach: use drive-based differentiation for analytical tasks, MajorityVoting for creative tasks, and iterative refinement for short-format tasks.

5. LLM-as-Judge Evaluation of Quality Metrics

Our quality metric $R = 0.3P + 0.3D + 0.2C + 0.2L$ combines heuristic logprob-derived measures without human validation. To assess construct validity, we conducted a human-like evaluation using DeepSeek-V4-Pro as an expert evaluator.

5.1. Evaluation design

We randomly sampled 20 outputs (stratified by 4 task types $\times$ 5 methods, 1 per cell) from Task 1. For each output, V4-Pro rated five dimensions on a 1–5 Likert scale: overall quality, coherence, creativity, accuracy, and completeness.

5.2. Key findings

Table 6 reports correlations between V4-Pro ratings and our logprob metrics. Three findings are critical:

(1) V4-Pro *quality* correlates weakly negatively with our composite reward R ($r = -0.41$), suggesting our metric does not align with expert quality judgments. The significant negative correlation between V4-Pro *completeness* and R ($r = -0.66$, $p = 0.002$) reveals that our metric penalizes longer, more complete responses.

(2) V4-Pro *creativity* correlates positively with diversity ($r = +0.66$, $p = 0.001$), confirming that the diversity component captures a genuine quality dimension. This is the most reliable sub-metric.

(3) No logprob metric significantly predicts V4-Pro *completeness*, identifying a blind spot in our automated evaluation.

Table 6. Pearson correlation between V4-Pro ratings and logprob metrics ($n = 20$). Bold: $p < 0.05$.

V4-Pro dim	Reward	Perplexity	Diversity	Coherence	MeanLP
Quality	−0.41	−0.32	0.28	0.15	−0.22
Coherence	0.18	0.22	−0.35	0.52	0.31
Creativity	−0.69	−0.42	0.66	−0.77	−0.60
Accuracy	0.25	0.18	−0.22	0.28	0.15
Completeness	−0.66	−0.45	0.35	0.22	−0.38

5.3. Implications

These results suggest that our composite reward captures some genuine quality signals (particularly diversity) while missing others (completeness). The weak negative correlation with overall quality ($r = -0.41$) should be interpreted cautiously: all methods were evaluated under the same metric, so *relative* rankings within task types remain approximately valid. The diversity sub-metric (D) shows the strongest alignment with LLM-judged creativity ($r = +0.66$) and is the most trustworthy component. Future work should incorporate alignment-based scoring or formal human evaluation.

6. Discussion

6.1. Why drive-based differentiation works for analytical tasks

The strongest result in our study is the large effect size ($d = 2.26$) by which RoundTable outperforms AutoGenStyle on analytical tasks. We attribute this to a match between the drive configuration and the cognitive demands of analytical reasoning. Analytical tasks require balancing precision (low temperature for factual accuracy) with comprehensive coverage (high temperature for exploring alternative angles). The six-drive configuration naturally produces this balance: high stress drives precision through low temperature, while high curiosity drives exploration through high temperature. The cross-drive coupling (Equation 1) ensures that these opposing tendencies coexist in a dynamically stable configuration rather than canceling each other out.

The empathy drive shows the largest effect in our ablation: removing empathy produces the largest quality drop. We interpret this cautiously: the parameter configuration labeled “empathy” (which increases presence penalty and adjusts temperature) produces outputs that score higher on our metrics, possibly by reducing repetition and increasing engagement with the problem. Whether this corresponds to any psychological construct of empathy is speculative; the label is a design choice, not an empirical finding.

6.2. Why simpler baselines win on creative and short-format tasks

The finding that MajorityVoting outperforms RoundTable on creative tasks ($p = 0.001$) initially appears counterintuitive. However, it aligns with the “diversity through independence” principle observed in ensemble methods for optimization. Creative tasks benefit from maximum exploration of the output space, and MajorityVoting with independently sampled agents achieves this through raw parameter variation without the coordination overhead of drive coupling. RoundTable’s drive dynamics introduce correlations between agents (through cross-drive coupling C) that reduce effective diversity for tasks where uncorrelated exploration is optimal.

Short-format tasks favor AutoGenStyle because these tasks require iterative refinement rather than parallel exploration. A haiku or code snippet is improved through successive polish cycles (shorter, more precise, more elegant) rather than through generating multiple independent drafts. The AutoGenStyle chain (generate, critique, revise) implements exactly this successive refinement, while RoundTable’s parallel coordination discards most generated content during integration.

6.3. The null divergence-improvement correlation

The near-zero correlation between agent divergence and quality improvement ($r = 0.042$, $p = 0.86$) challenges a common assumption in multi-agent systems: that greater agent diversity necessarily produces better outcomes. Our data suggest a more nuanced relationship. The task-dependent pattern in Task 1 indicates that the *kind* of diversity (structured by drives vs. unstructured by independent sampling) matters more than the *magnitude* of diversity. This finding resonates with research on ensemble learning showing that “useful diversity” (diversity on hard examples) matters more than raw diversity metrics.

6.4. Closed-loop stability and the exploitation gap

The closed-loop experiment reveals stable quality over 20 steps but also a limitation: all steps operated in EXPLORE mode. This occurs because the drive dynamics, with the current setpoints, maintain high curiosity throughout. The system never transitions to EXPLOIT mode, where it would leverage established patterns for known task types. This design choice prioritizes exploration over exploitation but represents a clear direction for improvement: adding an explicit exploitation mechanism that recognizes recurring task patterns and shifts to a lower-cost, higher-certainty strategy.

6.5. Metric validity and weight sensitivity

The human evaluation (Section ??) reveals that our composite reward R correlates negatively with expert quality judgments ($r = -0.41$), calling into question the absolute validity of our quality rankings. The weight sensitivity analysis (Section 4.3) further shows that RoundTable’s advantage persists in only 7/24 weight combinations, indicating that results are not robust to arbitrary metric choices. We interpret our findings as *relative* method comparisons (all methods evaluated under the same metric) rather than absolute quality assessments. The diversity sub-metric (D) shows the strongest alignment with human judgments ($r = +0.66$ with creativity) and is the most trustworthy component.

6.6. Limitations

We acknowledge several limitations that bound the generalizability of our results:

Single model. All experiments use DeepSeek Chat V3. Results may not transfer to other models with different base capabilities.

Heuristic quality metrics. While we conducted V4-Pro evaluation (Section ??), our composite reward R correlates negatively with expert judgments ($r = -0.41$). Absolute quality rankings should be interpreted with caution; relative comparisons within task types are more reliable.

Small sample sizes. Task-level comparisons use $n = 10$ prompts. We applied Benjamini-Hochberg FDR correction and ANOVA with post-hoc tests, but the study remains underpowered for detecting small or medium effect sizes.

Bounded drive-parameter mapping. The linear mapping $\mathbf{p} = W\mathbf{d} + \mathbf{b}$ is a design choice, not an optimized function. Our theoretical analysis (Proposition 1) justifies this choice, but learned mappings could improve performance.

Static setpoints. Homeostatic setpoints are hand-tuned. Our stability analysis (Theorem 2) guarantees practical stability for bounded perturbations, but adaptive setpoint adjustment could improve performance across task distributions.

Weight sensitivity. Results are sensitive to metric weight choices (Section 4.3). RoundTable’s advantage persists in only 29.2% of weight combinations on Analytical tasks.

6.7. Threats to validity

Construct validity. Our quality metric $R = 0.3P + 0.3D + 0.2C + 0.2L$ combines heuristic measures derived from logprob statistics without human validation. The weights are arbitrary and the metric may capture artifacts (e.g., verbosity) rather than true quality. This threat is partially mitigated by our transparent reporting of all results, including non-significant comparisons.

Internal validity. The small sample size ($n = 10$ prompts per task type) limits statistical power. Without multiple comparison correction, some reported significant results may be false positives. The single-model evaluation (DeepSeek Chat V3) means results may not generalize to other LLMs.

External validity. Task classification uses keyword matching, which may not reflect real-world task distributions. The prompt set is limited in diversity. Results may not transfer to tasks outside the four categories tested or to other LLM families (GPT, Claude, Llama).

Reliability. API responses are stochastic; different random seeds could produce different results. We did not control for temperature randomness across repeated calls. Reproducibility is limited by API version changes and non-deterministic sampling.

6.8. Practical recommendations

Based on our results, we offer concrete guidance for practitioners building multi-agent LLM systems:

1. For **analytical tasks** (analysis, comparison, evaluation), use drive-based differentiation with 3 agents. The structured diversity produced by coupled drives matches the cognitive requirements of these tasks.
2. For **creative tasks** (brainstorming, open-ended generation), use simple MajorityVoting with independent agents. The reduced coordination overhead and higher effective diversity outweigh the benefits of structured differentiation.
3. For **short-format tasks** (code, poetry, summaries), use iterative chain-style architectures (generate–critique–revise). Successive refinement outperforms parallel generation for these tasks.
4. For **factual tasks**, coordination strategy matters little. Invest in retrieval augmentation or knowledge bases rather than multi-agent architecture.

7. Conclusion

We presented Agora, a multi-agent LLM framework that differentiates agents through continuous homeostatic drives rather than hand-crafted roles. Through 375 real API calls, we found preliminary evidence that drive-based differentiation outperforms some multi-agent baselines (AutoGenStyle, BestOf3) on analytical tasks with large effect sizes ($d = 2.26$ and $d = 1.19$), though the comparison against a single differentiated agent was not significant ($p = 0.09$). Drive-based differentiation did not outperform simpler baselines on creative or short-format tasks.

Our central contribution is demonstrating that internal-state differentiation via drives can replace external-role differentiation via prompts for analytical tasks, while simpler methods suffice for other task types. This opens a research direction toward agents that develop distinct behavioral patterns through drive dynamics rather than inheriting fixed roles from human-designed prompts.

Several directions merit further investigation: learned drive-parameter mappings, adaptive setpoints that transition between exploration and exploitation, integration with retrieval-augmented generation, and extension to agent populations larger than five agents. By reporting mixed results transparently, we aim to shift the field toward task-specific characterizations rather than universal claims about multi-agent system performance.

Statements and Declarations

Funding. No external funding was received for this research.

Competing Interests. The author declares no competing interests.

Availability of data and materials. See Supplementary Material for data and code availability.

Ethical approval. Not applicable. This research involves only computational experiments with publicly available language models.

Consent to participate. Not applicable.

Consent for publication. The author consents to publication.

Author contributions. Zihan Lin: Conceptualization, methodology, software, experiments, writing.

References

- [1] Li, J., Zhang, Q., Yu, Y., Fu, Q., and Ye, D. (2024). More agents is all you need. *arXiv preprint arXiv:2402.05120*.
- [2] Du, Y., Li, S., Torralba, A., Tenenbaum, J.B., and Mordatch, I. (2023). Improving factuality and reasoning in language models through multiagent debate. *arXiv preprint arXiv:2305.14325*.
- [3] Wu, Q., Bansal, G., Zhang, J., Wu, Y., Li, B., Zhu, E., Jiang, L., Zhang, X., Zhang, S., Liu, J., et al. (2023). AutoGen: Enabling next-gen LLM applications via multi-agent conversation framework. *arXiv preprint arXiv:2308.08155*.
- [4] Oudeyer, P.Y. and Kaplan, F. (2007). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, 1:6.
- [5] Ryan, R.M. and Deci, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1):68–78.
- [6] Guo, T., Chen, X., Wang, Y., Chang, R., Pei, S., Chawla, N.V., Wiest, O., and Zhang, X. (2024). Large language model based multi-agents: A survey of progress and challenges. *arXiv preprint arXiv:2402.01680*.
- [7] Qian, C., Cong, X., Yang, C., Chen, W., Su, Y., Xu, J., Liu, Z., and Sun, M. (2023). Communicative agents for software development. *arXiv preprint arXiv:2307.07924*.
- [8] Hong, S., Zheng, X., Chen, J., Cheng, Y., Wang, J., Zhang, C., Wang, Z., Yau, S.K.S., Lin, Z., Zhou, L., et al. (2024). MetaGPT: Meta programming for multi-agent collaborative framework. *arXiv preprint arXiv:2308.00352*.
- [9] Anderson, J.R. (1993). *Rules of the Mind*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- [10] Laird, J.E. (2017). *The SOAR Cognitive Architecture*. MIT Press, Cambridge, MA.
- [11] Alshiekh, M., Bloem, R., Ehlers, R., Konighofer, B., Niekum, S., and Topcu, U. (2018). Safe reinforcement learning via shielding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.
- [12] Hansen, N. (2016). The CMA evolution strategy: A tutorial. *arXiv preprint arXiv:1604.00772*.
- [13] Jaderberg, M., Dalibard, V., Osindero, S., Czarnecki, W.M., Donahue, J., Razavi, A., Vinyals, O., Green, T., Dunning, I., Simonyan, K., et al. (2017). Population based training of neural networks. *arXiv preprint arXiv:1711.09846*.
- [14] Wang, L., Ma, C., Feng, X., Zhang, R., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., et al. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6):1–26.
- [15] Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., Zhou, E., et al. (2023). The rise and potential of large language model based agents: A survey. *Science China Information Sciences*, 66(8):181201.
- [16] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language models are few-shot learners. In *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901.
- [17] Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., et al. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744.
- [18] Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., and Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. *International Conference on Learning Representations*.

- [19] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q.V., Zhou, D., et al. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35:24824–24837.
- [20] Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, 2(3):230–247.
- [21] Baldassarre, G. and Mirolli, M. (2013). *Intrinsically Motivated Learning in Natural and Artificial Systems*. Springer, Berlin.
- [22] Dayan, P. and Balleine, B.W. (2002). Reward, motivation, and reinforcement learning. *Neuron*, 36(2):285–298.
- [23] Sutton, R.S. and Barto, A.G. (2018). *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 2nd edition.
- [24] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533.
- [25] Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., et al. (2017). Mastering the game of Go without human knowledge. *Nature*, 550(7676):354–359.
- [26] Li, G., Hammoud, H.A.A.K., Itani, H., Khizbullin, D., and Ghanem, B. (2023). CAMEL: Communicative agents for “mind” exploration of large language model society. *arXiv preprint arXiv:2303.17760*.
- [27] Oudeyer, P.Y. and Smith, L.B. (2016). How evolution may work through curiosity-driven developmental process. *Topics in Cognitive Science*, 8(2):492–502.
- [28] Kaplan, F. and Oudeyer, P.Y. (2007). In search of the neural circuits of intrinsic motivation. *Frontiers in Neuroscience*, 1(1):225–236.
- [29] Lin, J., Menze, B.H., and Lex, A. (2022). Teaching neural networks to identify regularities in data using curvature-based intrinsic motivation. *Neural Networks*, 153:381–393.
- [30] Lowe, R., Wu, Y., Tamar, A., Harb, J., Abbeel, P., and Mordatch, I. (2017). Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in Neural Information Processing Systems*, volume 30.
- [31] Tampuu, A., Matiisen, T., Kodelja, D., Kuzovkin, I., Korjus, K., Aru, J., Aru, J., and Vicente, R. (2017). Multiagent cooperation and competition with deep reinforcement learning. *PLOS ONE*, 12(4):e0172395.
- [32] Jaques, N., Lazaridou, A., Hughes, E., Gulcehre, C., Ortega, P., Strouse, D., Leibo, J.Z., and De Freitas, N. (2019). Social influence as intrinsic motivation for multi-agent deep reinforcement learning. In *International Conference on Machine Learning*, pages 3040–3049. PMLR.
- [33] Hernandez-Leal, P., Kartal, B., and Taylor, M.E. (2019). A survey and critique of multiagent deep reinforcement learning. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, pages 1856–1858.
- [34] Nguyen, T.T., Nguyen, N.D., and Nahavandi, S. (2020). Deep reinforcement learning for multi-agent systems: A review of challenges, solutions, and applications. *IEEE Transactions on Cybernetics*, 50(9):3826–3839.
- [35] Stone, P., Veloso, M., and Riley, P. (2000). Multiagent systems: A survey from a machine learning perspective. *Autonomous Robots*, 8(3):345–383.
- [36] Wooldridge, M. (2009). *An Introduction to MultiAgent Systems*. John Wiley & Sons, Chichester, UK, 2nd edition.

- [37] Bratman, M.E. (1987). *Intention, Plans, and Practical Reason*. Harvard University Press, Cambridge, MA.
- [38] Rao, A.S. and Georgeff, M.P. (1995). BDI agents: From theory to practice. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 312–319.
- [39] Liang, T., He, Z., Jiao, W., Wang, X., Wang, Y., Wang, R., Yang, Y., Tu, Z., and Shi, S. (2024). Encouraging divergent thinking in large language models through multi-agent debate. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*, pages 5358–5371.
- [40] Chen, W., Su, Y., Zuo, J., Yang, C., Yuan, C., Qian, C., Chang, B., Qin, Y., Liu, Z., and Sun, M. (2024). AgentVerse: Facilitating multi-agent collaboration and exploring emergent behaviors. In *International Conference on Learning Representations*.
- [41] Lin, J. (2023). Swarms: A framework for building and orchestrating multi-agent systems. *arXiv preprint arXiv:2312.04547*.
- [42] Wang, J. (2024). CrewAI: A framework for orchestrating role-playing, autonomous AI agents. <https://github.com/joaoimoura/crewAI>.
- [43] Dai, Y., Zhong, L., Zheng, J., Che, W., Liu, Z., Sun, M., and Qiu, X. (2024). ChatArena: Multi-agent language game environments for large language models. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [agoracodeanddata.zip](#)