

Emergent Intelligence: A Framework for Multi-Agent Systems Inspired by Minsky's Society of Mind

Abstract

This paper presents a conceptual framework for developing an Artificial Intelligence (AI) Multi-Agent System (MAS) based on Marvin Minsky's "Society of Mind" (SoM) theory. The central premise of Minsky's theory is that intelligence is not a monolithic entity but rather an emergent property arising from the interactions of numerous simple, non-intelligent agents. The paper contrasts Minsky's concept of "mindless" agents with the typically complex, autonomous agents found in modern MAS research. It highlights the core challenge of translating Minsky's philosophical ideas into a functional computational model.

The proposed framework is built on a bottom-up philosophy, emphasizing the simplicity of individual agents which would possess minimal internal logic and perform highly specialized functions. A key element of this framework is a computational analogue to Minsky's "K-lines," a mechanism for memory and learning that captures and reactivates successful configurations of agents, thereby facilitating structural learning without explicit programming.

To explore the emergence of intelligence within this framework, the paper suggests several computational approaches, including Multi-Agent Reinforcement Learning (MARL), Evolutionary Algorithms (EAs), and Agent-Based Modeling (ABM). It posits that a hybrid approach, combining the simulation strengths of ABM with the adaptive capabilities of EAs or RL, would be most effective.

A significant portion of the paper is dedicated to the challenges of evaluating such a system. It argues that traditional performance metrics are insufficient, and evaluation should instead focus on adaptability, robustness, and the emergence of hierarchical abstraction. The paper also addresses the profound difficulties of scalability, credit assignment, and bridging the gap from low-level agent interactions to high-level cognitive functions.

In conclusion, the paper advocates for a renewed investigation into Minsky's vision as a viable, albeit challenging, alternative path toward Artificial General Intelligence (AGI). It suggests that such an approach could lead to more robust and adaptable

systems, offering deep insights into the nature of cognition itself.

1. Introduction

The quest to understand and replicate human-level intelligence, often termed Artificial General Intelligence (AGI), remains one of the most profound challenges in science and engineering. While significant progress has been made in specialized AI domains, current approaches often struggle with the flexibility, adaptability, and common-sense reasoning characteristic of human cognition.¹ This suggests the potential value of exploring alternative paradigms beyond monolithic models or systems focused solely on optimizing predefined tasks.

One such compelling, albeit largely conceptual, alternative stems from the work of Marvin Minsky. His "Society of Mind" (SoM) theory posits that intelligence is not a singular entity but an emergent property arising from the complex interactions of numerous, individually simple, "mindless" processes, which he termed "agents".² Minsky argued against seeking a single, perfect principle for intelligence, instead proposing that its power originates from the vast diversity and specific organization of these simple components: "What magical trick makes us intelligent? The trick is that there is no trick. The power of intelligence stems from our vast diversity, not from any single, perfect principle".³

Minsky's vision resonates with contemporary interest in Multi-Agent Systems (MAS) within AI, particularly concerning emergent phenomena and bottom-up approaches to complexity.⁵ While modern MAS often focus on coordinating the capabilities of already sophisticated agents to solve specific problems⁸, Minsky's theory invites us to consider how intelligence itself might arise from the collective behavior of fundamentally simpler entities. However, translating Minsky's highly conceptual and often philosophical work³ into a concrete computational framework presents significant challenges. His agents are defined by their mindlessness and simplicity³, whereas modern MAS agents are typically characterized by their autonomy and complex capabilities.⁹ This fundamental difference means that building a Minsky-inspired system requires more than a direct mapping; it necessitates interpreting and operationalizing the core principles of SoM—emergence from simplicity—within current computational paradigms. Bridging this conceptual gap is a central task.

This report aims to develop such a conceptual framework for an AI MAS explicitly inspired by Minsky's Society of Mind. It will analyze the key differences between Minsky's notion of agents and those in contemporary MAS, explore computational

methodologies suitable for investigating emergent intelligence within this framework (specifically addressing Minsky's "very special ways" of agent organization ⁴), discuss the challenges of evaluating genuine emergence, and identify key limitations and potential future directions. The subsequent sections will delve into Minsky's theory, describe modern MAS, compare the two concepts of "agent," propose the SoM-inspired framework, discuss computational approaches and evaluation methods, review related work, analyze challenges, and offer concluding thoughts.

2. Marvin Minsky's Society of Mind

2.1 Core Thesis: Intelligence as an Emergent Society

Marvin Minsky's *Society of Mind* theory, primarily articulated in his 1986 book, presents a radical departure from viewing the mind as a unified, centralized processor.³ Instead, Minsky proposed that the mind is a vast "society" composed of myriad simple, specialized processes called "agents".² The central tenet is that intelligence, along with other cognitive phenomena like thought, common sense, emotion, and memory, is not an intrinsic property of any individual agent but rather an *emergent property* arising from the complex, dynamic interactions among these agents when they are organized in "certain very special ways".³ The complexity and power of the mind, therefore, stem from the sheer number, diversity, and intricate interconnectedness of these fundamentally simple components.³

2.2 Agents: The Mindless Building Blocks

A cornerstone of the SoM theory is the nature of the agents themselves. Minsky defined agents as simple, specialized processes, each capable of performing only a very basic function that, in itself, requires no intelligence or thought.³ Examples might include agents for detecting simple patterns, activating a muscle, comparing two values, or maintaining balance.¹⁵ This concept contrasts sharply with the search for a single, powerful, general-purpose algorithm underlying intelligence.³

Crucially, these agents are considered individually "mindless".³ Intelligence does not reside within the components but emerges from their collective activity. This "mindlessness" is not a limitation to be overcome but a fundamental aspect of the theory, suggesting that true intelligence can be built from non-intelligent parts. Minsky argued that the robustness and adaptability of intelligence stem precisely from this *diversity* of simple, specialized agents, rather than a single, potentially brittle, complex mechanism.³

2.3 Societies and Agencies: Organizing the Collective

These simple agents do not operate in isolation. They combine and interact to form larger functional units, which Minsky referred to as "societies" or "agencies".³ An agency is essentially a society of agents organized to perform a more complex function than any individual agent could achieve alone.¹⁴ For example, an agency for "building a tower" might involve agents for seeing blocks, grasping, moving, balancing, and recognizing completion.

Minsky suggested various organizational structures for these societies. He introduced the concept of "level-bands," implying a hierarchical organization where different processes operate at different levels of abstraction or detail simultaneously.¹² This allows, for instance, one part of the mind to focus on fine motor control while another handles high-level planning.¹² He also proposed layered structures, potentially involving K-societies built upon S-societies (agent societies).¹³ Furthermore, interactions could involve competition between agencies (e.g., for control or resources) or cooperation within an agency.¹³ A particularly intriguing idea is the distinction between "A-brains" and "B-brains," where the A-brain interacts with the external world, and the B-brain monitors the internal state of the A-brain, suggesting a mechanism for self-awareness or meta-cognition.¹³

2.4 Interaction Mechanisms: The Fabric of Thought

Minsky proposed several mechanisms through which agents and agencies could interact:

- **K-Lines (Knowledge-Lines):** Perhaps the most elaborated mechanism, K-lines are central to memory and learning in SoM.¹² A K-line is described as a "wirelike structure" created when a problem is solved or a memorable experience occurs.¹³ It attaches itself to the set of mental agents that were active during that event.²¹ When the K-line is later activated (e.g., by a similar situation or cue), it reactivates this specific configuration of agents, creating a "partial mental state" resembling the original one.¹⁹ This allows the mind to leverage past successful configurations to tackle new, related problems.¹⁴ K-lines effectively "chunk" experiences, capturing not just the solution but also aspects of the process, including false starts.¹⁴ Minsky suggested K-lines could connect to other K-lines, forming "K-societies" layered upon the primary agent societies.¹³ He also proposed refinements like the "Level-Band Principle" (K-lines connect to intermediate levels of an agency hierarchy) and the "K-Recursion Principle" (new K-lines primarily connect to existing active K-lines, building memories hierarchically).²² The formation and reactivation of K-lines represent a form of structural learning. When a particular configuration of agents leads to a successful outcome (solving a problem, having a useful idea), the creation of a K-line reinforces that

configuration by making it easier to reactivate later.¹³ This acts as an implicit credit assignment mechanism: success is credited to the *active society* of agents, and that society is made more persistent or recallable via the K-line, without needing an explicit external reward signal.

- **Frames:** To explain commonsense reasoning and the representation of stereotypical knowledge, Minsky introduced "frames".¹² A frame is a data structure, like a skeletal outline, representing a concept or situation (e.g., "bird," "room," "birthday party") with "slots" for specific details.¹² These slots can hold default assumptions (e.g., a bird usually flies) that can be overridden by specific information.¹² Frames help organize cultural knowledge, understand language, and make commonsense inferences by providing context and expectations.¹² Related concepts include polynomes, nemes (representing aspects of the world), and nomes (controlling representation processing).¹⁴
- **Competition, Cooperation, and Exploitation:** Agent interactions are not always harmonious. Minsky acknowledged that agencies might compete, perhaps for control or resources.¹³ Within an agency, agents cooperate. He also highlighted "exploitative" activity, where agents learn to use the outputs or capabilities of other agents without needing to understand their internal workings.¹³ This avoids the need for complex, universal communication protocols.¹⁸
- **Censors and Suppressors:** Minsky briefly mentioned mechanisms like censors and suppressors, likely involved in inhibiting or modulating the activity of certain agents or agencies, perhaps preventing undesirable actions or thoughts.¹⁷

2.5 Emergence of Intelligence: The "Very Special Ways"

The core message of SoM is that intelligence is not designed or programmed directly but *emerges* from the collective dynamics of the agent society.² The specific ways agents are interconnected and interact—the "very special ways"—are what lead to intelligent behavior.⁴ Minsky aimed to explain a wide spectrum of cognitive functions through this emergent process, including learning, different types of reasoning (distinguishing the apparent difficulty of logical reasoning from the deeper complexity of common sense¹²), memory, language understanding, consciousness, emotions, and self-awareness.³ His ideas were significantly influenced by his practical work in the 1970s attempting to build a robot that could perceive and build with children's blocks, where he found that no single algorithm sufficed and only a society of different processes seemed viable.³

3. Modern Multi-Agent Systems (MAS) in AI

3.1 Definition and Scope

In contemporary Artificial Intelligence, a Multi-Agent System (MAS) is typically defined as a computerized system comprising multiple interacting intelligent agents.⁸ These systems are often employed to tackle problems that are inherently distributed, too complex for a single agent, or require the integration of diverse capabilities.⁹ The agents within a MAS interact within a shared environment, pursuing individual or collective goals.⁸ The advent of powerful Large Language Models (LLMs) has spurred new research into LLM-based MAS, where agents leverage language capabilities for more sophisticated reasoning, communication, and coordination.⁹ While related to Agent-Based Modeling (ABM), MAS research often emphasizes engineering solutions and achieving specific tasks, whereas ABM frequently focuses on simulating and understanding emergent phenomena in natural or social systems.⁹

3.2 Agent Characteristics in Modern MAS

Agents in modern MAS are generally characterized by several key properties:

- **Autonomy:** Agents possess a significant degree of autonomy, meaning they can operate independently, perceive their environment, make decisions, and take actions without direct external control or intervention.⁹ They manage their own internal state and behavior.
- **Capabilities:** Unlike Minsky's simple agents, modern MAS agents are often endowed with substantial capabilities. These can include sensing the environment, acting upon it, reasoning, planning sequences of actions, learning from experience (e.g., via reinforcement learning), and communicating with other agents.⁹ Agent capabilities can range from simple reactive behaviors (active agents) to complex cognitive processing (cognitive agents).⁹ LLMs are increasingly used to provide advanced reasoning and communication abilities.⁹
- **Local Views:** Agents typically operate based on incomplete information; they possess only a local view of the overall system state and environment.⁹ This reflects the distributed nature of many real-world problems.
- **Goals:** Agents are usually goal-directed, working towards achieving specific objectives, which may be assigned individually or shared among a group of agents.⁸

3.3 Common Architectures

MAS can be organized according to different architectural patterns:

- **Centralized:** A single central agent or controller coordinates the activities of all other agents.²⁶ This simplifies management but introduces a potential single point of failure and scalability bottleneck.²⁴
- **Decentralized:** Agents interact directly with each other (peer-to-peer) without

any central authority.⁹ This architecture enhances robustness, fault tolerance, and scalability²⁷ but can make coordination more complex.

- **Hierarchical:** This architecture blends centralized and decentralized approaches, often organizing agents into groups or layers with varying levels of control or responsibility.²⁶ It aims to balance coordination efficiency with scalability and autonomy.
- **Other Paradigms:** Frameworks may utilize specialized middleware to manage resource access and coordination⁹, or employ indirect coordination mechanisms like digital pheromones left in the environment.⁹ LLM-based systems are leading to new architectures, such as frameworks that automatically generate and coordinate specialized agents for specific tasks (e.g., AutoAgents²⁵).

3.4 Interaction Paradigms

Interaction is fundamental to MAS functioning. Key paradigms include:

- **Communication:** Agents exchange information explicitly using predefined Agent Communication Languages (ACLs) like FIPA-ACL or KQML, and protocols such as message passing or accessing shared memory spaces (blackboards).⁹ Communication enables sharing knowledge, intentions, and coordinating actions.
- **Coordination:** This involves mechanisms to ensure agents' actions are coherent and aligned towards achieving system goals.⁸ Techniques include negotiation (where agents bargain to reach agreements), distributed planning (agents jointly create plans), synchronization, and establishing social conventions or norms.⁸
- **Collaboration:** Agents actively work together, pooling their resources, knowledge, and capabilities to achieve shared objectives.⁸ This is common in cooperative MAS designed for distributed problem-solving.
- **Competition:** In scenarios where agents have conflicting goals or compete for limited resources, interactions may be competitive.⁸ Game theory is often used to model and analyze such interactions.¹¹

3.5 System Properties

Well-designed MAS often exhibit desirable system-level properties:

- **Self-Organization:** Complex, adaptive, and globally coherent behavior can emerge from the local interactions of agents following relatively simple rules, without explicit top-down control.⁹
- **Adaptability/Flexibility:** MAS can often adapt to dynamic environments and changing requirements by modifying agent behaviors or system organization.¹⁰
- **Robustness/Fault Tolerance:** Due to decentralization and potential redundancy of agents, MAS can often continue functioning even if some individual agents fail.⁹

- **Scalability:** Decentralized and hierarchical architectures, in particular, can often scale to accommodate a large number of agents.¹⁰

3.6 Applications

The characteristics of MAS make them suitable for a wide array of applications, including autonomous driving and traffic management ⁸, multi-robot systems for manufacturing, logistics, or exploration ⁸, automated financial trading ⁸, complex simulations (e.g., social dynamics, epidemics) ⁹, smart grids ²⁷, disaster response coordination ²⁷, and interactive entertainment (computer games).⁸

The prevalent focus within much of modern MAS research lies in designing effective mechanisms for *orchestration*—that is, coordinating the communication, cooperation, and actions of agents that are often assumed to possess significant intelligence or capabilities already.⁸ The primary challenge is often viewed as how to harness and combine these existing capabilities effectively to solve a complex, distributed problem. This contrasts sharply with Minsky's objective, which was not to orchestrate pre-existing intelligence but to understand how intelligence *itself* could emerge from the interactions of fundamentally simple, non-intelligent components.³ While MAS provides powerful tools for building distributed systems, its standard frameworks and assumptions may need significant adaptation to address the foundational question of emergent intelligence posed by the Society of Mind.

4. Bridging Minsky and Modern MAS: A Comparative Analysis

To effectively design a Multi-Agent System inspired by the Society of Mind, it is crucial to first clearly delineate the fundamental differences between the concept of an "agent" as envisioned by Minsky and the "agent" typically encountered in contemporary AI MAS research. This comparison highlights the unique challenges and opportunities inherent in translating Minsky's conceptual framework into a computational reality.

The following table provides a comparative overview across several key dimensions:

Characteristic	Minsky Agent (Society of Mind)	Modern MAS Agent
Complexity	Simple, "Mindless" ³	Often Complex, "Intelligent" ⁹

Autonomy	Low / Reactive (Implied by simplicity)	High / Proactive, Autonomous ⁹
Goal / Purpose	Perform a specific, primitive function ⁶	Achieve pre-defined complex task/goal ⁸
Communication	Implicit / Limited (e.g., Activation/Suppression, K-lines) ¹³	Explicit Protocols (e.g., ACL, KQML) ⁹
Interaction	Often Local, "Exploitative" ¹³	Cooperative / Competitive / Negotiated ⁸
Learning	Emergent / Structural (via K-lines) ¹⁴	Often Explicit / Programmed (e.g., ML/RL) ⁹
Knowledge Representation	Simple / Implicit (e.g., Connections, Frames) ¹²	Often Explicit / Symbolic or Sub-symbolic ¹¹
System Goal	Emergence of Intelligence / Cognition	Distributed Problem Solving / Task Achievement

Analysis of Key Differences:

- Complexity and Mindlessness:** The most striking difference lies in complexity. Minsky deliberately proposed agents that are individually simple and non-thinking.³ The intelligence of the system is hypothesized to emerge *from* this simplicity when organized correctly. In contrast, agents in modern MAS are often designed with significant internal complexity, possessing capabilities for reasoning, planning, and learning.⁹ This fundamental difference dictates the design philosophy: SoM necessitates a bottom-up approach focused on achieving complexity through interaction, whereas MAS often employs top-down decomposition or focuses on coordinating existing complex capabilities.
- Autonomy and Control:** Minsky's simple agents often seem reactive, triggered by inputs from other agents or the environment. While they contribute to the overall system's behavior, their individual autonomy is limited. Modern MAS explicitly emphasizes high degrees of agent autonomy, where agents proactively make decisions and pursue goals.⁹ This difference impacts expectations regarding system control and predictability. A SoM-inspired system might be inherently less predictable at the micro-level but potentially more adaptive

globally.

- **Purpose and Goals:** Minsky agents serve highly specialized, primitive functions (e.g., "recognize edge," "activate grasp").¹⁴ Their "purpose" is to contribute to the functioning of their local agency. Modern MAS agents are typically designed to achieve more complex, often externally defined goals or sub-goals within a larger task.⁸ While both paradigms employ specialization, the *level* of specialization differs significantly. Minsky's specialization occurs at a near-atomic functional level, aiming to construct cognition itself.¹⁵ Modern MAS specialization is often task-oriented within an assumed cognitive framework (e.g., a "planning agent," a "database query agent"²⁵), focusing on efficient problem decomposition rather than the emergence of cognition. Simply using specialized agents does not equate to implementing Minsky's vision; the granularity and purpose of that specialization are critical.
- **Interaction and Communication:** Communication in SoM is often portrayed as implicit and indirect. K-lines, for example, activate entire sets of agents, functioning more like a state-setting mechanism than a message exchange.¹⁴ Agents might simply activate or suppress others¹⁹, or exploit outputs without deep understanding.¹³ Modern MAS, conversely, heavily relies on explicit communication using standardized languages and protocols, enabling complex negotiation, information sharing, and joint planning.⁹

Bridging the Gap:

The core challenge in creating a Minsky-inspired MAS lies in reconciling these differences. How can a system maintain Minsky's emphasis on agent simplicity and the emergence of intelligence from interaction, while leveraging the implementation techniques and scalability potential offered by modern MAS research? Can agents be designed to be computationally simple at the local level, yet collectively produce behavior that is complex, adaptive, and recognizably intelligent at the global level? This requires moving beyond simply coordinating complex agents and instead focusing on designing the *conditions* and *mechanisms* (the "very special ways") that allow intelligence to emerge from the interactions of simpler parts.

5. A Conceptual Framework for a Society of Mind-Inspired MAS

Building upon the principles of Minsky's Society of Mind and acknowledging the distinctions from standard MAS, this section outlines a conceptual framework for an AI system aimed at modeling the *emergence* of intelligence from the interaction of numerous simple, specialized agents. The design philosophy is fundamentally bottom-up, prioritizing agent simplicity and the dynamics of interaction as the source

of complexity and capability.

5.1 Agent Properties (Micro-level)

The foundation of the framework rests on the definition of individual agents:

- **Simplicity:** Agents must possess minimal internal state and processing logic. Potential implementations could include simple finite state machines, perceptron-like units, reactive rule-based systems (e.g., if condition A is met, activate output B), or basic signal processing nodes. They must embody Minsky's concept of "mindlessness," lacking complex internal models, planning capabilities, or symbolic reasoning.³
- **Specialization:** Each agent is dedicated to a highly specific, primitive function. Examples could include detecting a specific low-level feature (e.g., a color, an edge orientation), propagating activation, inhibiting another agent, generating a simple motor output, performing a basic comparison, or contributing to a simple accumulator.¹⁴ Complex cognitive functions are explicitly avoided at the individual agent level.
- **Connectivity:** Agents possess defined, potentially sparse, connections to a limited number of other agents. The output or activation state of an agent directly influences the state or input of its connected neighbors. Connections might be excitatory or inhibitory.
- **Activation State:** Agents possess a state indicating their level of activity, which could be binary (active/inactive) or graded. This activation level determines their influence on connected agents.²²

5.2 Interaction Mechanisms (Meso-level)

Interactions between agents form the crucial intermediate level where complexity begins to build:

- **Activation/Suppression Signals:** The most basic interaction involves active agents sending signals that either increase (excite) or decrease (inhibit) the activation level of connected agents. This mirrors neural processing and aligns with Minsky's mention of suppressors and censors.¹⁷
- **K-Line Analogue Mechanism:** A computational mechanism inspired by K-lines is essential for learning and memory within this framework.¹³ This mechanism must implement three core functions:
 1. **State Capture:** Identifying and "tagging" a configuration of agents that are simultaneously active during a significant event (e.g., achieving a system goal, reducing an error signal, encountering a novel stimulus associated with a positive outcome).

2. **Link Formation/Strengthening:** Creating or reinforcing connections associated with this tagged set of agents. This could manifest as a dedicated "K-agent" connected to the set, or as strengthened synaptic weights between the agents involved, effectively creating a persistent representation of that successful configuration.
 3. **Reactivation:** Allowing the K-line representation (K-agent or strengthened pathway) to be triggered later by relevant cues (e.g., a partial match of the agent state, an external signal associated with the original event). This triggering reactivates the associated agent set, recreating a "partial mental state" relevant to the past event.²¹ This reactivation effectively biases the system towards previously successful modes of operation. Computationally, activating a K-line analogue can be viewed as dynamically reconfiguring the network's effective topology or biasing the flow of activation. It temporarily increases the influence or activation probability of a specific sub-network (agency) that proved useful in the past, thereby setting a "mental context" or operational mode suited for similar situations.¹⁴
- **Resource Competition/Modulation:** Introducing constraints, such as a limited pool of "activation energy" or processing cycles, could lead to competition among agents or agencies. This could provide a mechanism for attentional focus or dynamic priority setting, where more salient or currently successful agencies dominate activity.
 - **Restricted Message Passing:** While avoiding complex symbolic communication like ACL⁹, highly constrained forms of message passing might be permissible. This could involve agents broadcasting simple state flags (e.g., "goal achieved," "error detected") or activation levels to their immediate neighbors, providing minimal contextual information without requiring sophisticated language processing.¹⁸

5.3 Organizational Structures (Macro-level)

The global organization of the agent society shapes the emergence of higher-level functions:

- **Agencies/Sub-Societies:** Agents should naturally group into functional assemblies or "agencies" based on their interactions and shared contributions to specific sub-tasks.¹² These agencies might not be rigidly defined but could be dynamic structures that form, adapt, and dissolve based on context and learning (e.g., through K-line reinforcement).
- **Layered/Hierarchical Structures:** The framework should support the emergence or explicit design of layered processing, reflecting Minsky's

"level-bands".¹² Lower layers could handle sensory processing and fine-grained actions, while higher layers integrate information over longer timescales and represent more abstract goals or plans. The A-brain/B-brain concept could be implemented as distinct layers or agencies for world interaction versus internal monitoring.¹³

- **Emergent Hierarchies:** Ideally, hierarchical control and abstraction should emerge from the interactions and learning processes (particularly the K-line analogue linking agent sets) rather than being entirely pre-programmed. K-lines connecting other K-lines (K-societies) could form the basis of such emergent hierarchical memory structures.¹³
- **Frame Analogues:** The framework should allow for the emergence of structures functionally similar to Minsky's frames.¹² These might arise as stable, strongly interconnected patterns of agent co-activation, potentially linked and activated by K-lines, representing stereotypical knowledge or situational contexts.

This framework provides a conceptual blueprint for building an MAS grounded in Minsky's core ideas. The emphasis is on agent simplicity, localized interactions, and the K-line analogue as a mechanism for structural learning and memory, aiming to create the conditions for intelligence to emerge as a property of the entire system.

6. Exploring Emergent Intelligence: Computational Approaches

Simulating the proposed Society of Mind-inspired MAS framework and investigating the potential emergence of intelligent behavior—Minsky's "very special ways"

⁴—requires appropriate computational methodologies. The chosen approach must handle a large number of simple, interacting agents and facilitate the observation and analysis of emergent, system-level properties. Several existing paradigms offer potential starting points, though likely requiring adaptation.

6.1 Multi-Agent Reinforcement Learning (MARL)

MARL techniques, where multiple agents learn concurrently through interaction with an environment and each other, could potentially be applied.¹¹ Agents within the SoM framework could learn simple policies based on their local observations (states of connected agents, local environmental signals) and reinforcement signals. However, standard MARL faces significant hurdles in this context:

- **Reward Design:** Defining appropriate reward signals is challenging. If rewards are tied only to external task completion, it may not encourage the development of the internal "society" structure Minsky envisioned. Rewards might need to be intrinsic, related to local coordination, information flow, or perhaps linked to the

successful formation and utilization of K-line analogues (reflecting the implicit credit assignment discussed earlier).

- **Credit Assignment:** The multi-agent credit assignment problem—determining which agent's actions contributed to overall success or failure—is notoriously difficult.³⁰ This is exacerbated in a SoM-MAS by the sheer number of agents, their simplicity, and the potentially delayed and indirect nature of emergent outcomes. Novel credit assignment schemes tailored to emergent phenomena would be necessary.
- **Non-Stationarity:** As numerous agents learn and adapt simultaneously, the environment from each agent's perspective is constantly changing, making learning unstable.
- **Potential Adaptations:** Cooperative MARL algorithms³⁶, techniques focusing on emergent communication, or methods incorporating intrinsic motivation might be more suitable than purely competitive or independent learning approaches.

6.2 Evolutionary Algorithms (EAs)

EAs, such as Genetic Algorithms (GAs) and Evolutionary Strategies (ES), appear naturally suited for exploring the design space of a SoM-MAS.³⁷ Evolution could operate at multiple levels:

- **Evolving Agent Rules:** EAs could optimize the simple behavioral rules or parameters of individual agents.
- **Evolving Connections/Topologies:** EAs could search for effective patterns of connectivity between agents or the structure of agencies.
- **Evolving K-Line Mechanisms:** The parameters governing the K-line analogue (e.g., activation thresholds, decay rates, connection strengths) could be evolved.
- **Fitness Evaluation:** Fitness could be based on the system's overall performance on a range of tasks, its adaptability to changing environments, or metrics related to the emergence of desired cognitive functions.
- **Evolutionary Reinforcement Learning (EvoRL):** Combining EAs with RL offers a powerful hybrid approach.³⁹ EAs could evolve aspects of the agent architecture, network structure, or learning rules, while RL could fine-tune the specific behaviors of agents within that evolved framework. This could balance large-scale exploration of the design space with local optimization of behavior.
- **Co-evolution:** Employing co-evolutionary algorithms, where different agent types or agencies evolve in response to each other, could foster diversity and complex interaction dynamics.¹¹

6.3 Complex Adaptive Systems (CAS) / Agent-Based Modeling (ABM)

Viewing the SoM-MAS explicitly as a Complex Adaptive System (CAS) allows the use of Agent-Based Modeling (ABM) techniques for simulation and analysis.⁴¹

- **Simulation Focus:** ABM is primarily a simulation methodology. It involves defining the simple rules governing individual agents and their interactions, then running the simulation to observe the emergent macro-level behaviors of the system as a whole.⁴¹
- **Exploring Emergence:** ABM is ideal for exploring the *conditions* under which intelligence or specific cognitive functions might emerge. Researchers can systematically vary agent rules, interaction mechanisms (including the K-line analogue), network topologies, and environmental conditions to see how these factors influence system-level outcomes.⁴¹
- **Observation over Optimization:** Unlike RL or EAs, the primary goal of ABM here is often understanding and observation rather than direct optimization towards a specific task performance metric. It helps identify the parameters and structures that lead to interesting emergent dynamics.

6.4 Connectionist Approaches / Neural Networks

The SoM framework shares philosophical ground with connectionism and neural networks, where complex functions emerge from the interaction of simple processing units.⁴⁵

- **Agents as Neurons:** Minsky's agents could be implemented as simple nodes (artificial neurons) in a large-scale network. Interactions would be mediated by weighted connections.
- **K-Lines as Learning/Attention:** The K-line analogue could be implemented using Hebbian learning principles (strengthening connections between co-active units) or attention mechanisms that dynamically modulate the influence of different network parts.
- **Dynamic Architectures:** Techniques like self-organizing maps, growing neural gas, or other dynamic/evolvable neural network architectures could be used to allow the "society" structure itself to emerge and adapt over time. Deep learning's success demonstrates the power of emergent representations learned by interconnected simple units.⁴⁶

A successful computational approach for exploring the SoM framework likely requires integrating elements from these paradigms. Minsky's theory involves both the *simulation* of interactions among simple agents (suited to ABM) and a mechanism for *learning and adaptation* based on the outcomes of those interactions (the K-line analogue, potentially implemented or guided by RL/EA principles). Simply simulating

interactions may not lead to adaptive intelligence, while simply applying standard learning algorithms to simple agents might miss the crucial societal and structural aspects of Minsky's theory. Therefore, a hybrid approach is likely necessary: simulating the emergent dynamics of the agent society while incorporating a learning mechanism (like the K-line analogue) that specifically captures, reinforces, and builds upon successful emergent configurations or states.

7. Evaluating Emergence: Metrics and Methodologies

Evaluating the success of a Society of Mind-inspired MAS presents unique challenges. Standard metrics used in AI and MAS, such as task completion rates, accuracy, or efficiency on specific benchmarks, may be insufficient or even misleading.³⁰ The primary goal is not just to solve a task, but to determine whether *intelligence* or complex *cognitive capabilities* are genuinely *emerging* from the bottom-up interactions of simple agents, rather than being implicitly pre-programmed or solely the result of optimization towards a narrow objective. Measuring whether the system is "thinking" in a Minskyan sense requires moving beyond conventional evaluation paradigms.

7.1 Beyond Task Performance Metrics

Evaluation must focus on hallmarks of general intelligence and cognitive function:

- **Adaptability and Generalization:** A key indicator of emergent intelligence is the system's ability to adapt to novel situations, unforeseen changes in the environment, and transfer knowledge or skills learned in one context to new, different domains.¹ Testing should involve tasks significantly different from those encountered during any training or evolutionary process.
- **Robustness and Resilience:** Given the distributed nature of SoM, the system should exhibit graceful degradation rather than catastrophic failure when some agents malfunction or are removed, or when inputs are noisy or incomplete.⁹ Resilience to perturbation is a potential sign of emergent system-level stability.
- **Complexity of Behavior:** Measuring the complexity of the system's emergent behavior over time can be informative. Does the system develop more sophisticated strategies, exhibit more intricate dynamics, or generate more complex outputs as it interacts and learns? Techniques from complex systems theory, information theory (e.g., integrated information³²), or network analysis could be applied to quantify internal dynamics or behavioral richness.
- **Emergence of Hierarchical Abstraction:** Evidence that the system is forming higher-level representations, control structures, or functional agencies from lower-level interactions would support the SoM model.¹² This might be assessed

by analyzing internal structures (e.g., K-line connectivity patterns) or designing tasks that require abstract reasoning or planning.

7.2 Adapting MAS Evaluation Metrics

While standard MAS metrics need careful interpretation, some categories can be adapted:

- **Emergent Collaboration/Coordination:** Metrics like decision synchronization³² or communication efficiency³² could be used, but the focus should be on *spontaneous* or *learned* coordination arising from agent interactions, not adherence to pre-defined protocols. How effectively does the "society" self-organize to achieve collective action?
- **Self-Organized Resource Utilization:** Efficiency metrics³² can be relevant if interpreted as measures of how well the system self-organizes to manage internal resources (e.g., activation, computation) or external resources in the environment.
- **Cognitively Plausible Output Quality:** Outputs should be assessed not just for factual accuracy but also for coherence, consistency, and potentially "common sense" qualities³², reflecting higher cognitive functions. Does the system exhibit frame-like understanding or reasoning?
- **Scalability of Emergence:** Evaluating how emergent intelligent behaviors scale as the number of agents increases is crucial.³⁰ Does intelligence increase, plateau, or collapse due to overwhelming complexity?

7.3 Methodologies for Assessing Emergence

Directly measuring emergent intelligence requires specific methodologies:

- **Longitudinal Behavioral Analysis:** Observing the system's behavior across a wide range of complex, dynamic environments over extended periods is essential. Qualitative analysis looking for significant shifts in capabilities, strategy changes, or the appearance of novel behaviors is needed.
- **Internal State Analysis and Visualization:** Developing tools to probe and visualize the internal dynamics of the agent society is critical. This could involve tracking agent activation patterns, analyzing the evolving network structure (including K-line analogue connections), and identifying emergent functional clusters (agencies).
- **AI Psychometrics:** Applying principles and tests from psychometrics offers a structured way to assess AI systems for psychological traits or cognitive capabilities.⁴⁹ Standardized cognitive tests (e.g., for reasoning, problem-solving, planning), adapted for the AI context, could potentially reveal emergent abilities. However, the validity and interpretation of such tests for non-human systems

require careful consideration and validation.⁴⁹

- **Comparative Analysis:** Benchmarking the system's behavior against human or animal performance on analogous cognitive tasks can provide a relative measure of emergent capability.
- **Targeted Cognitive Probes:** Designing specific experimental setups to test for the emergence of particular cognitive phenomena discussed by Minsky, such as frame-based reasoning (e.g., handling default assumptions and exceptions), commonsense inference, or self-monitoring capabilities (A/B brain analogue).

Minsky's assertion that intelligence emerges when agents are joined in "certain very special ways" ⁴ implies that evaluation should not be solely outcome-focused. It must also assess the *process* and *structure* of the emergent organization. Metrics are needed to characterize the quality of self-organization, the effectiveness of the K-line analogue in capturing and reusing useful system states, and the complexity and nature of the emergent agencies and hierarchies. Validating Minsky's theory computationally requires measuring *how* the system organizes itself, not just *what* it can achieve.

8. Related Work and Inspirations

While direct, large-scale computational implementations of the Society of Mind theory remain scarce, likely due to its conceptual nature and the inherent difficulty ¹⁷, Minsky's ideas have permeated and inspired various threads of AI and cognitive science research. Understanding these connections provides context for the proposed framework.

8.1 Direct Implementations and Conceptual Influence

Few projects have attempted a direct translation of SoM into code. One documented example involves applying the theory to the domain of Shogi (Japanese chess), outlining how agencies for recognition, lookahead, and learning might be constructed from simpler agents.¹⁷ However, the primary influence of SoM has been conceptual.⁷ It provided a powerful metaphor for thinking about intelligence as a distributed, emergent phenomenon arising from simpler components, influencing fields like agent-based AI, distributed AI, and research emphasizing emergence.⁷ The genesis of the theory itself was rooted in Minsky and Papert's practical work on a block-stacking robot in the late 1960s and early 1970s, where the need for diverse, interacting processes became apparent.³ Minsky's work on "frames" as knowledge structures also had a significant impact on knowledge representation and expert systems.¹²

8.2 Bottom-Up AI Approaches

The SoM philosophy aligns strongly with bottom-up approaches in AI, where complex global behavior arises from simple local rules and interactions:

- **Cellular Automata (CA):** Early work like Conway's Game of Life demonstrated how complex patterns could emerge from simple rules on a grid.⁴⁴ CAs share the principle of local interaction leading to global complexity.
- **Swarm Intelligence:** Fields like Ant Colony Optimization and Particle Swarm Optimization derive complex problem-solving capabilities from populations of simple agents following basic rules, often inspired by social insects or flocking behavior.²⁴ This resonates with the emergence of collective function from simple parts.
- **Artificial Life (ALife):** ALife research explicitly studies emergent phenomena, adaptation, and evolution in simulated systems, often using agent-based approaches. It shares SoM's interest in how complex, life-like behaviors can arise from simpler foundations.

8.3 Cognitive Architectures

Cognitive architectures aim to provide unified theories and computational models of the fixed structures and processes underlying human (or artificial) cognition.⁵¹

- **Symbolic Architectures (ACT-R, SOAR):** Architectures like ACT-R⁴⁵ and SOAR⁴⁵ typically rely on symbolic representations (e.g., production rules, declarative chunks) and pre-defined modules (e.g., for perception, memory, motor control).⁴⁵ While powerful for modeling specific cognitive tasks, they differ significantly from SoM's emphasis on emergence from non-symbolic, mindless agents and its less rigidly structured organization.⁴⁵ These architectures often face limitations in handling the sheer size and heterogeneity of human common-sense knowledge, relying on limited, task-specific knowledge bases.⁵² SoM, through its emergent approach, implicitly aims to circumvent these "knowledge level" limitations.
- **Emergent/Connectionist Architectures:** These architectures, inspired by neural networks, align more closely with SoM's bottom-up philosophy.⁴⁵ They emphasize learning and adaptation through networks of simple processing units, where cognitive capabilities emerge from interaction and experience rather than explicit programming.⁴⁵ The success of deep learning underscores the potential for complex functions to emerge from interconnected simple units⁴⁶, although bridging the gap to high-level, structured reasoning remains a challenge.⁵⁴
- **Hybrid Architectures & CoALA:** Some research explores hybrid architectures combining symbolic and connectionist elements.⁵⁴ Recent frameworks like CoALA (Cognitive Architectures for Language Agents) attempt to apply cognitive architecture principles (like working memory, long-term memory, action selection)

to structure agents based on LLMs⁵⁵, indicating continued interest in finding principled ways to organize complex AI systems, even if the underlying components (LLMs) are vastly different from Minsky's simple agents.

8.4 Theory of Mind (ToM) in MAS

Research incorporating Theory of Mind concepts from psychology into MAS focuses on enabling agents to model the mental states (beliefs, desires, intentions) of other agents (including humans).⁵⁶ This aims to improve communication, coordination, and collaboration in mixed human-agent or complex agent societies.⁵⁶ While distinct from SoM's focus on the internal society *constituting* a single mind, ToM research shares the theme of agents reasoning about other agents, which is relevant if a SoM-MAS were to develop higher-order capabilities involving modeling its own internal agencies or interacting with external agents.

Considering these related fields suggests a potential convergence. Minsky's vision of intelligence emerging from vast numbers of simple, interacting agents³ finds echoes in connectionism's success with simple units (neurons)⁴⁶ and the capabilities shown by modern LLM-based MAS using coordinated specialists.²⁵ Could a viable path involve constructing a SoM-inspired system using massively interconnected, extremely simple processing units (perhaps simpler than current neurons, or specialized micro-LLMs), where K-line analogues provide the crucial mechanisms for learning, memory, and context-setting? Such an approach might achieve the emergent cognitive functions currently pursued through large monolithic models or explicitly structured symbolic architectures, potentially offering a more decentralized, robust, and perhaps ultimately more scalable route towards AGI.

9. Challenges and Future Directions

Developing a functional MAS based on Minsky's Society of Mind framework, while conceptually appealing, faces substantial theoretical and practical challenges. These hurdles span computational feasibility, algorithmic design, knowledge representation, evaluation, and the fundamental difficulty of bridging the gap from simple interactions to high-level cognition.

9.1 Scalability

- **Computational Demands:** Minsky envisioned a mind composed of potentially vast numbers of agents.³ Simulating the interactions of potentially billions of simple agents, along with the dynamic creation and activation of K-line analogues, poses immense computational challenges.³⁰ The management of connections and state updates could quickly become intractable.

- **Interaction Overhead:** Even if individual agent computations are simple, the sheer volume of interactions in a large-scale society could lead to communication bottlenecks or prohibitive simulation times.³⁰ Minsky suggested that specialization into agencies might mitigate this by reducing overall interconnectedness¹⁸, but whether this scales sufficiently remains an open question.

9.2 Credit Assignment

As highlighted previously (Insight 2.1), determining which specific agents, interactions, or K-line formations contributed to an emergent success (or failure) is a core difficulty.³⁰ Standard RL credit assignment methods are often ill-suited for systems with numerous simple agents, delayed rewards, and emergent goals. The implicit credit assignment via K-lines needs to be computationally realized in a way that effectively reinforces useful societal configurations without explicit, global reward signals.

9.3 Designing Interaction Protocols ("Special Ways")

The effectiveness of the SoM model hinges on agents being joined in "certain very special ways".⁴ Discovering or designing these optimal interaction rules and organizational principles is non-trivial. It requires finding the right balance between agent simplicity (to encourage emergence) and sufficient interaction capability (to allow complex behaviors to arise). This likely necessitates extensive experimentation, possibly guided by evolutionary algorithms or large-scale ABM simulations, to explore the space of possible rules and structures. Defining a robust and effective computational K-line mechanism is particularly critical.

9.4 Knowledge Representation and Grounding

- **Emergence of Concepts:** A fundamental challenge is understanding how symbolic-like concepts, structured knowledge (akin to Minsky's Frames¹²), and abstract reasoning can emerge purely from the interactions of simple, sub-symbolic agents. This is a version of the classic symbol grounding problem.¹ How does the society transition from low-level processing to meaningful conceptual representation?
- **Knowledge Scope:** While potentially avoiding the homogeneity limitations of symbolic CAs⁵² by being emergent, it's unclear how a SoM-MAS could acquire, represent, and effectively utilize the vast, diverse, and nuanced common-sense knowledge possessed by humans. Is the proposed structure of agents, agencies, and K-lines sufficient for this scale and heterogeneity of knowledge?

9.5 Emergence versus Control

There exists an inherent tension between allowing behaviors to emerge freely from the bottom up and guiding the system towards specific intelligent capabilities or goals. Overly constraining the system with top-down objectives might stifle the very emergent processes the framework aims to capture. Finding mechanisms to gently guide or shape emergence without destroying it is a delicate balancing act.

9.6 Evaluation Difficulties

As detailed in Section 7, objectively measuring emergent intelligence and distinguishing it from complex but non-cognitive behavior is extremely difficult.³⁰ The lack of clear, agreed-upon metrics for emergent cognition hinders empirical progress and comparative evaluation.

9.7 Bridging the Gap to High-Level Cognition

Perhaps the most significant challenge is demonstrating convincingly how the collective activity of extremely simple agents can scale to produce the full spectrum of human cognition, including abstract thought, language, consciousness, self-awareness, and complex emotions, as Minsky suggested.³ Currently, this leap remains largely theoretical and speculative.²⁰ The mechanisms proposed (agents, agencies, K-lines) must be shown to be sufficient for this profound level of emergence. This echoes historical challenges faced by connectionism; Minsky himself co-authored "Perceptrons," which highlighted limitations of early neural networks.⁵⁰ While SoM offers different organizational principles, the underlying difficulty of bridging distributed, low-level processing to high-level, structured, symbolic-like cognition persists, a challenge even for modern deep learning.⁵⁴ The "special ways" need to be powerful enough to overcome these deep, historical hurdles in AI.

9.8 Technical and Ethical Risks

Standard MAS risks, such as agent malfunctions propagating through the system, unpredictable emergent behaviors leading to unintended consequences, and security vulnerabilities, apply equally to a SoM-inspired system.⁵⁹ The highly decentralized and emergent nature might make debugging and ensuring safety particularly difficult. Ethical considerations regarding autonomous decision-making and accountability also arise if such systems achieve significant capabilities.⁵⁹

9.9 Future Directions

Despite the challenges, pursuing research based on SoM principles offers intriguing

possibilities. Key future directions include:

- Developing and testing robust, scalable computational models of K-lines, agencies, and other core SoM mechanisms.
- Conducting large-scale simulations using ABM and EA approaches to explore the parameter space and identify conditions conducive to emergence.
- Designing and validating novel metrics and methodologies specifically for evaluating emergent cognitive capabilities.
- Investigating hybrid architectures that combine SoM principles (e.g., agent simplicity, K-lines) with techniques from neural networks, MARL, or cognitive architectures.
- Focusing research on specific cognitive domains (e.g., commonsense reasoning, language acquisition, problem-solving) as challenging testbeds for emergent intelligence.
- Fostering interdisciplinary collaboration between AI, cognitive science, neuroscience, and complex systems researchers.

10. Conclusion

Marvin Minsky's Society of Mind theory offers a provocative and enduring vision of intelligence as an emergent property of a vast society of simple, interacting agents. This report has outlined a conceptual framework for constructing an AI Multi-Agent System based on these principles, emphasizing agent simplicity, specialized functions, localized interactions, and a K-line-inspired mechanism for learning and memory through dynamic network reconfiguration. This framework stands in contrast to many contemporary MAS approaches that focus on orchestrating the capabilities of already complex agents.

The potential of a SoM-inspired approach lies in its offer of an alternative path towards understanding and potentially achieving artificial general intelligence—one rooted in emergence, diversity, and bottom-up self-organization. Such systems might exhibit greater robustness, adaptability, and perhaps a closer resemblance to natural intelligence than monolithic models. Exploring this framework could yield valuable insights into the fundamental nature of cognition, learning, and memory, irrespective of whether it ultimately leads to AGI.

However, the path forward is fraught with significant challenges. Operationalizing Minsky's often abstract concepts into concrete computational mechanisms, particularly the K-line analogue, is a primary hurdle. Ensuring scalability, managing the complexity of interactions, solving the deep credit assignment problem for emergent behaviors, and developing adequate evaluation methodologies are critical research

frontiers. The most profound challenge remains bridging the vast gap between the interactions of simple, mindless agents and the richness of high-level human cognition, a leap that remains largely theoretical.

Pursuing Minsky-inspired MAS research requires a long-term perspective and an interdisciplinary approach, drawing on insights from AI, cognitive science, complex systems theory, and computational neuroscience. While the practical realization of Minsky's full vision is uncertain, the exploration itself promises to deepen our understanding of intelligence and may uncover novel principles for designing more flexible, adaptive, and robust artificial systems. The "very special ways" agents might interact to produce intelligence remain a compelling area for investigation.

Works cited

1. Artificial general intelligence - Wikipedia, accessed April 14, 2025, https://en.wikipedia.org/wiki/Artificial_general_intelligence
2. www.ebsco.com, accessed April 14, 2025, <https://www.ebsco.com/research-starters/literature-and-writing/society-mind-marvin-minsky#:~:text=%22The%20Society%20of%20Mind%2C%22collectively%20handle%20tasks%20related%20to>
3. Society of Mind - Wikipedia, accessed April 14, 2025, https://en.wikipedia.org/wiki/Society_of_Mind
4. The Society of Mind: Minsky, Marvin: 9780671657130 - Amazon.com, accessed April 14, 2025, <https://www.amazon.com/Society-Mind-Marvin-Minsky/dp/0671657135>
5. Marvin minsky's timeless lessons on AI and collective intelligence - Capgemini, accessed April 14, 2025, <https://www.capgemini.com/insights/expert-perspectives/marvin-minskys-timeless-lessons-on-ai-and-collective-intelligence/>
6. The Society of Mind: Marvin Minsky - Books - Amazon.com, accessed April 14, 2025, <https://www.amazon.com/Society-Mind-Marvin-Minsky/dp/0434467588>
7. Marvin Minsky and His Pioneering Discoveries in Multi-Agent ..., accessed April 14, 2025, <https://blog.sparkengine.ai/posts/marvin-minsky-multi-agent-systems>
8. Multi-agent systems | The Alan Turing Institute, accessed April 14, 2025, <https://www.turing.ac.uk/research/interest-groups/multi-agent-systems>
9. Multi-agent system - Wikipedia, accessed April 14, 2025, https://en.wikipedia.org/wiki/Multi-agent_system
10. Multi-agent system: Types, working, applications and benefits - LeewayHertz, accessed April 14, 2025, <https://www.leewayhertz.com/multi-agent-system/>
11. AAMAS 2018 - IFAAMAS, accessed April 14, 2025, <https://www.ifaamas.org/AAMAS/aamas2018/callForPapers.html>
12. The Society of Mind by Marvin Minsky | EBSCO Research Starters, accessed April 14, 2025, <https://www.ebsco.com/research-starters/literature-and-writing/society-mind-m>

[arvin-minsky](#)

13. Marvin Minsky on Society of Minds - Philosophy Dictionary of Arguments, accessed April 14, 2025, https://philosophy-science-humanities-controversies.com/listview-details.php?id=2539860&a=a&first_name=Marvin&author=Minsky&concept=Society%20of%20Minds
14. Examining the Society of Mind, accessed April 14, 2025, <http://www.jfsowa.com/ikl/Singh03.htm>
15. The Society of Mind | work by Minsky - Britannica, accessed April 14, 2025, <https://www.britannica.com/topic/The-Society-of-Mind>
16. Society Of Mind - msg, accessed April 14, 2025, <https://www.msg.group/msg-wissen/curated-resources/society-of-mind>
17. The Society of Shogi - A research agenda - - Teu, accessed April 14, 2025, <https://www2.teu.ac.jp/gamelab/RESEARCH/SIG-GI-16.pdf>
18. Society of Mind Project - DTIC, accessed April 14, 2025, <https://apps.dtic.mil/sti/tr/pdf/ADA200313.pdf>
19. Examining the Society of Mind - ResearchGate, accessed April 14, 2025, https://www.researchgate.net/publication/2909614_Examining_the_Society_of_Mind
20. K-line (artificial intelligence) - Wikipedia, accessed April 14, 2025, [https://en.wikipedia.org/wiki/K-line_\(artificial_intelligence\)](https://en.wikipedia.org/wiki/K-line_(artificial_intelligence))
21. K-Lines: A Theory of Memory. - DTIC, accessed April 14, 2025, <https://apps.dtic.mil/sti/tr/pdf/ADA078116.pdf>
22. courses.csail.mit.edu, accessed April 14, 2025, <https://courses.csail.mit.edu/6.803/pdf/aim-516.pdf>
23. The Society of Mind | Electrical Engineering and Computer Science ..., accessed April 14, 2025, <https://ocw.mit.edu/courses/6-868j-the-society-of-mind-fall-2011/>
24. AI in Multi-Agent Systems: How AI Agents Interact & Collaborate, accessed April 14, 2025, <https://focalx.ai/ai/ai-multi-agent-systems/>
25. AutoAgents: A Framework for Automatic Agent Generation - IJCAI, accessed April 14, 2025, <https://www.ijcai.org/proceedings/2024/3>
26. What Are Multi-agent AI Systems? - SmythOS, accessed April 14, 2025, <https://smythos.com/ai-agents/multi-agent-systems/multi-agent-ai-systems/>
27. What is a Multi Agent System - Relevance AI, accessed April 14, 2025, <https://relevanceai.com/learn/what-is-a-multi-agent-system>
28. What is Agentic AI? Definition, Examples and Trends in 2025 - Aisera, accessed April 14, 2025, <https://aisera.com/blog/agentic-ai/>
29. Multi-Agent Coordination across Diverse Applications: A Survey - arXiv, accessed April 14, 2025, <https://arxiv.org/html/2502.14743v1>
30. Understanding Multiagent Systems: How AI Systems ... - Encord, accessed April 14, 2025, <https://encord.com/blog/multiagent-systems/>
31. Keynote Speakers - AAMAS 2025 Detroit, accessed April 14, 2025, <https://aamas2025.org/index.php/conference/program/keynote-speakers/>
32. A Comprehensive Guide to Evaluating Multi-Agent LLM Systems ..., accessed April 14, 2025, <https://orq.ai/blog/multi-agent-llm-eval-system>

33. Mastering Multi-Agent Eval Systems in 2025 - Botpress, accessed April 14, 2025, <https://botpress.com/blog/multi-agent-evaluation-systems>
34. Benchmarks and Use Cases for Multi-Agent AI - Galileo AI, accessed April 14, 2025, <https://www.galileo.ai/blog/benchmarks-multi-agent-ai>
35. www.cs.cmu.edu, accessed April 14, 2025, <https://www.cs.cmu.edu/~mmv/papers/MASsurvey.pdf>
36. A Comprehensive Survey on Multi-Agent Cooperative Decision-Making: Scenarios, Approaches, Challenges and Perspectives - arXiv, accessed April 14, 2025, <https://arxiv.org/html/2503.13415v1>
37. ALA 2023, accessed April 14, 2025, <https://alaworkshop2023.github.io/>
38. Tutorials – AAMAS 2025 Detroit, accessed April 14, 2025, <https://aamas2025.org/index.php/conference/program/tutorials/>
39. Evolutionary Reinforcement Learning: A Systematic Review and Future Directions - MDPI, accessed April 14, 2025, <https://www.mdpi.com/2227-7390/13/5/833>
40. Evolutionary Reinforcement Learning: A Systematic Review and Future Directions - arXiv, accessed April 14, 2025, <https://arxiv.org/html/2402.13296v1>
41. Agent-based modeling - (Cognitive Psychology) - Vocab, Definition, Explanations | Fiveable, accessed April 14, 2025, <https://fiveable.me/key-terms/cognitive-psychology/agent-based-modeling>
42. faculty.sites.iastate.edu, accessed April 14, 2025, <https://faculty.sites.iastate.edu/tesfatsi/archive/tesfatsi/ABMTutorial.MacalNorth.JOS2010.pdf>
43. Agent-Based Modeling in Psychology - SmythOS, accessed April 14, 2025, <https://smythos.com/ai-industry-solutions/healthcare/agent-based-modeling-in-psychology/>
44. Agent-based model - Wikipedia, accessed April 14, 2025, https://en.wikipedia.org/wiki/Agent-based_model
45. Cognitive Agent Architectures: Revolutionizing AI with ... - SmythOS, accessed April 14, 2025, <https://smythos.com/ai-agents/agent-architectures/cognitive-agent-architectures/>
46. The Emergence of Intelligence as a Natural Phenomenon: An Interdisciplinary Review, accessed April 14, 2025, <https://stevenmilanese.com/the-emergence-of-intelligence-as-a-natural-phenomenon-an-interdisciplinary-review/>
47. Human Cognitive Architecture as an Intelligent Natural Information Processing System, accessed April 14, 2025, <https://www.mdpi.com/2076-328X/15/3/332>
48. Measuring Agent Effectiveness in Multi-Agent Workflows - Galileo AI, accessed April 14, 2025, <https://www.galileo.ai/blog/analyze-multi-agent-workflows>
49. Evaluating the Psychological Reasoning of Large ... - ScholarSpace, accessed April 14, 2025, <https://scholarspace.manoa.hawaii.edu/bitstreams/7e9c1382-9efc-45d3-b859-e6a52a95ed4e/download>
50. AI Origins: Marvin Minsky - Datategy, accessed April 14, 2025, <https://www.datategy.net/2024/07/22/ai-origins-marvin-minsky/>

51. Cognitive architectures: Research issues and challenges - Electrical Engineering and Computer Science, accessed April 14, 2025, <https://web.eecs.umich.edu/~soar/sitemaker/docs/pubs/cogarch.cogsys08.pdf>
52. Representational Limits in Cognitive Architectures - CEUR-WS, accessed April 14, 2025, https://ceur-ws.org/Vol-1855/EUCognition_2016_Part4.pdf
53. philarchive.org, accessed April 14, 2025, <https://philarchive.org/archive/LIETKL>
54. Bridging Generative Networks with the Common Model of Cognition - arXiv, accessed April 14, 2025, <https://arxiv.org/html/2403.18827v1>
55. Cognitive Architectures for Language Agents - arXiv, accessed April 14, 2025, <http://arxiv.org/pdf/2309.02427>
56. Distributed Theory of Mind in Multi-Agent Systems - SciTePress, accessed April 14, 2025, <https://www.scitepress.org/Papers/2024/125634/125634.pdf>
57. Applying Theory of Mind to Multi-Agent Systems: A Systematic Review - ResearchGate, accessed April 14, 2025, https://www.researchgate.net/publication/374199650_Applying_Theory_of_Mind_to_Multi-Agent_Systems_A_Systematic_Review
58. LLM Multi-Agent Systems: Challenges and Open Problems - arXiv, accessed April 14, 2025, <https://arxiv.org/html/2402.03578v1>
59. Challenges in Multi-Agent Systems: Navigating Complexity in Distributed AI - SmythOS, accessed April 14, 2025, <https://smythos.com/ai-agents/multi-agent-systems/challenges-in-multi-agent-systems/>
60. Everything you need to know about multi AI agents in 2025: explanation, examples and challenges - Springs, accessed April 14, 2025, <https://springsapps.com/knowledge/everything-you-need-to-know-about-multi-ai-agents-in-2024-explanation-examples-and-challenges>
61. What are the risks and benefits of 'AI agents'? - The World Economic Forum, accessed April 14, 2025, <https://www.weforum.org/stories/2024/12/ai-agents-risks-artificial-intelligence/>