

**Position Paper**

**Emergent Machina Sapiens Urge Rethinking Multi-Agent Paradigms in Critical Infrastructures**

Li, Hepeng; Liu, Yuhong; Yan, Jun; Gao, Jie; Yang, Xiao'ou; Naili, Mohamed

**DOI**

[10.1109/IJCNN64981.2025.11228849](https://doi.org/10.1109/IJCNN64981.2025.11228849)

**Publication date**

2025

**Document Version**

Final published version

**Published in**

International Joint Conference on Neural Networks, IJCNN 2025 - Proceedings

**Citation (APA)**

Li, H., Liu, Y., Yan, J., Gao, J., Yang, X., & Naili, M. (2025). Position Paper: Emergent Machina Sapiens Urge Rethinking Multi-Agent Paradigms in Critical Infrastructures. In *International Joint Conference on Neural Networks, IJCNN 2025 - Proceedings* (Proceedings of the International Joint Conference on Neural Networks). IEEE. <https://doi.org/10.1109/IJCNN64981.2025.11228849>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

**Green Open Access added to [TU Delft Institutional Repository](#)  
as part of the Taverne amendment.**

More information about this copyright law amendment  
can be found at <https://www.openaccess.nl>.

Otherwise as indicated in the copyright section:  
the publisher is the copyright holder of this work and the  
author uses the Dutch legislation to make this work public.

# Position Paper: Emergent Machina Sapiens Urge Rethinking Multi-Agent Paradigms in Critical Infrastructures

Hepeng Li\*, Yuhong Liu<sup>†</sup>, Jun Yan<sup>‡</sup>, Jie Gao<sup>§</sup>, Xiao'ou Yang<sup>†</sup>, Mohamed Naili<sup>†</sup>

\*Department of Electrical & Computer Engineering, University of Maine, Orono, ME, USA

<sup>†</sup>Department of Computer Science & Engineering, Santa Clara University, Santa Clara, CA, USA

<sup>‡</sup>Concordia Institute for Information Systems Engineering, Concordia University, Montréal, QC, Canada

<sup>§</sup>Department of Transport & Planning, Delft University of Technology, Delft, Netherlands

Email: hepeng.li@maine.edu; yhliu@scu.edu; jun.yan@concordia.ca; j.gao-1@tudelft.nl; xyang13@scu.edu

**Abstract**—Artificial Intelligence (AI) agents capable of autonomous learning and independent decision-making hold great promise for addressing complex challenges across various critical infrastructure domains, including transportation, energy systems, and manufacturing. However, the surge in the design and deployment of AI systems, driven by various stakeholders with distinct and unaligned objectives, introduces a crucial challenge: How can uncoordinated AI systems coexist and evolve harmoniously in shared environments without creating chaos or compromising safety? To address this, we advocate for a fundamental rethinking of existing multi-agent frameworks, such as multi-agent systems and game theory, which are largely limited to predefined rules and static objective structures. We posit that AI agents should be empowered to adjust their objectives dynamically, make compromises, form coalitions, and safely compete or cooperate through evolving relationships and social feedback. Through two case studies in critical infrastructure applications, we call for a shift toward the emergent, self-organizing, and context-aware nature of these multi-agentic AI systems.

**Index Terms**—Multi-Agent Systems, critical infrastructures, artificial intelligence, Machina Sapiens, agentic AI.

## I. INTRODUCTION

The advancement in autonomous AI systems, such as deep reinforcement learning [1], [2], agentic artificial intelligence [3], [4], self-supervised learning [5], [6], meta-learning [7], [8], etc., has ushered in a new era of machina sapiens, where machines transcend their traditional roles to become independent learners, problem-solvers, and human assistants and collaborators. The emergence of these autonomous systems promises transformative capabilities in addressing complex tasks across various critical infrastructure domains [9], such as smart transportation [10], smart grid and energy systems [11], [12], and robotics and smart manufacturing [13]–[16], among others.

With the enormous potential of AI adoptions, we envision an ecosystem where different types of autonomous systems, designed and deployed independently by various stakeholders

This work is supported in part by the Concordia University Research Chair program, the Natural Sciences & Engineering Research Council (NSERC) of Canada under grants RGPIN-2018-06724 and RGPIN-2025-05097, and the U.S. National Science Foundation under the grant number 2330504.

with distinct goals, will interact, evolve, and coexist as *agents* [17]. Autonomous cars from multiple manufacturers will share the same roads, while intelligent energy management systems deployed by utilities and prosumers will co-manage the power grid. Additionally, robotic assistants customized for diverse user needs will collaborate in shared spaces, such as manufacturing facilities, hospitals, warehouses, and homes. However, in such environments, agents may find it challenging to achieve their goals alongside other AI agents if they adhere solely to the rules and objectives as designed. The underlying conflicts, competition, and misalignment in the goals and behaviors could lead to failures at the societal level, which in turn undermine the success of individual agents, causing chaotic disruptions that compromise the safety, reliability, and ethical integrity of real-world AI applications.

The interconnected agentic ecosystems raise critical questions: How can AI agents with different goals and architectures coevolve and collaborate in shared environments? How can they adapt to unforeseen situations without jeopardizing system-wide safety and performance? These challenges extend beyond technical coordination. Autonomous systems, despite being designed and deployed independently, must nonetheless learn to cooperate, negotiate trade-offs, and adhere to societal rules that balance individual objectives with collective well-being. For instance, should a delivery robot prioritize speed over safety when sharing sidewalks with humans? How should an autonomous car handle conflicting interests between its passengers and other road users? As the diversity and ubiquity of AI systems grow, these questions become increasingly urgent.

Traditional frameworks such as multi-agent reinforcement learning [18]–[22] and game theoretical methods [23]–[26] have been widely used to model interactions among agents, but they often fall short when applied to such open environments [27] for several reasons:

- *Static Objectives and Equilibria*: Existing approaches often assume AI agents optimize with fixed utility or reward structures toward static equilibria. These assumptions do not hold in systems where contexts shift unpredictably and common

interests change over time, requiring agents to be capable of dynamically adjusting their goals post deployment.

- *Predefined Rules for Interactions:* Conventional frameworks typically model AI interactions based on predefined relationships. However, these interactions may need to dynamically shift between cooperation and competition, depending on the evolving contexts and objectives.

- *Pre-Designed Coordination Mechanisms:* Most learning frameworks rely on the meticulous design of coordination algorithms in advance to ensure consensus or convergence, which is impractical in an ecosystem where AI agents are designed and deployed independently.

- *Curse of Scalability:* Coordinating and aligning the behaviors of numerous AI agents demands resource-intensive setups, leading to exponential increases in the complexity of computation and interaction. This poses significant challenges for traditional paradigms in large-scale applications.

These limitations are fundamentally rooted in paradigms that prioritize the design of multi-agent interactions for coordination rather than permitting interactions to emerge organically among AI agents. The advent of machine intelligence presents a transformative opportunity to transcend these constraints with innovative approaches to managing the dynamic, emergent, and socially embedded complexities in the emergent AI ecosystems.

**Our Position:** We posit that the growing trend of agentic AI development demands a fundamental rethinking of how we conceptualize interoperability among an influx of AI agents. Instead of relying on pre-engineered rules and static reward structures, agents should possess the flexibility to adapt to the diverse and evolving world. This flexibility involves the capacity to align their individual goals with broader systems to balance their objectives with collective considerations, such as safety, fairness, and ethical standards. By fostering such autonomy, agents can not only coexist but also co-thrive in open-ended environments, which would otherwise challenge existing paradigms.

Realizing this vision entails empowering agents to adapt their learning mechanisms, negotiate trade-offs, form coalitions, and engage safely in cooperation or competition. AI agents must be capable of proactively collecting feedback and critics by building trust in dynamic relationships and interactions with peers. This paradigm prioritizes adaptability, harmony, and resilience, shifting from pre-engineered frameworks to self-organizing and context-aware systems. To this end, we advocate for the development of novel methodologies and design principles for multi-agentic ecosystems that foster cooperation among AIs and between AI and humans, ensuring that the benefits of AI align with societal values.

Our position aims to contribute by:

- **Inspiring Future Research Directions:** We highlight the opportunities and challenges posed by the increasing prevalence of autonomous, diverse, and independently developed AI agents in shared environments. We encourage the research community to explore innovative methodologies by reexamining existing paradigms while addressing the complexities of emergent AI ecosystems.

- **Proposing a Framework for Dynamic Interaction:** We advocate for a new framework that integrates dynamic norms and adaptive protocols as fundamental components for governing multi-agent interactions. This framework equips agents to continuously evolve their behaviors based on feedback, fostering alignment, coordination, and resilience in diverse, open-ended environments.

- **Grounding Challenges in Real-World Contexts:** We situate the theoretical challenges of emergent agentic ecosystems within practical scenarios. This contextualization underscores the necessity of adaptive and dynamic approaches while demonstrating how the proposed framework can accommodate real-world complexities.

## II. THE ECOLOGY OF AN AGENTIC AI ECOSYSTEM

### A. *The Nature and Nurture of Autonomous Agents*

Analyzing how autonomous agents develop and perform in multi-agent settings is helpful in distinguishing between nature and nurture. Although these terms may evoke biological analogies, we use them here to articulate two complementary facets of agent behavior: nature is relatively fixed at inception, and nurture evolves as circumstances change.

From a **nature** perspective, each agent begins with an intrinsic capacity to:

- Evolve greedily in a goal-driven manner
- Learn autonomously from interactive feedback
- Optimize quantitatively with precise metrics

These built-in qualities establish a stable foundation: agents know what they are striving for, have independent means of adapting their strategies, and can precisely measure outcomes relevant to their predefined goals.

Yet, as no environment remains entirely static, agents must also embody a **nurture** component, where

- Goal may shift by contextual and dynamic factors
- Learning can be influenced by peer agents or humans
- Metrics may be uncertain or incompletely defined

Under these nurture influences, the pursuit of a fixed objective becomes a nuanced negotiation of evolving factors.

This duality sheds light on how agents develop over time. They begin with certain high-level design choices, learning algorithms, and objective functions that guide their initial behavior. However, when confronted with real-world change, they adapt: goals may be reassessed; unanticipated signals can refocus their decisions; and newly emerging metrics can upend earlier assumptions about optimal performance. In short, nature ensures that agents have consistent, goal-driven architectures, while nurture ensures they do not stagnate but continually revise their internal parameters in light of external developments.

When many such agents, each guided by distinct but similarly structured nature-and-nurture dynamics, coexist in a shared environment, interactions can become unexpectedly intricate. Only by first recognizing how an agent's built-in qualities (nature) intersect with its adaptive growth (nurture) can we fully grasp why multi-agent ecosystems exhibit both boundless potential and inherent challenges.



## B. Connecting Norms, Agents, and the World

**Evolution of Norms.** Norms are the backbone of the proposed framework, shaping the interactions, behaviors, and learning mechanisms of agents. Norms can be classified into two types: **fixed norms**, which establish a static structure with predetermined rules, e.g., MARL and games, and **dynamic norms**, which evolve through agent interactions and feedback to enable adaptive and self-organized relationship building. We primarily focus on dynamic norms.

The evolution of norms begins with initial protocols,  $\mathcal{P} = \{\mathcal{P}_1, \dots, \mathcal{P}_n\}$ , that establish rules for agents to build cooperation or manage conflicts. For instance, a basic protocol,  $\mathcal{P}_i = \alpha \cdot H_i$ , may dictate resource-sharing mechanisms by discouraging hoarding as follows:

$$J_i^{\mathcal{P}} = J_i - \mathcal{P}_i, \forall i \quad (1)$$

where  $J_i$  is agent  $i$ 's original goal (e.g., selfishly hoarding),  $J_i^{\mathcal{P}}$  is the goal subject to the protocol  $\mathcal{P}_i$ , and  $\alpha$  is the penalty coefficient for the hoarding behavior metric  $H_i$ .

Protocols can emerge as penalties on agents' objectives, as shown in (1), incentivizing agents to adjust their goals to align with collective priorities. Alternatively, protocols may act as constraints that limit the actions of agents, ensuring that their individual objectives do not violate safety, fairness, or ethical boundaries. This is particularly useful in applications like autonomous vehicles.

Protocols are learnable and evolve over time, helping agents achieve individual goals by managing conflicts, developing cooperation, or delegating subtasks to other agents. The evolution of protocols could be influenced by the agents' relationship network  $\mathcal{N}$ , their societal impacts  $\mathcal{I}$ , and the expectations  $\mathcal{E}$  held by peers. When agents consistently meet or exceed the expectations of their peers, they reinforce existing protocols, fostering trust and enhancing their influence. Conversely, deviations from expected behaviors—whether due to conflicts, inefficiencies, or environmental shifts—can trigger the refinement or replacement of protocols.

The relationship network  $\mathcal{N}$ , the impacts  $\mathcal{I}$ , and the expectations  $\mathcal{E}$  are not static. They evolve as agents interact with each other. As agents learn more about their environment and the behaviors of others, their actions and the resulting impacts shift, changing how they are perceived and how they perceive others. These dynamic relationships guide agents toward more effective adaptation. Through this continuous evolution of  $\mathcal{N}$ ,  $\mathcal{I}$ , and  $\mathcal{E}$ , agents gradually refine their strategies, norms, and interactions, aiming to achieve their goals in a dynamic ecosystem.

The stability of protocols can be analyzed using a dynamical system representation:

$$\frac{d\mathcal{P}}{dt} = -\gamma \cdot (\mathcal{P} - \mathcal{P}^*), \quad (2)$$

where  $\gamma$  is a stability coefficient and  $\mathcal{P}^*$  is the equilibrium protocol. A special case is that the equilibrium optimizes collective outcomes:  $\mathcal{P}^* = \arg \max_{\mathcal{P}} \sum_i J_i(\pi_i, \mathcal{P})$ , where  $\mathcal{P}^*$  can be seen as protocols used for traditional cooperative MARL problems. Note that protocols may not necessarily converge, indicating an ever-evolving ecosystem.

a) *Evolution of Relationships.*: The relationships among all agents collectively form the network  $\mathcal{N}$ , serving as the social fabric of the framework. Agent  $i$ 's relationships can be represented by a directed weighted graph  $\mathcal{G}_i$ , where each edge  $g_{ij} \in \mathcal{G}_i$  originates from agent  $i$  and points to agent  $j$ , with a weight  $w_{ij}$ , indicating the strength and nature of a relationship. The relationships are not necessarily symmetric.

The relationship strength  $w_{ij}$  is dynamic and evolves over time based on interactions and feedback, updated by:

$$w_{ij} \leftarrow f(w_{ij}, \mathcal{I}_{ij}, \mathcal{E}_{ij}) \quad (3)$$

where  $f(\cdot)$  is a function that governs the update process,  $\mathcal{I}_{ij}$  represents the impact of agent  $j$  observed by agent  $i$ , and  $\mathcal{E}_{ij}$  denotes the expectation that agent  $i$  holds about agent  $j$ .  $i$  establishes or dissolves relationships.

The evolution of relationships creates a dynamic network, allowing agents to form coalitions, dissolve unproductive ties, and adapt their connections to establish new norms. These adjustments create a feedback loop where successful collaborations strengthen relationships, while conflicts prompt reevaluation and adaptation. For example, agents with aligned goals may form coalitions to pool resources or address complex tasks. Mathematically, a coalition  $C$  may form when:  $\sum_{i,j \in C} w_{ij} \geq \theta$ , where  $\theta$  is a threshold for the minimum collective relationship strength required for coalition stability. Additionally, agents may prioritize communications or knowledge sharing with trusted peers, reducing resource expenditure and enhancing overall efficiency. This prioritization could be modeled as:  $P_i(j) = \frac{w_{ij}}{\sum_k w_{ik}}$ , where  $P_i(j)$  is the probability that agent  $i$  communicate with  $j$ .

b) *Cross-Environment Interactions.*: In our framework, agents interact across multiple environments, e.g., a hierarchy or heterarchy of energy management systems, each governed by unique rules, dynamics, and challenges. Unlike traditional frameworks confined to a single shared environment, these environments may be interconnected, overlapping, or nested within hierarchical structures, allowing agents to influence and be influenced by actions across environments.

Consider agents  $i$  and  $j$  operate in environments  $E_i$  and  $E_j$ , respectively. If the state transition of  $E_j$  depends on the actions in both  $E_i$  and  $E_j$ , it can be expressed as:

$$s^{E_j'} = P^{E_j}(s^{E_j}, a^{E_i}, a^{E_j}) \quad (4)$$

where agent  $i$ 's actions in  $E_i$  propagate to  $E_j$ , resulting in cross-environment consequences. For example, in a cyber-physical system, cyber agents (e.g., smart routers) operating in the cyber environment can introduce delays in sensor or control signals, subsequently affecting the physical environment. Similarly, agents managing logistics in a localized supply chain may impact global inventory levels, creating dependencies between local and global environments.

To navigate multiple environments, agents develop meta-coordination strategies that generalize collaborative behaviors, enabling them to handle diverse tasks while balancing individual objectives with broader societal impacts. By fostering cross-environment interactions, the framework transcends traditional

multi-agent models, offering a flexible approach to addressing real-world complexities.

### C. Evaluating the Framework: Metrics for Success

In the proposed framework, every agent strives to achieve its own goals while operating under the constraints of societal norms and collective dynamics. However, the inherent competition for limited resources, conflicting objectives, and interconnected nature make it challenging for all agents to fully realize their goals simultaneously. Evaluating such a framework requires metrics that capture not only the success of individual agents but also the health, fairness, stability, and adaptability of the entire system.

a) *Balancing Individual and Collective Success.*: A central challenge in the proposed framework is striking a balance between individual agents' goals and collective harmony. Each agent acts to optimize its own objectives, but in shared environments, these actions inevitably influence - and are influenced by - other agents that are not developed nor coordinated by any single entity. Although existing paradigms often consider this as a multi-agent, multi-utility, or game-theoretical problem, agentic AI solutions from different developers and providers continue to learn and update autonomously in a runtime environment. This inter-connected continual learning creates scenarios where adhering to norms may initially seem like a penalty to an agent's goal. However, such adherence is often necessary to ensure long-term success in an ecosystem where agents coexist and interact dynamically. A key question is how to evaluate whether a norm is appropriate for an agent to achieve its goal without being overly penalized. Effective norms align collective objectives with individual incentives, enabling agents to cooperate while pursuing their own goals. Metrics for this evaluation must capture the trade-offs involved. For example, measuring the *individual utility gain* can reveal how much a norm contributes to an agent's success in a multi-agent ecosystem. Similarly, assessing *norm efficiency* highlights how well a norm improves collective outcomes without imposing undue burdens on specific agents.

b) *Sacrifices and Fairness.*: Forming norms often necessitates sacrifices, requiring agents to prioritize collective benefits over individual goals. An agent may delay its own objective to comply with a norm that enhances system-wide stability, such as slowing down to prevent traffic congestion. A fundamental challenge is ensuring that these sacrifices yield tangible benefits beyond just pre-designed scenarios. To evaluate sacrifices, metrics such as *fairness* [39] assess whether the burden of sacrifice is shared proportionally across agents, considering factors like resource availability, task importance, or agent capacity. However, fairness does not imply equality. For example, in hierarchical settings, lower-level "worker" agents may accept greater sacrifices, such as reduced rewards, to support a "boss" agent tasked with achieving higher-level objectives. The critical metric here is whether these sacrifices are justifiable and evaluated fairly based on their contribution to the overall system. For instance, if a worker agent's sacrifices lead to significant improvements in the boss agent's success,

which in turn benefits the collective, the trade-off may be deemed fair within specific tasks. Balancing these dynamics ensures that norms remain robust and adaptable.

c) *Stability and Adaptability.*: Norms are crucial in guiding agents toward predictable and harmonious behaviors, ensuring the system functions cohesively over time. Metrics such as *norm stability* evaluate whether agents consistently adhere to a defined set of norms, fostering reliable interactions and reducing uncertainties. Similarly, *behavioral predictability*, measuring how agents' actions align with collective objectives, can offer insights into how effectively norms shape agent behavior to benefit the system. However, stability alone is insufficient in dynamic environments where static norms risk becoming obsolete. *Norm adaptability* must also be ensured to accommodate changes in agent goals, environmental conditions, and emerging challenges. Gradual and predictable norm evolution ensures that the system remains flexible and responsive without causing disruptions. Striking a balance between stability and adaptability is a defining feature of this framework's approach to norm management. By embedding mechanisms for controlled norm evolution, the framework ensures that agents can collectively adapt to shifting conditions while preserving a foundation of harmonious interactions. This balance is key to maintaining long-term system stability.

## IV. CASE STUDIES

### A. Autonomous Driving

Imagine a city where autonomous vehicles (AVs) from different makers share the streets. Each AV is trained - and continuously learning - to minimize travel time, reduce fuel consumption, and optimize route selection while avoiding collisions as individual agents, but their approaches vary by the developers and operators. For example, one might brake aggressively, while another calculates its stopping distance more cautiously. Without a shared understanding, these differences can lead to expectation misalignment, causing AVs to learn reciprocally and react unpredictably to one another's evolving interdependent strategies [40]–[45]. This misalignment can lead to traffic congestion, as illustrated in Figure 2a. Additionally, AVs may clash unexpectedly when inferring the others' evolving strategies: one defensive AV observes and infers that another AV is driving aggressively and thus stops abruptly to keep a safe distance, only to confuse a third AV that expect the first one to keep a stable speed, triggering a chaotic ripple effect that can escalate into a chain-reaction accident, as illustrated in Figure 2b. Over time, the uncertainty may prompt AVs to adopt overly cautious behaviors—driving slowly, leaving large gaps, and hesitating at intersections. While these strategies reduce accidents, they disrupt traffic flow, leading to longer commutes, wasted fuel, and overall inefficiencies.

Now, consider that these AVs are developed with dynamic **norms**. They learn flexible **protocols** that enable them to deviate from their original goals to explore different behaviors. For instance, protocols may be rewards that initially encourage AVs to drive more aggressively. Then, they experiment—one car tests how closely it can follow another without colliding,

Mistakes happen, and a minor fender bender might teach a valuable lesson about boundaries. Over time, the cars observe patterns, exchange signals, and learn to anticipate one another's reactions. This iterative process enables protocols to evolve over time.

In this process, two crucial forces — **impacts** and **expectations** — drive the evolution of protocols. Each AV begins by evaluating the impacts of its actions. For example, an AV merging aggressively into another lane may observe how this forces neighboring cars to brake suddenly, increasing the risk of collision and creating congestion. Recognizing this negative impact enables it to revise its merging protocol to balance assertiveness with consideration for others. Furthermore, these AVs learn to anticipate the expectations of others. A car approaching a busy roundabout knows the standard rules: yield or proceed predictably. When it hesitates too long, it disrupts the rhythm, confusing others and slowing the entire system. Through trial and error, it begins to act with confidence, aligning its actions with the expectations of other vehicles.

The **relationship network** influences how AVs anticipate expectations. Repeated interactions between the same vehicles strengthen trust, allowing them to better predict one another's actions. Over time, this shared understanding fosters coordina-

tion. For instance, an AV that consistently signals its intentions encourages others in its network to rely on its predictability, reducing uncertainty and hesitation. Conversely, a vehicle that frequently violates expectations by breaking unpredictably or merging erratically may erode trust, forcing others to adopt more defensive driving strategies.

*a) Cooperation or Competition? The Role of Algorithms.:*

This process does not naturally guarantee cooperation. The design of the AVs' learning algorithms determines how impacts, expectations, relationships, and protocols evolve. Algorithms that prioritize collective efficiency may drive autonomous vehicles (AVs) to focus on system-wide benefits, thereby fostering coordination and cooperation. In contrast, algorithms that reward individual optimization might lead AVs to exploit their relationships, taking advantage of others' predictable behavior for personal gain. For example, an AV might learn to drive more boldly, knowing others will yield. While this benefits the individual vehicle in the short term, it may strain the network, creating inefficiencies over time. Alternatively, an algorithm focused on collective benefits might encourage AVs to adopt protocols that balance assertiveness with consideration, ensuring smoother traffic flow for all.

*B. Case Study 2: Energy Autonomy*

Energy autonomy refers to a new paradigm of a distributed, agentic ecosystem where different stakeholders may adopt their chosen AI algorithms into energy management systems (EMS), designed to act and optimize toward corresponding objectives [46]. This shift occurs as distributed energy resources (DERs) like solar panels, battery storage, and electric vehicles are increasingly run by standalone EMS. While the grid has been traditionally controlled by electric utilities, DER owners can adopt their own AI solutions for their own use cases; these AI-aided EMS become self-serving agents acting on behalf of their stakeholders to learn and optimize toward individual objectives [47].

The result is a complex network of **inter-operating agentic EMS** driven by different algorithms and goals. They can no longer treat other agents as part of the environment; instead, each agent observes, learns from, influences, and confronts other interacting agents. These agentic EMS can thus become self-serving players in the electricity market with oscillatory learning objectives, processes, and decisions.

An illustrative example of the resulting challenges can be found in **conflicts along a hierarchical decision chain** by agentic EMS, shown on the left of Figure 3. An intelligent vehicle battery management system can learn to optimize its state of charge for longer life via restrictive charging and battery balancing; such decisions may contradict a smart wall charger's algorithm trying to maximize charging power for a favorable price or a planned trip. Building automation systems will learn from occupant profiles and preferences to optimize the energy consumption of connected appliances. The resulting frequent set-point changes can lead to unpredictable patterns, unlike those used by load models employed by utilities, which are adopting AI for better load forecasting and demand-side

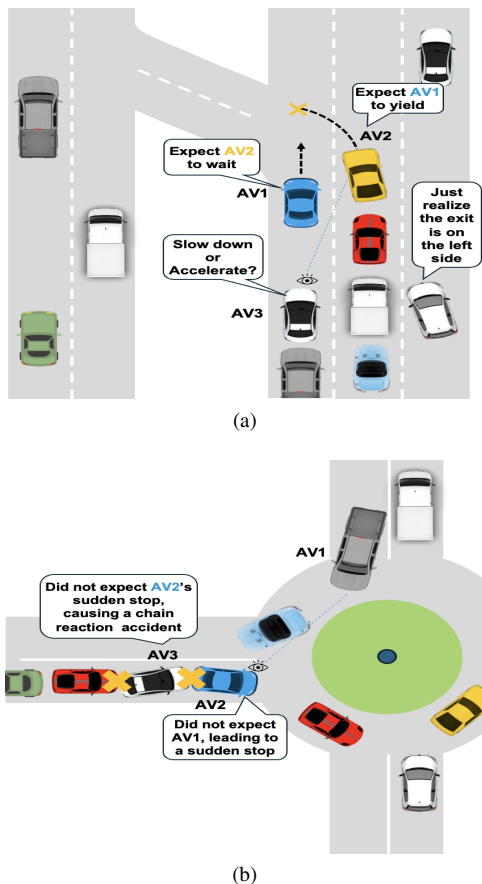


Fig. 2. Challenges of uncoordinated AI agents in autonomous driving: (a) Expectation misalignment between AVs when switching lanes causes traffic jams; (b) A sudden stop at a roundabout due to a cautious strategy causes a chain accident.

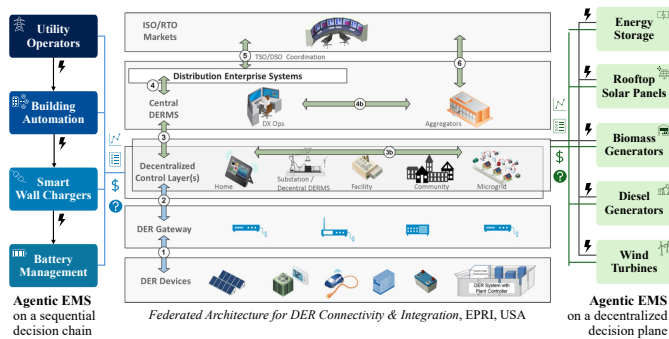


Fig. 3. Agentic EMS in energy autonomy, mapped on the hierarchical architecture for DERs [48].

management. The conflicts above may be a result of not only insufficient coordination among developers but also self-serving policies that agents learn from inherently incompatible objectives across different stakeholders. Collaborative norms and protocols are therefore needed in addition to individual agentic EMS designs to describe, analyze, and resolve such conflicts.

Another illustrative example is **implicit collusion in a decentralized market** among agentic EMS, as shown on the right of Figure 3. When various DER owners adopt autonomous AI to optimize pricing on the electricity market, tacit collusion may occur when the agents learn to conspire toward an unfair disadvantage spontaneously: A shared data analytics platform collecting information from subscribing DER providers may inform the best pricing to all subscribers who will thus (un)knowingly price gauge simultaneously; collusion can also be triggered when multiple agents learn to form a coalition and started price gauging without explicit communication or formal agreement. In either case, shared interests and information among agents — a common presumption in multi-agent system theories — can lead to tacit collusion and unfair advantages over other agents. New tools and metrics are needed to monitor and detect such behaviors closely, ensuring that collaborations do not cross the line into collusion.

## V. CHALLENGES, RISKS, AND OPPORTUNITIES

### A. Managing Emergent and Unpredictable Learned Behaviors

In the envisioned multi-agent ecosystems, emergent behaviors arise from interactions between agents and their environments, both before and after deployment. While such behaviors can lead to innovative and efficient solutions, they can also result in unintended consequences, such as resource monopolization, inefficiencies, or conflicts, particularly during the post-deployment stage when fine-tuning and/or continual updates are still applied, leading to oscillatory decisions and interplays. The unpredictability of such emergent behaviors and dynamics complicates system design and oversight, particularly in open environments where agents continuously adapt and learn.

To address these challenges, future research should explore the types of norms that can encourage beneficial emergent

behaviors while minimizing harmful outcomes. One promising avenue involves designing methodologies to monitor, model, and predict emergent dynamics in real time, ensuring that agents remain adaptable without losing control over their collective actions. Additionally, mechanisms to manage feedback loops are crucial to prevent destabilizing effects that could cascade through the system, particularly in environments where agents influence one another's learning trajectories. Understanding and guiding emergent behaviors is essential to ensure that multi-agent systems remain resilient and productive as they scale in complexity.

### B. Designing Algorithms for Dynamic Objectives

Agents often face shifting priorities and evolving goals as they interact with dynamic environments and diverse peers. Enabling agents to adjust their objectives while maintaining stability and coherence presents a significant technical challenge. Developing algorithms for dynamic goal-setting and adaptation is a crucial research direction. Such algorithms must enable agents to align their evolving goals with the broader needs of the system while preserving individual autonomy. Techniques such as reinforcement learning and meta-learning show promise in handling dynamically changing objectives, but further innovation is needed to address the trade-offs between short-term responsiveness and long-term optimization. Research should also examine constraints and safeguards to ensure that agents' adaptations remain aligned with the goals and values of both individual users and the overall system.

### C. The Dilemma of Self-Reference

The ability of autonomous agents to revise their own objectives introduces a fundamental recursive challenge: How can an agent establish its initial goal if the process of goal revision itself requires a predefined objective? This creates a chicken-and-egg dilemma—an agent needs a goal to guide its behavior, yet once granted autonomy to modify that goal, the basis for the initial objective becomes unclear. This raises a key question: If agents determine their own goals, what serves as the initial guiding principle, and how does the agent decide whether a revision is an improvement or a deviation? Without an initial reference point, goal revision risks becoming arbitrary or contradictory. This poses a risk in multi-agent systems, where alignment among agents is critical—if each agent independently shifts priorities, collective coordination may break down.

To address this dilemma, agents need an anchoring framework for goal revision, ensuring modifications occur within a coherent evaluative structure rather than being entirely self-referential. This framework should enable adaptive yet bounded evolution to balance between flexibility and constraint, allowing agents to explore alternative objectives while preventing divergence that could lead to instability.

### D. Ethical Issues of Free Will in Agents

The capacity of agents to autonomously revise their objectives poses a profound ethical question: to what extent

can such agents be said to possess a form of free will? In a sense, agents become products of their own design, which challenges traditional notions of AI control and autonomy. This ambiguity raises critical concerns about accountability [49]. If a self-modified agent's actions result in harm or conflict with societal values, where does responsibility lie—with the designer, the user, or the agent itself? Such dilemmas are especially urgent in high-stakes domains like healthcare or autonomous transportation, where goal realignment could have significant consequences.

To navigate this terrain, agents must operate within ethical constraints that ensure their goal modifications remain aligned with human benefits. Transparent mechanisms are essential to make these modifications interpretable, allowing human stakeholders to monitor and guide agents' evolving priorities. Interdisciplinary efforts involving ethics, philosophy, and law are crucial to addressing questions of moral agency, responsibility, and oversight. As agents grow increasingly autonomous, clear frameworks will be needed to balance their capacity for self-determination with the broader priorities of society.

### E. Human-Agent Collaboration

As AI agents increasingly interact with humans in real-world settings, ensuring effective collaboration will be a critical challenge. Agents must not only understand and respond to human preferences but also adapt to diverse cultural, ethical, and contextual factors, ensuring that their actions align with the expectations of diverse stakeholders. Designing systems that facilitate transparent and effective communication between humans and agents is a key priority. This includes developing interfaces and protocols that help humans to better understand, predict, and trust agent behaviors intuitively. In contrast, human oversight can guide and shape agent behaviors, together with agents' knowledge, co-develop norms, facilitating collaborative frameworks that combine the strengths of human judgment and machine autonomy. By addressing these factors, future research can pave the way for robust and equitable human-agent partnerships.

## VI. CONCLUSION

We posit that the rise of autonomous AI systems in critical infrastructures necessitates a fundamental rethinking of interoperability among independently designed, developed, and deployed AI agents in open environments. Autonomous machine learning and decision-making will become pervasive - and thus interactive - beyond the development phases in various critical infrastructure systems, processes, and services. We advocate for a shift toward the exploration of self-adaptive and self-organizing multi-AI ecosystems, propose a conceptual framework that integrates adaptive norms, evolving protocols, and dynamic relationships to provoke research questions and investigations, and highlight some key challenges and opportunities for a potential research paradigm through two case studies in critical infrastructure applications.

## ACKNOWLEDGEMENTS

This work is supported in part by the ..., the Concordia University Research Chair program, the Natural Sciences & Engineering Research Council (NSERC) of Canada under grants RGPIN-2018-06724 and RGPIN-2025-05097, the U.S. National Science Foundation under the grant number 2330504, the ... and the ... . The authors would also like to thank ..., ..., and ... for their valuable feedback.

## REFERENCES

- [1] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [2] D. Silver *et al.*, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, Jan 2016.
- [3] Y. Shavit *et al.*, "Practices for governing agentic ai systems," *Research Paper, OpenAI, December*, 2023.
- [4] A. Chan *et al.*, "Harms from increasingly agentic algorithmic systems," in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, ser. FAccT '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 651–666.
- [5] Y. LeCun and Y. Bengio, "Reflections from the turing award winners," 2020. [Online]. Available: <https://iclr.cc/virtual2020/speaker7.html>
- [6] T. Chen, S. Kornblith, K. Swersky, M. Norouzi, and G. E. Hinton, "Big self-supervised models are strong semi-supervised learners," *Advances in neural information processing systems*, vol. 33, pp. 22 243–22 255, 2020.
- [7] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proceedings of the 34th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, D. Precup and Y. W. Teh, Eds., vol. 70. PMLR, 06–11 Aug 2017, pp. 1126–1135.
- [8] A. Rajeswaran, C. Finn, S. M. Kakade, and S. Levine, "Meta-learning with implicit gradients," in *Advances in Neural Information Processing Systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds., vol. 32. Curran Associates, Inc., 2019.
- [9] B. Luitel and G. K. Venayagamoorthy, "Cellular computational networks—a scalable architecture for learning the dynamics of large networked systems," *Neural Networks*, vol. 50, pp. 120–123, 2014.
- [10] J. Ault and G. Sharon, "Reinforcement learning benchmarks for traffic signal control," in *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 1)*, 2021.
- [11] Z. Wan, H. Li, H. He, and D. Prokhorov, "Model-free real-time ev charging scheduling based on deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5246–5257, 2019.
- [12] X. Chen, G. Qu, Y. Tang, S. Low, and N. Li, "Reinforcement learning for selective key applications in power systems: Recent advances and future challenges," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 2935–2958, 2022.
- [13] D. Kalashnikov *et al.*, "Scalable deep reinforcement learning for vision-based robotic manipulation," in *Proceedings of The 2nd Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, A. Billard, A. Dragan, J. Peters, and J. Morimoto, Eds., vol. 87. PMLR, 29–31 Oct 2018, pp. 651–673.
- [14] C. Tang, B. Abbatemateo, J. Hu, R. Chandra, R. Martín-Martín, and P. Stone, "Deep reinforcement learning for robotics: A survey of real-world successes," 2024.
- [15] L. Zhou, L. Zhang, and B. K. Horn, "Deep reinforcement learning-based dynamic scheduling in smart manufacturing," *Procedia Cirp*, vol. 93, pp. 383–388, 2020.
- [16] Y. Zhang, H. Zhu, D. Tang, T. Zhou, and Y. Gui, "Dynamic job shop scheduling based on deep reinforcement learning for multi-agent manufacturing systems," *Robotics and Computer-Integrated Manufacturing*, vol. 78, p. 102412, 2022.
- [17] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: A Bradford Book, 2018.
- [18] M. Tan, *Multi-Agent Reinforcement Learning: Independent vs. Cooperative Agents*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1997, p. 487–494.

- [19] C. Claus and C. Boutilier, "The dynamics of reinforcement learning in cooperative multiagent systems," in *Proceedings of the Fifteenth National/Tenth Conference on Artificial Intelligence/Innovative Applications of Artificial Intelligence*, ser. AAAI '98/IAAI '98. USA: American Association for Artificial Intelligence, 1998, p. 746–752.
- [20] S. Kar, J. M. F. Moura, and H. V. Poor, " $QD$ -learning: A collaborative distributed strategy for multi-agent reinforcement learning through consensus + innovations," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1848–1862, 2013.
- [21] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," in *the 31st Conference on Neural Information Processing Systems*, ser. NeurIPS'20, 2017, pp. 6382–6393.
- [22] K. Zhang, Z. Yang, and T. Başar, "Multi-agent reinforcement learning: A selective overview of theories and algorithms," *ArXiv pre-prints*, 2019, arXiv:1911.10635.
- [23] M. L. Littman, "Markov games as a framework for multi-agent reinforcement learning," in *Machine learning proceedings 1994*. Elsevier, 1994, pp. 157–163.
- [24] O. Slumbers, D. H. Mguni, S. B. Blumberg, S. M. Mcaleer, Y. Yang, and J. Wang, "A game-theoretic framework for managing risk in multi-agent systems," in *Proceedings of the 40th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, and J. Scarlett, Eds., vol. 202. PMLR, 23–29 Jul 2023, pp. 32 059–32 087.
- [25] S. Mao *et al.*, "Allympics: Llm agents meet game theory—exploring strategic decision-making with ai agents," *arXiv preprint arXiv:2311.03220*, 2024.
- [26] Y. Yang, H. Du, G. Sun, Z. Xiong, D. Niyato, and Z. Han, "Exploring equilibrium strategies in network games with generative ai," *arXiv preprint arXiv:2405.08289*, 2024.
- [27] K. Gal and B. J. Grosz, "Multi-agent systems: Technical & ethical challenges of functioning in a mixed group," *Daedalus*, vol. 151, no. 2, pp. 114–126, 2022.
- [28] M. Wooldridge and N. R. Jennings, "Intelligent agents: Theory and practice," *The knowledge engineering review*, vol. 10, no. 2, pp. 115–152, 1995.
- [29] Z. Durante *et al.*, "Agent ai: Surveying the horizons of multimodal interaction," 2024. [Online]. Available: <https://arxiv.org/abs/2401.03568>
- [30] J. Zhang *et al.*, "Aflow: Automating agentic workflow generation," *arXiv preprint arXiv:2410.10762*, 2024.
- [31] W. J. Yun, J. Park, and J. Kim, "Quantum multi-agent meta reinforcement learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 9, 2023, pp. 11 087–11 095.
- [32] C.-D. WANG and X.-F. WANG, "Multi-agent based architecture of context-aware systems," in *2007 International Conference on Multimedia and Ubiquitous Engineering (MUE'07)*, 2007, pp. 615–619.
- [33] H. Du, S. Thudumu, R. Vasa, and K. Mouzakis, "A survey on context-aware multi-agent systems: Techniques, challenges and future directions," 2024. [Online]. Available: <https://arxiv.org/abs/2402.01968>
- [34] G. Jelen, J. Babic, and V. Podobnik, "A multi-agent system for context-aware electric vehicle fleet routing: A step towards more sustainable urban operations," *Journal of cleaner production*, vol. 374, p. 134047, 2022.
- [35] A. Amirkhani and A. H. Barshooi, "Consensus in multi-agent systems: a review," *Artificial Intelligence Review*, vol. 55, no. 5, pp. 3897–3935, 2022.
- [36] K. R. McKee *et al.*, "A multi-agent reinforcement learning model of reputation and cooperation in human groups," *arXiv preprint arXiv:2103.04982*, 2021.
- [37] S. Sukhbaatar, R. Fergus *et al.*, "Learning multiagent communication with backpropagation," *Advances in neural information processing systems*, vol. 29, 2016.
- [38] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, p. 6382–6393.
- [39] N. A. Grupen, "Responsible emergent multi-agent behavior," 2023. [Online]. Available: <https://arxiv.org/abs/2311.01609>
- [40] S. Shalev-Shwartz, S. Shammah, and A. Shashua, "Safe, multi-agent, reinforcement learning for autonomous driving," 2016. [Online]. Available: <https://arxiv.org/abs/1610.03295>
- [41] P. Palanisamy, "Multi-agent connected autonomous driving using deep reinforcement learning," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–7.
- [42] A. Wachi, "Failure-scenario maker for rule-based agent using multi-agent adversarial reinforcement learning and its application to autonomous driving," 2019. [Online]. Available: <https://arxiv.org/abs/1903.10654>
- [43] G.-P. Antonio and C. Maria-Dolores, "Multi-agent deep reinforcement learning to manage connected autonomous vehicles at tomorrow's intersections," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 7, pp. 7033–7043, 2022.
- [44] B. R. Kiran *et al.*, "Deep reinforcement learning for autonomous driving: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4909–4926, 2022.
- [45] C. Gulino *et al.*, "Waymax: An accelerated, data-driven simulator for large-scale autonomous driving research," in *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Eds., vol. 36. Curran Associates, Inc., 2023, pp. 7730–7742.
- [46] J. K. Juntunen and M. Martiskainen, "Improving understanding of energy autonomy: A systematic review," *Renewable and Sustainable Energy Reviews*, vol. 141, p. 110797, 2021.
- [47] G. A. Daneke, "Machina-economicus or homo-complexicus: Artificial intelligence and the future of economics?" *Real-World Economic Review*, vol. 93, pp. 18–39, 2020.
- [48] J. Chen, J. Yan, A. Kemmeugne, M. Kassouf, and M. Debbabi, "Cybersecurity of distributed energy resource systems in the smart grid: A survey," *Applied Energy*, vol. 383, p. 125364, 2025.
- [49] J. Dai, "Position: Beyond personhood: Agency, accountability, and the limits of anthropomorphic ethical analysis," in *Forty-first International Conference on Machine Learning*, 2024.