IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received 14 November 2024, accepted 10 December 2024, date of publication 25 December 2024, date of current version 10 January 2025.

Digital Object Identifier 10.1109/ACCESS.2024.3522362

PERSPECTIVE

Autonomous Mental Development at the Individual and Collective Levels: Concept and Challenges

MARCO LIPPI[1], STEFANO MARIANI[2], MATTEO MARTINELLI[2], AND FRANCO ZAMBONELLI[2], (Fellow, IEEE)

[1]Dipartimento di Ingegneria dell'Informazione, Università di Firenze, 50139 Florence, Italy
[2]Dipartimento di Scienze e Metodi dell'Ingegneria, Università di Modena e Reggio Emilia, 42122 Reggio Emilia, Italy

Corresponding author: Franco Zambonelli (franco.zambonelli@unimore.it)

This work was supported in part by Italian Ministero Università e Ricerca (MUR), Programmi di Ricerca di Interesse Nazionale (PRIN) 2017 Program, Project Fluidware.

ABSTRACT The increasing complexity and unpredictability of many ICT scenarios let us envision that future systems will have to dynamically learn how to act and adapt to face evolving situations with little or no a priori knowledge, both at the level of individual components and at the collective level. In other words, such systems should become able to autonomously (i.e., self-) develop mental models of themselves and of their environment. Autonomous mental development includes: learning models of own capabilities; learning how to act purposefully towards the achievement of specific goals; and learning how to act in the presence of others, i.e., at the collective level. In this paper, we introduce a conceptual framework for autonomous mental development in ICT systems – at both the individual and collective levels – by framing its key concepts and illustrating suitable application domains. Then, we overview the many research areas that are contributing or can potentially contribute to the realization of the framework, and identify some key research challenges.

INDEX TERMS Autonomous agents, learning, multiagent systems, self-adaptation, self-organization.

I. INTRODUCTION

Since their early months, human infants start experiencing their own bodies, moving hands, touching objects, and interacting with people around them. Such activities are part of an overall process of *autonomous mental development* (aka self-development), which lets them gradually develop cognitive and behavioral capabilities [69]. These skills include the capability to recognize situations around, the sense of self [17], the sense of agency (i.e., understanding the effect of own actions in an environment) [43], the capability to act purposefully towards a goal, and some primitive social capabilities (i.e., knowing how to act in the presence of others).

Moving from humans to machines, the possibility of building ICT systems capable of autonomously developing their own mental and social models and acting purposefully in an environment is increasingly recognized as a key challenge in many areas of artificial intelligence (AI) [83], such as robotics [41], [86], intelligent IoT [46], autonomous vehicles management [89].

Indeed, for small-scale and static scenarios and for simple goal-oriented tasks, it is possible to "hardwire" a model of the environment within a system, alongside some pre-designed plans of action. However, for larger and more dynamic scenarios, and for complex tasks, individual components of ICT systems should be able to autonomously (i.e., without human supervision): *(i)* build environmental models and continuously update them as situations evolve; *(ii)* develop the capability of recognizing and modelling the effect of their own actions on the context (which variables of the environment can or cannot be directly affected by which actuators, which variables and actuators relate to each other); *(iii)* learn to achieve goals on this basis and depending on the current situation; and eventually *(iv)* learn to interact with others to coordinate actions.

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

The ambitious idea of building systems capable of autonomous development is not new, and its opportunity has already been advocated for several years [86]. Also, many similar research ideas have been conceived under different names. For instance, researches under the names of "Self-aware computing systems" [39], "Autonomic computing" [33], "Organic computing" [58], "Lifelike computing systems" [77], and finally, "Self-improving system integration", [10]. All of the abbove foster systems that are able to learn models of their own structure, behavior, and relationship with their operational environment, and use such models to plan actions achieving their intended goal despite the potential disruption of environment changes. Autonomous mental development differs as it emphasizes the continual learning process underlying autonomous mental development, and the explicit nature of the mental models built with it. Additionally, we consider both the individual and collective dimensions, and we interpret the process more as bottom-up rather than top-down—of course, traditional top-down engineering is not ruled out.

The topic is now even more timely. Many recent research results in areas such as causal analysis and causal learning [1], [52], reinforcement learning [56] multiagent learning [19], and self-organizing behaviors [72], have started shedding light on the various mechanisms that have to be involved in the process of autonomous development. Such results, in addition to the recent breakthroughs in the area of large language models [21], [38] and large generalist agents built on top of generative AI technologies [24], [95], hint at the fact that what once was only a vision (at least in specific application areas) is getting closer to become a reality. Furthermore, unfolding the key concepts and mechanisms underlying the very notion of autonomous development can also contribute to understanding some of the many mental mechanisms behind artificial general intelligence [14], [49].

Against this background, the contribution of this paper is to frame the key concepts of autonomous mental development in ICT systems and to identify challenges and promising research directions. More in particular: Section II introduces a general conceptual framework for the (continuous and adaptive) process of autonomous mental development, both at the individual and at the collective level; Section III sketches some key application scenarios; Section IV analyzes the most promising approaches in the area of machine learning, multiagent systems, and collective adaptive systems that can contribute with fundamental building blocks towards realizing the vision of autonomous mental development, each *per se* challenging; Section V identifies additional horizontal challenges to be attacked, emphasizing their inherent cross-disciplinary nature.

We emphasize that an earlier and much shorter version of this paper appeared in [45]. This new version contains more detailed discussions, updated to cover recent research results in the AI field, as well as more references to related works and more representative and updated figures.

II. CONCEPTUAL FRAMEWORK

Autonomous mental development, besides being the process that infants carry out during the early stages of their life [69], also involves any "agent" whenever it is incarnated in a new body and immersed in a new environment, be it a physical, virtual, or a mixed cyber-physical one.

As an example, to quickly and intuitively introduce our general framework (Figure 1), let us consider what we do whenever we start playing a new videogame. At first, we observe the game environment on the screen and the commands we have available on the joystick; that is, we get acknowledged with our *embodiment* in and *perception* of the videogame. We spend a few seconds trying the commands to assess their effects in the game environment; that is, we try to acquire a *sense of agency*. Then, we understand what is the goal of the game and how we can use the commands to achieve it; that is, we start acting and planning future actions in a *goal-oriented* way.

Typically, we recognize in the videogame the presence of other "agents", virtual characters that are not under our control; that is, we distinguish between *self* and *non-self*. The acquisition of such a skill implies that we acknowledge that we should tune our actions also in dependence on the actions of these other agents (*strategic thinking*). All this process is typically repeated in a cyclic way (i.e., when reaching a new level in the game) to adapt to new environments, new situations, new tools available to play with, new goals, and new virtual characters (e.g., enemies) appearing.

In the case of multiplayer games, besides recognizing the presence of players different from ourselves, and recognizing the need to act also accounting for them, we should understand: whether we have *communication* tools available, and how to use these tools to affect and influence the actions of others, i.e., to *coordinate* with them, so that eventually structured (i.e. *institutional*) ways to act together towards a goal can be established. Again, this process may be cyclically repeated as the game advances.

Truly intelligent and adaptive ICT systems should undergo a similar process and autonomously develop their artificial minds through similar phases. Thus, in the following, we analyzes each of such phases, whose key characteristics are summarized in 1.

A. THE INDIVIDUAL LEVEL

Let us first consider a single agent X (purely software or physically embodied) immersed in a (virtual or physical) environment. The agent can observe a set of environmental variables $\mathcal{V} = \{v_1, v_2, \ldots, v_m\}$. For simplicity and without loss of generality, internal variables of the agent itself (i.e., its current status and configuration) are included in the set. In addition, the agent has a set of actions that it can choose from $\mathcal{A} = \{a_0, \ldots, a_{n-1}, \mathit{null}\}$, including the *null* action.

1) EMBODIMENT AND PERCEPTION

In this very early phase, the agent should autonomously recognize the existence of $\mathcal{A}$ and $\mathcal{V}$, that is, it should

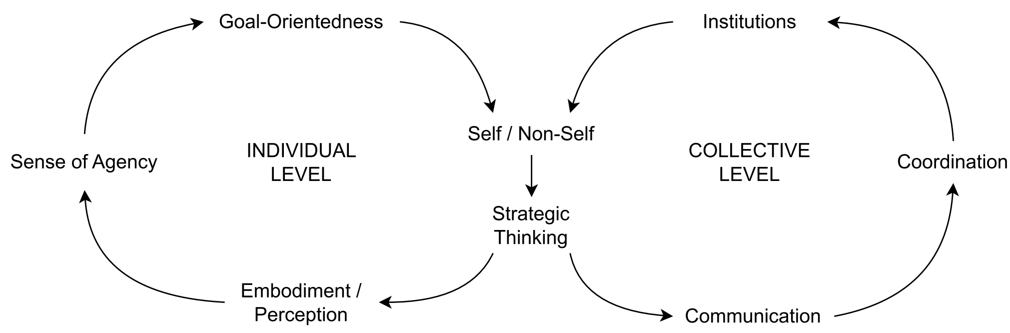


FIGURE 1. **The conceptual framework of autonomous mental development.**

get acquainted with its actuation and sensory capabilities. Without resorting to complex AI techniques, methods from the reflective and self-aware programming systems [71] can effectively apply in this phase to let the agent dynamically self-inspect its capabilities ($\mathcal{A}$) and start analyzing the observed variables ($\mathcal{V}$). Still in this phase, the agent can also start acquiring some understanding of the relations between variables in $\mathcal{V}$ over time, as well as some simple prediction capabilities.

2) SENSE OF AGENCY

In this exploratory phase, the agent starts trying to understand what are the effects of $\mathcal{A}$ on $\mathcal{V}$, by trying to apply actions (even without any specific goal in mind) to see their effects. That is, it will eventually recognize that, given the current state of $\mathcal{V}$ ($\mathcal{V}_{current}$), the application of an action a_i (or of a sequence of actions) in $\mathcal{A}$ will eventually lead (with some probability) to state $\mathcal{V}_{next}$. This mechanism enables the construction of the basic sense of agency [69], and of the sense of causality from $\mathcal{A}$ to $\mathcal{V}$.

3) GOAL-ORIENTEDNESS

In this exploitation phase, the agent starts applying $\mathcal{A}$ with goals in mind. That is, given the current state $\mathcal{V}_{current}$ and a desired future state $\mathcal{V}_g$ (the goal, aka the desired "state of the affairs"), the agent resorts to the acquired sense of agency by applying those actions in $\mathcal{A}$ that can possibly lead to $\mathcal{V}_g$. This also involves achieving the capability of planning the required sequence of actions to achieve $\mathcal{V}_g$.

4) SELF AND NON-SELF

As soon as an individual agent starts exploring its own actions in $\mathcal{A}$, and recognizes that such actions have effect on $\mathcal{V}$, it also understands that there are effects that are not under its own control (i.e. that does not belong to its $\mathcal{A}$). That is, there are "non-self" entities acting in the environment, too, either with the same $\mathcal{A}$ or a (partially) disjoint set $\mathcal{A}'$. By learning how to apply $\mathcal{A}$, the agent also learns the limits of such actions because of the non-self entities affecting some variable v_i in $\mathcal{V}$.

5) STRATEGIC THINKING

The agent has built a model of the world, that is, of how $\mathcal{A}$ affects $\mathcal{V}$, and it starts including the mental models of others (non-self) [79] while acting, as well as while designing strategies. That is, it can recognize that there are goals that it can possibly (or hopefully) attain only by accounting for the actions of others, be them cooperative or competitive.

As in the videogame example, autonomous development is not to be conceived as a "once-and-for-all" process. Rather, it is a continuous, never-ending loop: environmental conditions can change, new sensors may become available to enable more detailed observations (thus expanding $\mathcal{V}$), new actions become feasible (hence expanding $\mathcal{A}$) – or vice versa, some sensors and actions may no longer be available – and new virtual characters may appear or disappear. This requires agents to re-tune their learned sense of agency, and re-think how to achieve goals in isolation and in the presence of non-self entities.

We note that the learning mechanism that we envisage for driving the autonomous mental development of an agent is significantly different from the learning paradigm of foundation models such as LLMs. Specifically, the learning of such models is either self-supervised [96], using large corpora of raw text to learn next-word completion, or strongly based on human supervision, as in the case of reinforcement learning from human feedback (RLHF) [20]. Curiosity-driven learning and autonomous development are currently not among the principles that regulate the learning of foundation models. The collective level, which we present in the next section, is another dimension that has been started being investigated only recently by the most recent models [98].

B. THE COLLECTIVE LEVEL

In the presence of multiple agents acting in the same environment, agents could recognize that there are goals that cannot be achieved in isolation or by simply applying strategic thinking. Thus as part of their autonomous development, they should collectively develop some forms of "autonomous social engagement".

TABLE 1. Core skills considered in autonomous mental development, along with examples of their practical application.

Skill	Individual	Collective	Description
Embodiment/Perception	✓	–	Recognizing the existence of an environment, and own affordances
Sense of Agency	✓	–	Acquiring causal knowledge by exploring the effects of actions
Goal-Orientedness	✓	–	Applying actions purposefully meant to achieve a specific goal
Self/Non-Self	✓	✓	Understanding that something is out of own control, thus recognising the existence of other agents
Strategic thinking	✓	✓	Incorporating knowledge or expectations about other actions in own decision making, for cooperation or competition
Communication	–	✓	Exploring and using communication forms (explicit and implicit) and understanding the effect of communications on other agents
Coordination	–	✓	Applying communications to acquire information and request actions for (collective) goal achievement
Institutions	–	✓	"Istitutionalisation" of emergent coordination patterns useful to achieve collective goals

Formally, this corresponds to considering a set of K agents $X_0, \ldots, X_{K-1}$, where *(i)* each agent can choose the actions to perform from its own set $\mathcal{A}^k$ (possibly disjoint or only partially overlapping from those of the other agents); *(ii)* not necessarily all the agents can observe the whole set of environmental variables, but more likely each agent X_j has the capability to perceive and/or control a subset $\mathcal{V}^k$ of them. Thus, for specific goals $\mathcal{V}_g$ to be achieved, there is the need to properly combine and sequence actions by different agents, e.g. X_i executes a_w^i whereas X_j executes a_z^j, and so on.

1) COMMUNICATION

To overcome the limitations of strategic thinking, agents should be provided with a specific set of *communication actions*, i.e., actions that are devoted to influencing the actions of other agents. These could take the form of explicit communication acts, i.e., messages, that the agent should learn how to receive and send as an additional – social – form of perception and action. However, they could also take the form of more indirect actions aimed at affecting the behavior of others, i.e., leaving signs in the environment (stigmergy) or adopting peculiar behaviors aimed at being noticed by others (behavioral implicit communication) [50]. All these cases can be formalized by augmenting the $\mathcal{A}$ set to include communication actions, and possibly $\mathcal{V}$, to include observable signs in the environment.

2) COORDINATION

By exploring its available communication actions, agents start understanding how such acts can be used to get access to and affect some of the variables of the environment, and in particular those that are not observable and controllable by themselves. For instance, agent X_k can learn how to use communication acts to get access to the value of some non-observable variables $v_i \notin \mathcal{V}^k$, or to direct other agents in executing the actions $a_w \notin \mathcal{A}^k$ that can affect its values $v_{j \neq i} \in \mathcal{V}^k$ as required for a goal to be achieved. In other words, such explorations enable learning basic forms of coordination, which can be thought of as a social form of the sense of agency.

3) INSTITUTIONS

Eventually, after exploring coordination protocols, the agents can "institutionalize" their patterns of interaction towards collective actions. That is, they will learn those acceptable social patterns of coordination, and the set of social norms and social incentives, that enable them to systematically achieve goals together [57]. Formally, this corresponds to having agents in the collective recognize and adhere to a set of constraints $\mathcal{C}(\mathcal{A} \mid \mathcal{V})$ ruling the way actions $\mathcal{A}$ (there included communication actions) can be performed in specific conditions $\mathcal{V}$, as well as the commitments and expectations each communication action sets on the agents participating in the protocol.

As for the case of the individual level, the dynamics of the environment or of the agent population may require the above collective process to assume a continuous cyclic nature. This is what happens in video games played in groups: after a first acquaintance of involved players (i.e., recognition of self/non-self and strategic thinking phases), they start to coordinate with each other through communication means of various types, that can be reused several times if they appear to be effective toward the shared goal, letting them become *institutionalized*. The cyclic nature of this process is due to the fact that communication tools or strategies may emerge in the game, and they can be used to learn new ways to coordinate and new institutionalized group strategies.

We note that communication, coordination, and institutions are not *strictly* necessary to *promote* complex goal-oriented collective actions, according to some literature [44]. Nevertheless, whenever communication mechanisms are available, learning to exploit them is a natural part of the autonomous mental development process In fact, that can both facilitate and improve outcomes of the development of effective and purposeful social ways of acting and can give to the system designer more control over such collective behavior, therefore more guarantees on expected outcomes.

III. APPLICATION DOMAINS

There are diverse application scenarios that can potentially take advantage of systems capable of autonomous mental development, at the individual and/or collective level.

A. ROBOTICS

Robotics is the area which first identified the profitability of building robots cables of autonomously building "by experience" a model of their own capabilities and consequently learning to achieve specific goals [13], [78].

In general, designers define a formalized model of their robots and can easily wire such model directly into embedded software without necessarily having the robot learn it in autonomy. However, autonomous learning may become necessary when the robot gets damaged while in operation, sometimes even partially losing its original capabilities. In that case, the robot should autonomously learn a new model to understand what it can do according to its residual operational capabilities, and how it can re-learn to achieve goals with them [13].

A different situation is that of modular robots [92]. There, the robot can reconfigure its shape to serve different tasks, and having the model designer foresee all possible shapes can be time-consuming and prevent emergent (not previously envisioned) shapes – functional to peculiar tasks – to be identified. In this case, having the modular robot try to assume a variety of forms and understand its action capabilities for the different shapes could be very useful before deployment (other than after deployment, to recover from injuries and from the loss of some of its modules).

At the level of collective robotics (i.e., a group of robots living together in an environment and having to cooperate to achieve a collective task), most current approaches rely on coordination schemes defined at design time. However, it has been argued that the autonomous evolution of communication and coordination capabilities can be of fundamental importance to acquire the capability of the collective to act in unknown and dynamically changing scenarios [18].

B. SMART FACTORIES

Similarly to a collective robotic system, a complex manufacturing system can be seen as an aggregated group of components that act together in order to achieve a production goal. Besides their basic scheme of functioning, defined at design time, if one component of the manufacturing system breaks or has some unexpected behavior, the manufacturing system should ideally adapt to the new situation, so as to overcome the problem without undermining production.

Exploring in advance all possible contingencies and hard-wiring possible solutions to them in the system is almost impossible. The system should rather learn autonomously how to act upon changing conditions (whether of a temporary or ultimate nature) to maintain its overall functioning. For example, the system may explore the possibility of deviating the material flow from the broken component to another one, to learn the effect of such actions, and its overall impact on the production. In doing so, a local or global production re-scheduling might be necessary, with the initiative collectively coming from the different components of the manufacturing system and without involving the production planning office. Clearly, such capabilities of autonomous adaptation can also apply at the level of individual components of the system, whenever these adaptations do not impact other components.

The need for integrating adaptability and flexibility in manufacturing systems is explicitly recognized as a key challenge in Industry 4.0 initiatives [94], and some examples of agent-based production control systems – exhibiting some limited forms of adaptivity – can be already found, e.g., in the automotive industry [16]. However, these are still far from the adaptivity level that could be reached by fully-fledged autonomous mental development approaches.

C. SMART HOMES

Buildings and homes are increasingly being enriched with sensors and actuators (i.e., IoT devices, in general), to facilitate our interactions with the environment and, by monitoring our activities and habits, to increase our safety and comfort [2]. However, such systems exploit design-time decisions w.r.t. deployment of devices, their interactions, and the types of services to be provided. Learning capabilities are typically limited to monitoring user activities and adapting the parameters of services (e.g., the levels of light and heating) accordingly [91].

From the perspective of autonomous mental development, we envision that once IoT devices are deployed in an environment, they should be activated in order to autonomously explore their own individual and collective capabilities (i.e., the individual sense of agency and the impact of inter-device interactions). This will enable them to eventually learn how they can affect the home environment and how, and apply such capabilities once users start populating the environment. Then, the overall smart home/building system will continue to dynamically and continuously modify its functioning to adapt to the presence of different users with different profiles, of users whose habits tend to evolve over time, or simply to react to contingencies (e.g., modifications in the number and type of available devices, or structural modifications to the environment).

We have conducted some simple preliminary experiments to show the potential feasibility of such a vision in a simple two-rooms smart home testbed [46]. In particular, we have shown that IoT devices in a room, when left free to explore the effects of their actions, can eventually build a sound causal model of the room, and can use such a model to actuate specific environmental conditions in that room. Also, by merging the models built for the two rooms, devices can learn to cooperatively actuate specific house-level environmental conditions. Exploring larger and more complex environments and different learning techniques, also in the presence of users, will give us better clues on the general applicability of the approach and its possible limitations.

D. SMART CITIES

Most of the considerations made above for smart home and building scenarios can, in theory, be transferred to the larger scenario of smart cities. That is, to all those ICT and IoT systems that can be deployed to automate and regulate the activities of modern cities, e.g., mobility, energy management, and garbage collection. Indeed, the need to deploy robust systems capable – with limited design efforts and limited human intervention – of dynamically adapting their behavior to continuously changing urban conditions is widely recognized as a key challenge for harmonic and sustainable urban development [80].

The substantial difference between smart homes/buildings and smart cities is that cities already exist and are already inhabited. Thus, one cannot think to let a smart city system free to explore the effect of its actions and interactions to eventually become capable of acting in a goal-oriented way (which you can instead do before a new building becomes inhabited).

However, one can think of exploiting a simulation-based approach to this purpose. Given that accurate and reliable simulators exist to study different urban aspects, one can think of having system components explore and learn in a simulated environment towards full mental development, before being eventually deployed in the real world [89].

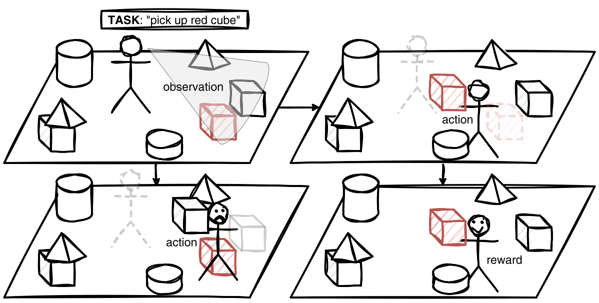


FIGURE 2. Reinforcement learning in a nutshell.

IV. STATE OF THE ART: OPPORTUNITIES AND CHALLENGES

The idea of autonomous mental development, at both the individual and collective level, has been widely investigated in areas such as cognitive psychology, neuroscience, and philosophy, but has been also argued as a key property that general artificial intelligence approaches should feature [64], [86], especially with the recent advent of a new breed of autonomous/cognitive agents based on generative AI technologies [84], [87].

We hereby focus on the computational perspective, and in particular on a set of selected approaches (Figure 3) that we think can contribute to attack and implement different aspects of the overall autonomous mental development vision, thus making it possible to turn the vision into a reality. However, although these approaches will play a fundamental role in future implementations, several challenges have still to be addressed to become practical tools for future autonomously developing systems.

We do not focus here on the basic levels of individual development, i.e., perception and embodiment, in that tools already exist to give agents sophisticated sensing abilities (e.g., convolutional neural networks to recognize objects, scenes, and activities [42]) and the capability of controlling their own actuators purposefully [51].

A. REINFORCEMENT LEARNING

The broad area of reinforcement learning shares with our vision the objective of designing machines (i.e., agents) capable of autonomously learning to act in a goal-oriented way in a specific context [36]. The key idea is, given a goal implicitly assigned to an agent, to provide the agent with "rewards" measuring the goodness of the actions it performs toward the goal achievement (Figure 2). By properly balancing purely exploratory actions and actions directly focused on maximizing rewards, the agent eventually learns how to achieve goals (i.e., it learns a *policy* guiding which actions to apply in which situations).

Approaches based on reinforcement learning, and in particular those based on deep learning [55], have indeed recently achieved amazing results, such as agents capable of achieving complex goals in complex environments. However, the majority of current research efforts do not aim at building systems with an explicit sense of agency and capable of developing an interpretable (i.e., symbolic) world model: the policy learned by the agent is *model-free* (i.e., sub-symbolic), hence it tells nothing to the agent about the effects of its actions, being only finalized at achieving a very specific goal. This makes most approaches highly ineffective in transferring the learned policy to scenarios where the goals assigned to agents change over time when the agent is immersed in a different environment, or simply in an ever-changing environment.

Curriculum-based approaches to machine learning go somewhat in the direction of gradually developing the capability to act in complex scenarios [12]. The agent is first trained on simple tasks, and the gained knowledge (i.e., its policy) is accumulated and exploited in increasingly complex scenarios, where further skills can thus be effectively learned. Yet again, most of these approaches are model-free and do not focus on the development of a world model and of an explicit sense of agency.

Reinforcement learning approaches based on *intrinsic* rewards [73], instead, more closely exploit the idea of exploring the world to develop a sense of agency. While in traditional reinforcement learning rewards are *extrinsic*, i.e., designed by a "teacher" as the scores in a videogame, intrinsic rewards are developed by the agent itself to satisfy

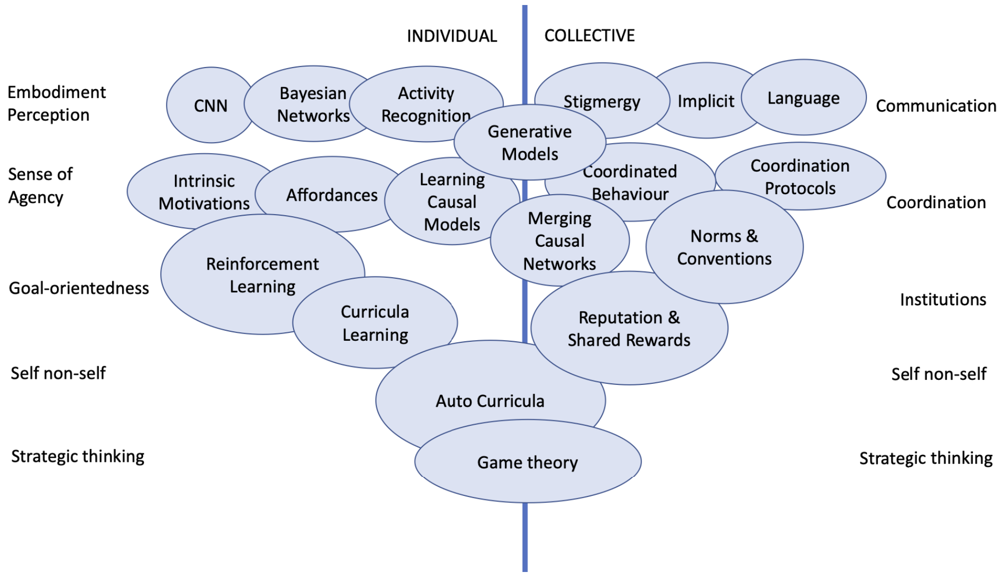


FIGURE 3. The galaxy of autonomous mental development.

its curiosity, i.e., when it autonomously discovers how to achieve specific tasks. For example, in [15] intrinsic rewards are computed as the error in forecasting the consequence of the action performed by the agent given its current state.

In general, one can think of splitting the concept of the environment into an *external* environment, with which the agent interacts through action and perception, and an *internal* one, which basically represents an interaction with the self. This distinction allows to model both intrinsic and extrinsic motivations within the same framework (i.e., through the interaction with some environment). Yet, we remark that intrinsic rewards do not necessarily derive from interactions with the internal environment. Some intrinsic rewards, in fact, are direct consequences of the concept of *surprise* or *novelty*, as felt by the agent when observing something in the external environment. In the well-known exploration vs. exploitation trade-off that is typical of reinforcement learning, intrinsic rewards can often be seen as an incentive for agents to perform exploration, and thus to improve their knowledge of the world and of their sense of agency.

Recent approaches based on the theory of affordances [37] propose to have agents gradually learn the effects of their actions: by having them act in constrained environments where only a limited set of actions apply, they eventually develop an explicit model of how their individual actions affect the environment. That is, they develop a sense of agency, which can be later exploited to build policies for the achievement of specific goals.

Generally speaking, reinforcement learning can be a useful tool to support the conceptual framework of autonomous mental development. It follows the same principle of a gamer exploring a new video game: the selection of a flow of actions balanced between *exploration* and *exploitation*, the continuous learning and refinement of strategies operated by the gamer facing difficulties in the game, the building of an internal *knowledge* about the game and the set of *strategies* and *behaviors* (at the individual and collective level) needed to achieve the final goal set by designers of that particular game. Nevertheless, to bring the autonomous mental development closer to reality, is necessary to go beyond the perimeter of a specific game or a specific task the agent is designed for, to give him the ability to *generalize* and continuously *adapt* his behavior to several contexts and situations it might be immersed to, exactly as humans do.

In summary, all the above reinforcement learning approaches attack the key challenge of building general tools to learn how to effectively achieve goals in an environment. At the same time, though, it is increasingly recognized that model-free approaches can hardly be used to let agents autonomously develop generalizable and interpretable policies of action, which would instead require the agents to develop an explicit sense of agency, i.e., a symbolic causal model of the world.

B. CAUSAL MODELS

Understanding and leveraging *causality* is recognized as a key general challenge for AI in the coming years [74]. In particular, Pearl [65] has proposed the idea of a "causal hierarchy" (also named "ladder of causation") to define different levels of causality recognition and exploitation by an intelligent agent (Figure 4).

The first level of the ladder consists of simply detecting relations as associations (correlations), whereas the second one assumes the possibility to intervene in the environment (such as in reinforcement learning) and observe the (causal) effects of the actions taken. Finally, the third level enables reasoning and planning on the basis of counterfactual analysis. Such layers correspond to some of the phases of

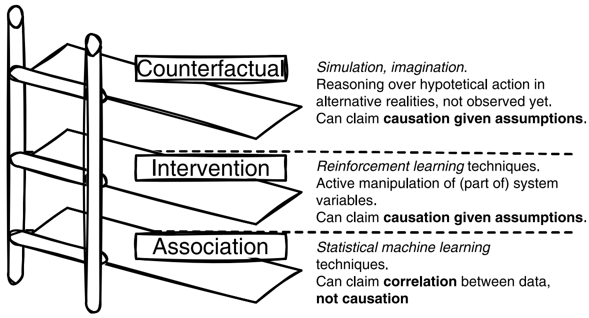


FIGURE 4. Judea Pearl's "ladder of causality" [66].

the autonomous development loop we defined: the first one is mostly involved in the perception phase, whereas the second one is associated with the development of a sense of agency and to recognition of self and non-self. The final layer clearly enables goal-oriented behavior, strategic thinking, and collective coordination.

Graphical models such as Bayesian and causal networks are among the ones that are most widely exploited in order to build interpretable models of the world. Although traditional Bayesian networks do not properly represent causal links, but just conditional dependence relations between random variables, they can be easily extended towards identifying causal relations so that a link between two variables "A is related to B" can be eventually recognized as a causal link, i.e., "A causes B" [27]. In other scenarios, domain knowledge can be exploited in order to assume that some relations depicted in a Bayesian network actually represent causal links, e.g., because some of the variables represent the status of some actuators (causes) and others the state of some sensors (effects) [46].

It is worth remarking that causality can be learned from observational data alone but only with restrictive assumptions about the data generation process and resulting distributions, and about the ground truth model to be learnt. Instead, it can be inferred with more relaxed assumptions after experimental intervention. By freezing the values of some variables, counterfactual analysis can be applied in order to observe the behavior of the other variables: only in this way, causal relations can be discovered.

Regarding the autonomous mental development vision, causality learning and counterfactual analysis can be a very useful tool. Indeed, they can be exploited to build a generic model of the world, as well as to build effective action planning towards an individual or shared goal, fixing the wanted value of a target variable and understanding how to tune the causes to obtain the desired outcome. Moreover, observing how other agents behave and tuning variables to obtain desired effects can support the strategic thinking phase of autonomous development, strengthing the cycle at the individual level. At the collective level, instead, causality can support communication, coordination and institutions, since it helps an agent understand what request made to the social community brings to a desired outcome, what requests are therefore needed to achieve a desired collective goal, and as a final instance, what set of shared actions bring effectively in time to achieve a desired state of the affairs.

A recent contribution that is in line with the ideas we envision for autonomous development is the application of curriculum learning to the problem of learning the structure of Bayesian networks [97], or even causal networks. On a pure sub-symbolic level, on the other hand, another recent work proposes to learn causal models in an online setting [34], with the aim of finding (and strengthening) causal links between input and output variables.

We argue that key challenges in this area concern, again, understanding how to synergistically exploit symbolic and sub-symbolic model-free approaches to learn, represent, and evolve causal models in autonomous development scenarios, and how to use them to adaptively achieve goals.

C. MULTI-AGENT REINFORCEMENT LEARNING AND EVOLUTIONARY APPROACHES

When multiple agents act in a shared environment, their actions and their effectiveness in achieving goals are affected by what others do. Game-theoretic approaches to strategic thinking and their extension to population dynamics (i.e. Evolutionary Game Theory) [31] have deeply investigated this problem and the decision-making processes behind it. In this context, it has also been shown that agents can effectively learn in autonomy to improve their performance in dealing with others [62].

However, when moving from theoretical settings (e.g., the prisoner's dilemma) to complex and realistic scenarios where agents have complex goals, peculiar phenomena arise. The more one agent learns, the more it challenges others, triggering a continuous increase in the complexity of behavior, ultimately enabling it to incrementally learn more sophisticated means to act. This somewhat resembles the increase of complexity that agents face in curriculum approaches to reinforcement learning. The key difference is that, in the presence of multiple agents, the increase in complexity and capabilities of agents is promoted and self-sustained by the system itself, hence the term *autocurricula* [44].

Recently, autocurricula-based approaches have produced stunning results in multiagent environments, both cooperative and competitive. For instance, in a hide-and-seek scenario [5], agents moving in a complex simulated environment have learned how to effectively compete (hiders against seekers) and cooperate (coalitions of cooperating seekers/hiders) in very elaborated ways, in a continuous self-sustained learning process. Indeed, we consider such approaches fundamental towards the autonomous development of complex agent societies. Yet, a deep understanding of the process that drives the evolution of individual and collective behaviors is still missing and is a key challenge for the next few years. To this end, providing agents with an *explicit* modelling (possibly in *causal* terms) of the others' behavior and of the overall societal behavior, may be necessary [79]. Also, autocurricula approaches do not currently account for the possibility of explicitly interacting (e.g., through speech acts) with other agents, which may prove fundamental to improve collective learning and "crystallize" useful behaviors into coordination protocols.

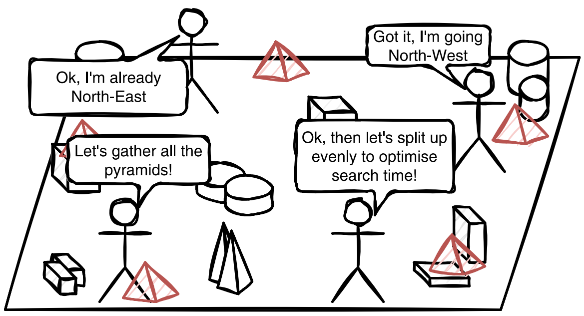


FIGURE 5. Learning to coordinate in a multi-agent system.

D. LEARNING TO COMMUNICATE AND COORDINATE

As already mentioned, agents may communicate and coordinate: by explicit messages [40], by leaving traces in the environment [48], or implicitly [50].

These forms of communication are already exploited in multiagent learning, mostly to improve the individual learning process by letting agents share information (e.g., for merging their individual causal models of the world [53]) and coordinate actions [59]. However, these communication approaches are usually assumed as *innate* capabilities of agents, rather than one to be learnt. That is, agents have an *a-priori* sense of agency with respect to communication actions, whereas in our vision it should be developed by learning (Figure 5).

For example, with reference to explicit communication acts, [30] proposes a voting game to let agents learn to share a communication language and to develop a strategy to communicate. In [28], it is shown that reinforcement learning can be effectively applied to let agents learn how to communicate in order to achieve a specific effect. Or, in [11] it is shown how more traditional prediction techniques, aimed at guessing the right sequence of interactions among agents (or compositions amongst services), can be applied to learn to coordinate effectively based on past experience.

In the case of *implicit* communication, instead, forms of implicit behavioral communications have been shown to emerge in simple system components that purposefully move in an environment [29], as they learn to affect others with *ad-hoc* actions. Learning to use stigmergy to effectively coordinate is under-explored in the literature, which instead focuses on the opposite – using stigmergy to boost learning. Nevertheless, some early experiments show that it appears indeed possible to learn stigmergic forms of communication [51]. On the opposite, stigmergic mechanisms appear underexplored in the literature, yet, as are the general mechanisms that can let a population of agent learn to use environmental or behavioral signals to coordinate their behavior [3], [75].

In any case, the development of general approaches to let agents develop fully-fledged forms of communication and coordination is still under-explored. Doing so may call for agents to develop not only a model of the world but an overall model of the society supporting the autonomous development at the collective level, i.e., a social sense of agency explicitly modelling how communications and coordination actions affect other agents in the shared environment.

Coordination approaches built on the notion of *social commitment* [6] may be interpreted as somewhat going in this direction, as they are concerned with modelling the pre-requisites and effects that social actions (such as communications) require to the action performer and expect on the action receiver, respectively. Hence, similarly to how causal relationships between actuators and sensors are built by modelling the structural relationships between actions and their observable effects (as perceived by sensors), commitments model the relationships between interacting agents in terms of social actions (cause) and expectations (effect). Nevertheless, these approaches are not currently exploring the possibilities of *autonomously learning* such commitments and their effects, as they are concerned with giving an external designer the means to engineer interaction protocols at design time.

E. EMERGENCE OF INSTITUTIONS

Whereas learning to communicate is about understanding how to use communication to coordinate actions with others, enabling and sustaining global collective achievement of goals requires "institutionalized" means of acting at the collective level, i.e., a set of shared beliefs and of shared social conventions and norms aimed at ruling collective actions [26].

The mechanisms leading to the spontaneous emergence of institutions in human society [63], including the mechanisms to promote and sustain altruistic and cooperative behavior (e.g., reputation and shared rewards [61]) have been widely investigated [60]. However, most approaches to building multiagent systems assume such mechanisms as *explicitly designed* [26].

Yet, some promising studies related to the emergence of institutionalized behaviors in multiagent systems have been undertaken (see [57] for a recent survey). Such a research topic is concerned, on the one hand, with understanding the mechanisms and the necessary conditions that make social norms arise in agent societies, and, on the other hand, with understanding how to individually engineer agents so that norms may arise spontaneously without the need for prescriptive approaches defining the norms beforehand, at design time.

In the autonomous development perspective, the spontaneous emergence of recurrent social behavior supports the collective level loop, letting agents learn how to communicate, coordinate and, as the last step, how to correctly adhere to emerging norms. In this context, a social norm can be described as "the belief that a given behavior is prescribed within the population". Such a belief may be acquired by an agent in a variety of ways, e.g. by observing peers' actions, by asking for / inferring their strategy behind an action, or by learning from past experience through repeated coordination games [54]. For instance, [93] proposes a collective learning framework where agents learn to adopt norms in repeated coordination. In this way, agents eventually *learn* that a social norm has emerged, and "institutionalize" their behavior in their (social) decision-making processes *implicitly*, by behaving so as to comply with the norm. Another interesting work [8] integrates rational thought, reinforcement learning, and social interactions to model norms emergence in a society: agents incrementally develop a social behavior (a social norm) while *internalizing* it within their cognitive model. The emergence of a shared norm and its institutionalization, naturally let agents recognize the existence of others in the self/non-self step. Then, due to its cyclic nature, the process is eventually repeated, letting the agent refine its collective or individual behavior concerning its purpose as individual or as part of the community.

However, most of current research efforts only consider *implicit representations* of norms, that is, do not bother in making agents capable of explicitly modeling an emergent norm as soon as agents somewhat recognize its existence, so as to exploit such representation in future reasoning. This deprives agents of a useful model of the social world, that could help shape their social sense of agency in terms of the causes and effects of their social actions on other agents in the society, as defined by norms themselves. Also, the development of general models and tools to support the proper learning and evolution of institutionalized mechanisms of coordination, through the construction of explicit norms representations and their adoption by agents within their cognitive models is still missing, and so are the solutions to the many problems involved in this process. For instance: how to avoid that an agent learns that free-riding is better than abiding by norms? How to avoid inconsistencies and misunderstandings in norms interpretation? Can the collective level loop of the autonomous development framework be enough to balance such challenges?

F. GENERATIVE TECHNOLOGIES

Generative AI focuses on creating "original" content. This primarily includes textual documents (there including software code), but is now expanding to images, video, and audio [85]. It operates primarily in three phases: *training* of a "large" *foundation model* [88], that is, a huge deep neural network architecture (billions of parameters) trained on a huge corpus of heterogeneous documents; *tuning* of the model, that is, refining its training by tailoring it to the specific application domain and intended usage; *usage and feedback*, where consumer applications or end users exploit the model while (explicitly or not) giving positive or negative feedback in return (used to evaluate and improve the model).

Given the relevance of language for humans, in terms of both reasoning ("inner monologue") and verbal and written communication, it is natural that especially Large Language Models quickly started to pervade any application domain in which textual understanding and text generation are either the objective of the domain (e.g. contents writing) or a preferred means of interaction with human users (e.g. chatbots). Then Multi-Modal models followed, for their ability to work with visual data, another cornerstone sensory modality for humans.

Accordingly, it is already apparent that generative AI can play a central role in autonomous mental development at least in the perception and communication "steps". For the former, the capability of interpreting multi-modal inputs can greatly enhance the ability of intelligent software agents to perceive their environment and make sense of complex situations, by seamlessly merging multiple data streams. Also, they can learn to recognize novel situations while in operation (via fine-tuning and in-context learning), again from multi-modal sources. For the latter, not only software agents can now interact more naturally with humans via text, speech, and visual input/output, but they could leverage the same capability to interact with other agents as well. Systems where multiple generative agents self-coordinate to achieve collective goals are already appearing [22], and they complement the traditional communication and coordination protocols (such as those conceived and developed in the multi-agent systems community [9]) with more flexible and fluid conversations, often resembling human negotiation. A generative AI can even dynamically create new software agents (by producing code instead of narrative text), to delegate them specific tasks.

It is also worth emphasising that being language and vision two especially prominent factors in human-like reasoning, they can effectively facilitate and support planning actions toward the achievement of a goal [82], and coordinating activities based on capabilities, requirements, and roles [22]. This adds sense of agency and goal-orientedness to the "steps" of our proposed conceptual framework that generative AI can impact. For instance, a debate is currently unfolding about the capability of generative AI technologies to perform causal reasoning [32], [35], [81]: confirming this achievement would be a breakthrough as this is a cornerstone of autonomous mental development in humans, underlying planning abilities and goal achievement.

V. HORIZONTAL CHALLENGES

The presented approaches and techniques are still at the research stage, and many research challenges have been identified for each of them in the previous section. Moreover, it is possible to identify several additional "horizontal" challenges, i.e., of a general nature independently of the specific approach.

The nature of such challenges, in our opinion, makes them specifically suited for being pursued by the most diverse research communities in the area of machine learning, reinforcement learning, autonomous agents and multi-agent systems.

A. ENGINEERING

Many of the presented approaches are grounded in machine learning, a discipline with plenty of years of research behind it, but in which good engineering practice is still often disregarded, and traditional software engineering problems are sometimes considered mundane. Systems are often developed *ad-hoc* for a specific task or problem domain, with little attention to modularity, reusability, and dependability, thus missing the flexibility to adopt them across different domains, tasks, datasets [67]. In addition, given that the diverse approaches presented can each contribute important pieces to the overall vision of autonomous development, sound engineering approaches are needed to try to integrate such a heterogeneous plethora into a coherent whole.

A notable example of the difficulty of providing reusable, robust, and fully controllable integrations between machine learning-based approaches and traditional software engineering paradigms (e.g. the agent-oriented one) is given by recent literature about generative AI-based autonomous/cognitive agents [84], [87]. These represent multi-faceted and horizontal research challenges that, in our opinion, could and should be profitably attacked by the those research communities interested in the engineering of complex systems.

B. CONTROLLING EVOLUTION

Autonomous development raises the issue of somewhat controlling how behaviors evolve, as an individual learns new skills and tasks, and as the collective learns new ways of coordinating and acting together [44]. How can we *steer* a learning process towards desired outcomes without putting bias in it? How can we *constrain* the boundaries within which individual and collective behaviors should stay (e.g., in terms of safety)? What *interventions* can we make to re-direct an agent or a collective that has taken an unpredictable or unsafe development path?

Again, generative AI is a recent notable example of the difficulty of answering these questions, as ‘‘hallucination‘‘, bias, and misinformation are proving resistant to any attempt to mitigate them. Experience in self-adaptive components based on feedback, as well as in the study of emergent behaviors in self-organising systems can help in finding proper technical answers, and – why not – *ethical* ones [90].

C. INNATE VS. ACQUIRED KNOWLEDGE

Besides the trade-off between autonomous learning and ‘‘safety boundaries’’, that should be preserved no matter what, another trade-off needs to be set for autonomous development approaches to suit the real world: what should be learned from scratch and what could be assumed as given? In other terms, what is a realistic and effective trade-off between innate and acquired knowledge? The answer is difficult to give as it depends on many factors and is highly domain-dependent [7].

It can be argued that, in most domains, knowledge about the available sensory and actuation capabilities (intended both in a physical or purely computational world) could be taken as innate knowledge that agents have either since design or automatically acquired upon deployment: this would account for embodiment and perception in our conceptual framework. The same could be said for basic communication skills, both direct and indirect, such as sending and receiving messages or depositing and reacting to signals in a shared environment. However, the ability to compose such basic skills into coordination patterns and protocols, or even further to collectively and explicitly construct shared norms out of them could be left to appropriate learning routines. Similarly, at the individual level, the ability to operate on actuators to affect the environment, and the ability to do so as part of a sequence of activities meant to achieve a prescribed goal, can be delegated to autonomous development. In any case, the methodologies and tools to precisely assess the impact of such choices on the process of autonomous development itself are still missing.

D. HUMANS IN THE LOOP

The more autonomous development technologies will advance, the more humans will have to actively interact with them. This interaction will raise technical issues (Will we have ‘‘handles’’ to control or block such systems in some ways and to some extent? Could large language models be the key to interact with and control such systems?) and ethical problems (Will we be rather ‘‘handled’’ by these systems and subjects to their decisions?). Some of these problems already emerged, like in the *moral machine* experiment [4] or in AI-based hiring technology.

Technical challenges nurture research in the HCI and distributed systems communities (there included the self-organising and self-adaptive systems). Ethical and moral ones will be meat for politicians and lawyers, although deep joint work with technical experts will always be necessary. A key ingredient involves institutions since they represent humans as a group: laws and regulations need to be developed to regulate the global actors in day-by-day technology usage. Nevertheless, a deeper interaction between researchers in science and technology and public institutions is needed to support the regulation design phase.

E. SUSTAINABILITY

Algorithms for autonomous development will most likely require extensive computational resources. For example, the mentioned ‘‘hide and seek’ experiment by OpenAI involved a distributed infrastructure of 128,000 preemptible CPU cores and 256 GPUs on GCP [5]: the default model optimized over 1.6 million parameters taking 34 hours to reach the fourth stage over six of agents skills progression. Generative AI technologies make these numbers seem like a joke: OpenAI involved the use of 25,000 NVIDIA A100 GPUs to train GPT-4. Each server with these GPUs consumes about 6.5 kW, leading to an estimated energy usage of 50 GWh during the training phase. For context, all data centres in Sweden currently consume about 3,000 GWh of energy. The energy used for a single training session of GPT-4, which is 50 GWh, would represent approximately 2% of this total capacity. OpenAI also disclosed that GPT-3 runs on a server cluster that uses a total of 91 GWh per year for serving with GPT-4.

These examples are a sort of best-in-class project; anyway, it is clear that if self-developing systems will be based on similar learning approaches, they will require massive amounts of computational resources. Therefore, a key challenge for the community will be to devise algorithmic and system-level means to make autonomous development systems sustainable, and affordable by others other than the big technology players.

F. EXPLAINABILITY

Being able to inspect and explain the decision-making process of AI systems is already a hot topic, so much so that the entire research field of XAI, eXplainable AI, has been born [25].

We already commented several times on how such problems should be accounted for also for autonomous development, possibly with the help of causal models. This is indeed a key challenge for self-adaptive and, especially, for self-organising research, too, where explaining global behaviors, patterns, and configurations emerging from local interactions is mostly still considered the ''holy grail''.

G. PERFORMANCE ASSESSMENT

Properly assessing the performance of autonomous development is another challenge per se. It is not straightforward to identify well-established benchmarks and metrics that cover the wide spectrum of properties and skills that developmental agents need to acquire as they evolve. Yet, some inspiration can come from related fields. For example, recently there has been an effort to develop benchmarks for autonomous agents interacting with a given environment, even with natural language specifications [68].

Similar ideas come from the related field of curriculum learning and continual learning, where novel benchmarks are designed to capture also ad-hoc metrics, such as forgetting (i.e., avoiding forgetting old tasks and skills when learning new ones), memory overhead and computational efficiency [23], [76]. Other metrics coming from the area of reinforcement learning, like average reward and time (i.e., episodes) needed to fulfil a task might be appropriate [70].

VI. CONCLUSION

In this paper, we have elaborated upon the vision of autonomous development, at both the individual and collective level. Although the road towards fully realizing the vision is still a long one, several ideas in the areas of learning, causality, and multiagent systems, are already showing its potential feasibility.

From our side, with the aim of starting implementing some of the concepts presented in the paper, we are currently experimenting with Bayesian networks and causal models to learn dependencies between variables that represent sensors and actuators within a smart environment. In a simplified smart home setting, we showed how an agent is able to learn the effect of one of its own actions, thus acquiring the sense of agency, the necessary precondition towards goal-orientedness [46]. The training set consists of a collection of observations where the agent performs random actions and observes their effect on the rest of the environment. Once the learning phase is completed, the agent is eventually able to understand what to do to reach the desired state of affairs. At the collective level, our preliminary experiments show how different agents are able to learn to cooperate to achieve a goal they could not achieve individually. We assumed that the agents could share their observations, thus providing training examples to a single data set that can be used to learn a single, general model. By learning from the joint set of observations and actions, the two agents learn that they need to cooperate and coordinate their actions.

As a continuation of this strand of research, we are now moving to a distributed learning setting, where agents do not fully share their observations to agree on a single global (causal) model of their shared environment [52]. Rather, they cooperate to refine their own local causal models whenever they recognize partial, missing, or wrong information, by organizing a coordinated distributed intervention protocol meant to obtain the additional information needed to disambiguate, refine, complete, or correct their own local models.

As part of our future work, we plan to investigate how digital twins could enable the learning paradigms described so far [47]. In particular, in many application domains such as smart factories, one could envisage a hierarchical architecture where digital twins collect and integrate data coming from heterogeneous physical devices, building more and more abstract models and representations. In addition, as from Subsection V-G, we plan to work towards identifying suitable metrics to assess and test the effectiveness of our proposals.

REFERENCES

[1] K. Ahuja, D. Mahajan, Y. Wang, and Y. Bengio, ''Interventional causal representation learning,'' in *Proc. Int. Conf. Mach. Learn.*, 2023, pp. 372–407.
[2] M. Alaa, A. A. Zaidan, B. B. Zaidan, M. Talal, and M. L. M. Kiah, ''A review of smart home applications based on Internet of Things,'' *J. Netw. Comput. Appl.*, vol. 97, pp. 48–65, Nov. 2017.
[3] R. Aras, A. Dutech, and F. Charpillet, ''Stigmergy in multiagent reinforcement learning,'' in *Proc. 4th Int. Conf. Hybrid Intell. Syst. (HIS)*, Dec. 2004, pp. 468–469.

[4] E. Awad, S. Dsouza, R. Kim, J. Schulz, J. Henrich, A. Shariff, J.-F. Bonnefon, and I. Rahwan, "The moral machine experiment," *Nature*, vol. 563, no. 7729, pp. 59–64, Nov. 2018.

[5] B. Baker, I. Kanitscheider, T. Markov, Y. Wu, G. Powell, B. McGrew, and I. Mordatch, "Emergent tool use from multi-agent autocurricula," 2019, *arXiv:1909.07528*.

[6] M. Baldoni, C. Baroglio, R. Micalizio, and S. Tedeschi, "Programming agents by their social relationships: A commitment-based approach," *Algorithms*, vol. 12, no. 4, p. 76, Apr. 2019.

[7] E. Bates, J. L. Elman, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett, "Innateness and emergentism," in *A Companion to Cognitive Science*. Hoboken, NJ, USA: Blackwell Publishing Ltd, 2017, pp. 590–601.

[8] R. Beheshti, "Normative agents for real-world scenarios," in *Proc. Int. Conf. Auton. Agents Multi-Agent Syst.*, 2014, pp. 1749–1750.

[9] F. Bellifemine, A. Poggi, and G. Rimassa, "Developing multi-agent systems with a FIPA-compliant agent framework," *Softw., Pract. Exper.*, vol. 31, no. 2, pp. 103–128, Feb. 2001.

[10] K. Bellman, J. Botev, A. Diaconescu, L. Esterle, C. Gruhl, C. Landauer, P. R. Lewis, P. R. Nelson, E. Pournaras, A. Stein, and S. Tomforde, "Self-improving system integration: Mastering continuous change," *Future Gener. Comput. Syst.*, vol. 117, pp. 29–46, Apr. 2021.

[11] H. Ben Mahfoudh, G. Di Marzo Serugendo, N. Naja, and N. Abdennadher, "Learning-based coordination model for spontaneous self-composition of reliable services in a distributed system," *Int. J. Softw. Tools Technol. Transf.*, vol. 22, no. 4, pp. 417–436, Aug. 2020.

[12] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009, pp. 41–48.

[13] J. Bongard, V. Zykov, and H. Lipson, "Resilient machines through continuous self-modeling," *Science*, vol. 314, no. 5802, pp. 1118–1121, Nov. 2006.

[14] S. Bubeck, V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. T. Lee, Y. Li, S. Lundberg, H. Nori, H. Palangi, M. T. Ribeiro, and Y. Zhang, "Sparks of artificial general intelligence: Early experiments with GPT-4," 2023, *arXiv:2303.12712*.

[15] Y. Burda, H. Edwards, D. Pathak, A. Storkey, T. Darrell, and A. A. Efros, "Large-scale study of curiosity-driven learning," in *Proc. Int. Conf. Learn. Represent.*, 2018, pp. 1–15.

[16] S. Bussmann and K. Schild, "Self-organizing manufacturing control: An industrial application of agent technology," in *Proc. 4th Int. Conf. MultiAgent Syst.*, Jul. 2000, pp. 87–94.

[17] G. Butterworth, "Origins of self-perception in infancy," *Psychol. Inquiry*, vol. 3, no. 2, pp. 103–111, Apr. 1992.

[18] N. Cambier, R. Miletitch, V. Frémont, M. Dorigo, E. Ferrante, and V. Trianni, "Language evolution in swarm robotics: A perspective," *Frontiers Robot. AI*, vol. 7, p. 12, Feb. 2020.

[19] L. Canese, G. C. Cardarilli, L. Di Nunzio, R. Fazzolari, D. Giardino, M. Re, and S. Spanò, "Multi-agent reinforcement learning: A review of challenges and applications," *Appl. Sci.*, vol. 11, no. 11, p. 4948, May 2021.

[20] S. Casper et al., "Open problems and fundamental limitations of reinforcement learning from human feedback," *Trans. Mach. Learn. Res.*, pp. 1–42, Dec. 2023. [Online]. Available: https://jmlr.org/tmlr/papers/bib/bx24KpJ4Eb.bib

[21] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, H. Chen, X. Yi, C. Wang, Y. Wang, W. Ye, Y. Zhang, Y. Chang, P. S. Yu, Q. Yang, and X. Xie, "A survey on evaluation of large language models," *ACM Trans. Intell. Syst. Technol.*, vol. 15, no. 3, pp. 1–45, Jan. 2024.

[22] W. Chen, Y. Su, J. Zuo, C. Yang, C. Yuan, C.-M. Chan, H. Yu, Y. Lu, Y.-H. Hung, C. Qian, Y. Qin, X. Cong, R. Xie, Z. Liu, M. Sun, and J. Zhou, "AgentVerse: Facilitating multi-agent collaboration and exploring emergent behaviors," 2023, *arXiv:2308.10848*.

[23] N. Díaz-Rodríguez, V. Lomonaco, D. Filliat, and D. Maltoni, "Don't forget, there is more than forgetting: New metrics for continual learning," 2018, *arXiv:1810.13166*.

[24] Z. Durante et al., "An interactive agent foundation model," 2024, *arXiv:2402.05929*.

[25] R. Dwivedi, D. Dave, H. Naik, S. Singhal, R. Omer, P. Patel, B. Qian, Z. Wen, T. Shah, G. Morgan, and R. Ranjan, "Explainable AI (XAI): Core ideas, techniques, and solutions," *ACM Comput. Surv.*, vol. 55, no. 9, pp. 1–33, Sep. 2023.

[26] M. Esteva, J. A. Rodriguez-Aguilar, C. Sierra, P. Garcia, and J. L. Arcos, "On the formal specification of electronic institutions," in *Agent Mediated Electronic Commerce*. Berlin, Germany: Springer, 2001, pp. 126–147.

[27] K. Fadiga, É. Houzé, A. Diaconescu, and J.-L. Dessalles, "To do or not to do: Finding causal relations in smart homes," in *Proc. IEEE Int. Conf. Autonomic Comput. Self-Organizing Syst. (ACSOS)*, Sep. 2021, pp. 110–119.

[28] J. N. Foerster, Y. M. Assael, N. de Freitas, and S. Whiteson, "Learning to communicate with deep multi-agent reinforcement learning," 2016, *arXiv:1605.06676*.

[29] N. A. Grupen, D. D. Lee, and B. Selman, "Low-bandwidth communication emerges naturally in multi-agent learning systems," 2020, *arXiv:2011.14890*.

[30] S. Gupta and A. Dukkipati, "Winning an election: On emergent strategic communication in multi-agent networks," in *Proc. Int. Conf. Auto. Agents Multiagent Syst.*, 2020, pp. 1861–1863.

[31] T. A. Han, "Emergent behaviours in multi-agent systems with evolutionary game theory," *AI Commun.*, vol. 35, no. 4, pp. 327–337, Sep. 2022.

[32] Z. Hu, Y. Zhang, R. Rossi, T. Yu, S. Kim, and S. Pan, "Are large language models capable of causal reasoning for sensing data analysis?" in *Proc. Workshop Edge Mobile Found. Models (EdgeFM)*, Tokyo, Japan, 2024, pp. 24–29.

[33] M. C. Huebscher and J. A. McCann, "A survey of autonomic computing—Degrees, models, and applications," *ACM Comput. Surv.*, vol. 40, no. 3, p. 7, 2008.

[34] K. Javed, M. White, and Y. Bengio, "Learning causal models online," 2020, *arXiv:2006.07461*.

[35] Z. Jin, Y. Chen, F. Leeb, L. Gresele, O. Kamal, L. Y. U. Zhiheng, K. Blin, F. G. Adauto, M. Kleiman-Weiner, M. Sachan, and B. Schölkopf, "Cladder: Assessing causal reasoning in language models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 36, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Eds., Red Hook, NY, USA: Curran Associates, 2023, pp. 31038–31065.

[36] L. P. Kaelbling, M. L. Littman, and A. Moore, "Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 4, pp. 237–285, May 1996.

[37] K. Khetarpal, Z. Ahmed, G. Comanici, D. Abel, and D. Precup, "What can I do here? A theory of affordances in reinforcement learning," in *Proc. 37th Int. Conf. Mach. Learn.*, 2020, pp. 5243–5253.

[38] J. Kirchenbauer, J. Geiping, Y. Wen, J. Katz, I. Miers, and T. Goldstein, "A watermark for large language models," in *Proc. Int. Conf. Mach. Learn.*, Jan. 2023, pp. 17061–17084.

[39] S. Kounev, P. Lewis, K. L. Bellman, N. Bencomo, J. Camara, A. Diaconescu, L. Esterle, K. Geihs, H. Giese, S. Götz, P. Inverardi, J. O. Kephart, and A. Zisman, "The notion of self-aware computing," in *Self-Aware Computing Systems*. Cham, Switzerland: Springer, 2017, pp. 3–16.

[40] Y. Labrou, T. Finin, and Y. Peng, "Agent communication languages: The current landscape," *IEEE Intell. Syst.*, vol. 14, no. 2, pp. 45–52, Mar. 1999.

[41] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman, "Building machines that learn and think like people," *Behav. Brain Sci.*, vol. 40, p. e253, Jan. 2017.

[42] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[43] R. Legaspi and T. Toyoizumi, "A Bayesian psychophysics model of sense of agency," *Nature Commun.*, vol. 10, no. 1, pp. 1–11, Sep. 2019.

[44] J. Z. Leibo, E. Hughes, M. Lanctot, and T. Graepel, "Autocurricula and the emergence of innovation from social interaction: A manifesto for multi-agent intelligence research," 2019, *arXiv:1903.00742*.

[45] M. Lippi, S. Mariani, M. Martinelli, and F. Zambonelli, "Individual and collective self-development: Concepts and challenges," in *Proc. 17th Conf. Comput. Sci. Intell. Syst. (FedCSIS)*, Sofia, Bulgaria, Sep. 2022, pp. 15–21.

[46] M. Lippi, S. Mariani, and F. Zambonelli, "Developing a 'sense of agency' in IoT systems: Preliminary experiments in a smart home scenario," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops Affiliated Events (PerCom Workshops)*, Mar. 2021, pp. 44–49.

[47] M. Lippi, M. Martinelli, M. Picone, and F. Zambonelli, "Enabling causality learning in smart factories with hierarchical digital twins," *Comput. Ind.*, vol. 148, Jun. 2023, Art. no. 103892.

[48] M. Mamei and F. Zambonelli, "Programming stigmergic coordination with the TOTA middleware," in *Proc. 4th Int. Joint Conf. Auton. Agents Multiagent Syst.*, 2005, pp. 415–422.

[49] G. Marcus and E. Davis, "Insights for AI from the human mind," *Commun. ACM*, vol. 64, no. 1, pp. 38–41, Jan. 2021.

[50] S. Mariani and A. Omicini, "Anticipatory coordination in socio-technical knowledge-intensive environments: Behavioural implicit communication in *MoK*," in *AI*IA 2015 Advances in Artificial Intelligence*. Cham, Switzerland: Springer, 2015, pp. 102–115.
[51] S. Mariani and F. Zambonelli, "Learning stigmergic communication for self-organising coordination," in *Proc. IEEE Int. Conf. Autonomic Comput. Self-Organizing Syst. (ACSOS)* Toronto, ON, Canada, Sep. 2023, pp. 47–56.
[52] S. Mariani and F. Zambonelli, "Distributed discovery of causal networks in pervasive environments," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops Affiliated Events (PerCom Workshops)*, Biarritz, France, Mar. 2024, pp. 1–6.
[53] S. Meganck, S. Maes, B. Manderick, and P. Leray, "Distributed learning of multi-agent causal models," in *Proc. IEEE/WIC/ACM Int. Conf. Intell. Agent Technol.*, Sep. 2005, pp. 285–288.
[54] M. Mihaylov, K. Tuyls, and A. Nowé, "A decentralized approach for convention emergence in multi-agent systems," *Auton. Agents Multi-Agent Syst.*, vol. 28, no. 5, pp. 749–778, Sep. 2014.
[55] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
[56] T. M. Moerland, J. Broekens, A. Plaat, and C. M. Jonker, "Model-based reinforcement learning: A survey," *Found. Trends Mach. Learn.*, vol. 16, no. 1, pp. 1–118, 2023.
[57] A. Morris-Martin, M. De Vos, and J. Padget, "Norm emergence in multiagent systems: A viewpoint paper," *Auto. Agents Multi-Agent Syst.*, vol. 33, no. 6, pp. 706–749, Nov. 2019.
[58] C. Müller-Schloer, "Organic computing: On the feasibility of controlled emergence," in *Proc. 2nd IEEE/ACM/IFIP Int. Conf. Hardw./Softw. Codesign Syst. Synthesis (CODES+ISSS)*, A. Orailoglu, P. H. Chou, P. Eles, and A. Jantsch, Eds. NY, USA: IEEE, Aug. 2004, pp. 2–5.
[59] T. T. Nguyen, N. D. Nguyen, and S. Nahavandi, "Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications," *IEEE Trans. Cybern.*, vol. 50, no. 9, pp. 3826–3839, Sep. 2020.
[60] M. A. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, Dec. 2006.
[61] M. A. Nowak and K. Sigmund, "Evolution of indirect reciprocity," *Nature*, vol. 437, no. 7063, pp. 1291–1298, Oct. 2005.
[62] A. Nowé, P. Vrancx, and Y. M. De Hauwere, "Game theory and multi-agent reinforcement learning," in *Reinforcement Learning*. Berlin, Germany: Springer, 2012, pp. 441–470.
[63] E. Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge, U.K.: Cambridge Univ. Press, 1990.
[64] P.-Y. Oudeyer, F. Kaplan, and V. V. Hafner, "Intrinsic motivation systems for autonomous mental development," *IEEE Trans. Evol. Comput.*, vol. 11, no. 2, pp. 265–286, Apr. 2007.
[65] J. Pearl, "The seven tools of causal inference, with reflections on machine learning," *Commun. ACM*, vol. 62, no. 3, pp. 54–60, Feb. 2019.
[66] J. Pearl and D. Mackenzie, *The Book of Why: The New Science of Cause and Effect*. NY, USA: Basic Books, 2018.
[67] B. Porter and R. Rodrigues Filho, "Distributed emergent software: Assembling, perceiving and learning systems at scale," in *Proc. IEEE 13th Int. Conf. Self-Adaptive Self-Organizing Syst. (SASO)*, Jun. 2019, pp. 127–136.
[68] C. Rawles, S. Clinckemaillie, Y. Chang, J. Waltz, G. Lau, M. Fair, A. Li, W. Bishop, W. Li, F. Campbell-Ajala, D. Toyama, R. Berry, D. Tyamagundlu, T. Lillicrap, and O. Riva, "AndroidWorld: A dynamic benchmarking environment for autonomous agents," 2024, *arXiv:2405.14573*.
[69] P. Rochat, "Self-perception and action in infancy," *Exp. Brain Res.*, vol. 123, nos. 1–2, pp. 102–109, 1998.
[70] C. Romac, R. Portelas, K. Hofmann, and P. Y. Oudeyer, "TeachMyAgent: A benchmark for automatic curriculum learning in deep RL," in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 9052–9063.
[71] M. Salehie and L. Tahvildari, "Self-adaptive software: Landscape and research challenges," *ACM Trans. Auton. Adapt. Syst.*, vol. 4, no. 2, pp. 1–42, 2009.
[72] Ş. Sarkadi, "An arms race in theory-of-mind: Deception drives the emergence of higher-level theory-of-mind in agent societies," in *Proc. IEEE Int. Conf. Autonomic Comput. Self-Organizing Syst. (ACSOS)*, Toronto, ON, Canada, Sep. 2023, pp. 1–10.
[73] J. Schmidhuber, "Formal theory of creativity, fun, and intrinsic motivation (1990–2010)," *IEEE Trans. Auto. Mental Develop.*, vol. 2, no. 3, pp. 230–247, Sep. 2010.
[74] B. Schölkopf, F. Locatello, S. Bauer, N. R. Ke, N. Kalchbrenner, A. Goyal, and Y. Bengio, "Toward causal representation learning," *Proc. IEEE*, vol. 109, no. 5, pp. 612–634, May 2021.
[75] S. Shaw, E. Wenzel, A. Walker, and G. Sartoretti, "ForMIC: Foraging via multiagent RL with implicit communication," *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 4877–4884, Apr. 2022.
[76] T. Srinivasan, T.-Y. Chang, L. P. Alva, G. Chochlakis, M. Rostami, and J. Thomason, "CliMB: A continual learning benchmark for vision-and-language tasks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 35, 2022, pp. 29440–29453.
[77] A. Stein, S. Tomforde, J. Botev, and P. R. Lewis, "Special issue on lifelike computing systems," *Artif. Life*, vol. 29, no. 4, pp. 390–393, Nov. 2023.
[78] A. Stoytchev, "Self-detection in robots: A method based on detecting temporal contingencies," *Robotica*, vol. 29, no. 1, pp. 1–21, 2011.
[79] B. Subagdja and A. H. Tan, "Beyond autonomy: The self and life of social agents," in *Proc. 18th Int. Conf. Auto. Agents MultiAgent Syst.*, 2019, pp. 1654–1658.
[80] Z. Ullah, F. Al-Turjman, L. Mostarda, and R. Gagliardi, "Applications of artificial intelligence and machine learning in smart cities," *Comput. Commun.*, vol. 154, pp. 313–323, Mar. 2020.
[81] J. Vallverdú, *Generative AI and Causality*. Singapore: Springer, 2024, pp. 55–61.
[82] K. Valmeekam, S. Sreedharan, M. Marquez, A. Olmo, and S. Kambhampati, "On the planning abilities of large language models (A critical investigation with a proposed Benchmark)," 2023, *arXiv:2302.06706*.
[83] D. Vernon, G. Metta, and G. Sandini, "A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents," *IEEE Trans. Evol. Comput.*, vol. 11, no. 2, pp. 151–180, Apr. 2007.
[84] L. Wang, C. Ma, X. Feng, Z. Zhang, H. Yang, J. Zhang, Z. Chen, J. Tang, X. Chen, Y. Lin, W. X. Zhao, Z. Wei, and J. Wen, "A survey on large language model based autonomous agents," *Frontiers Comput. Sci.*, vol. 18, Mar. 2023, Art. no. 186345.
[85] X. Wang, G. Chen, G. Qian, P. Gao, X.-Y. Wei, Y. Wang, Y. Tian, and W. Gao, "Large-scale multi-modal pre-trained models: A comprehensive survey," *Mach. Intell. Res.*, vol. 20, no. 4, pp. 447–482, Jun. 2023.
[86] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen, "Autonomous mental development by robots and animals," *Science*, vol. 291, no. 5504, pp. 599–600, Jan. 2001.
[87] Z. Xi et al., "The rise and potential of large language model based agents: A survey," 2023, *arXiv:2309.07864*.
[88] S. Yang, O. Nachum, Y. Du, J. Wei, P. Abbeel, and D. Schuurmans, "Foundation models for decision making: Problems, methods, and opportunities," 2023, *arXiv:2303.04129*.
[89] Y. Yang, M. Taylor, J. Luo, Y. Wen, O. Slumbers, D. Graves, H. B. Ammar, and J. Wang, "Diverse auto-curriculum is critical for successful real-world multiagent learning systems," in *Proc. 20th Int. Conf. Auto. Agents Multiagent Syst.*, 2021, pp. 1–6.
[90] A. Yapo and J. Weiss, "Ethical implications of bias in machine learning," in *Proc. 51st Hawaii Int. Conf. Syst. Sci.*, 2018, pp. 1–8.
[91] J. Ye, S. Dobson, and F. Zambonelli, "Lifelong learning in sensor-based human activity recognition," *IEEE Pervasive Comput.*, vol. 18, no. 3, pp. 49–58, Jul. 2019.
[92] M. Yim, Y. Zhang, and D. Duff, "Modular robots," *IEEE Spectr.*, vol. 39, no. 2, pp. 30–34, Feb. 2002.
[93] C. Yu, M. Zhang, and F. Ren, "Collective learning for the emergence of social norms in networked multiagent systems," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2342–2355, Dec. 2014.
[94] J. Zhang, X. Yao, J. Zhou, J. Jiang, and X. Chen, "Self-organizing manufacturing: Current status and prospect for industry 4.0," in *Proc. 5th Int. Conf. Enterprise Syst. (ES)*, Sep. 2017, pp. 319–326.
[95] W. Zhang, "Large decision models," in *Proc. 32nd Int. Joint Conf. Artif. Intell.*, 2023, pp. 7062–7067.
[96] W. Xin Zhao et al., "A survey of large language models," 2023, *arXiv:2303.18223*.
[97] Y. Zhao, Y. Chen, K. Tu, and J. Tian, "Learning Bayesian network structures under incremental construction curricula," *Neurocomputing*, vol. 258, pp. 30–40, Oct. 2017.

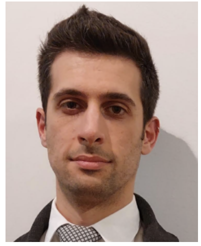
[98] M. Zhuge et al., ''Mindstorms in natural language-based societies of mind,'' 2023, *arXiv:2305.17066*.

MARCO LIPPI received the Ph.D. degree in computer and automation engineering from the University of Florence, in 2010.

He was a Research Assistant with the Universities of Florence, from 2010 to 2011; Siena, from 2011 to 2014; and Bologna, from 2014 to 2016. From March to June 2014, he was a Visiting Scholar with the Laboratoire d'Informatique Paris 6 (LIP6), Université Pierre et Marie Curie, Paris. He has been an Assistant and Associate Professor with the University of Modena and Reggio Emilia, since 2016. Since 2024, he has been an Associate Professor with the University of Florence. His research interests include machine learning and artificial intelligence, natural language processing, explainability, and neuro-symbolic learning, with applications in many fields, and especially in healthcare and the legal domain. In 2012, he received the ''E. Caianiello'' Prize for the Best Ph.D. Thesis in the field of neural networks.

STEFANO MARIANI received the Ph.D. degree in computer science and engineering from the University of Bologna, in 2012.

From April to July 2013, he was a Visiting Scholar with the Technical University of Wien, Austria. Following, he had postdoctoral research positions with the Universities of Bologna and the University of Modena and Reggio Emilia. He is currently a tenure-track Researcher with the University of Modena and Reggio Emilia. He has published one monograph, more than 20 journal articles, and over 60 conference papers, where he delivered over 20 talks. His research interests include distributed computing, pervasive systems, multi-agent systems, socio-technical systems, coordination, self-organization, multi-agent reinforcement learning, and causal discovery, with applications to the Internet of Things, healthcare, and traffic management.



MATTEO MARTINELLI received the Ph.D. degree in industrial innovation engineering from the University of Modena and Reggio Emilia, in 2024.

From October 2022 to April 2023, he was a Visiting Scholar with the Institut für Steuerungstechnik der Werkzeugmaschinen und Fertigungseinrichtungen (ISW), University of Stuttgart, Germany. He is currently holding a postdoctoral research position with the University of Modena and Reggio Emilia, with a research focus on the application of digital twins, distributed, intelligent systems and modern technologies in the industrial domain. In these areas, he has published in several international conferences and journals.

FRANCO ZAMBONELLI (Fellow, IEEE) received the Laurea degree in electronic engineering and the Ph.D. degree in computer science and engineering from the University of Bologna, in 1992 and 1997, respectively.

He is currently a Full Professor with the University of Modena and Reggio Emilia. He has published over 130 articles in peer-reviewed journals and has been invited speaker at many conferences and workshops. He has been the Scientific Manager of the EU FP6 Project CASCADAS and a coordinator of the EU FP7 Project SAPERE and the PRIN 2017 Project Fluidware. His current research interests include pervasive computing, multi-agent systems, self-adaptive, and self-organizing systems, with applications to healthcare and smart cities.

Prof. Zambonelli is an ACM Distinguished Scientist and a member of the Academia Europaea. He received the 2018 Influential Paper Award of the International Foundation for Autonomous Agents and Multiagent Systems. He is on the Editorial Board of the *ACM Transactions on Autonomous and Adaptive Systems*, *SN Computer Science*, *IEEE Technology and Society Magazine*, *Computer Journal* (BCS), and *International Journal of Pervasive Computing and Communications*.

Open Access funding provided by 'Università degli Studi di Modena e Reggio Emilia' within the CRUI CARE Agreement