

Individual and Collective Self-Development: Concepts and Challenges

Marco Lippi, Stefano Mariani, Matteo Martinelli, Franco Zambonelli
 Department of Sciences and Methods for Engineering
 University of Modena and Reggio Emilia, Italy
 Email: name.surname@unimore.it

Abstract—The increasing complexity and unpredictability of many ICT scenarios will represent a major challenge for future intelligent systems. The capability to dynamically and autonomously adapt to evolving and novel situations, with a partial or limited knowledge of the domain, both at the level of individual components and at the collective level, will become a crucial need for smart devices acting in many application domains. In this paper, we envision future systems able to self-develop mental models of themselves and of the environment they act in. Key properties will include: 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 our work, we will introduce the vision of self-development in ICT systems, by framing its key concepts and by illustrating suitable application domains. Then, we overview the many research areas that are contributing or can potentially contribute to the realisation of the vision, and identify some key research challenges.

Index Terms—Self-development, sense of agency, learning, self-adaptation, self-organization.

I. INTRODUCTION

HUMAN infants, since their early months, start experiencing with their own body, moving hands, touching objects, and interacting with people around. Such activities are part of an overall process of self-development (a.k.a. autonomous mental development), which lets them gradually develop cognitive and behavioural capabilities [1]. These skills include the capability to recognize situations around, the sense of self, the sense of agency (i.e., understanding the effect of own actions in an environment), the capability to act purposefully towards a goal, and some primitive social capabilities (i.e., knowing how to act in the presence of others).

The possibility of building ICT systems capable – as humans – of self-developing their own mental and social models and to act purposefully in an environment, is increasingly recognized as a key challenge in many areas of artificial intelligence (AI), such as robotics [2], intelligent IoT and smart environments [3], [4], autonomous vehicles management [5], [6].

Indeed, for small-scale and static scenarios, and for simple goal-oriented tasks, it is possible “hardwire” a model of the environment within a system, alongside some pre-designed plans of action. However, for larger and 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); and (iii) learn to achieve goals on this basis and depending on the current situation; (iv) learn how to organize and coordinate actions among multiple distributed components whenever necessary.

The main contribution of this paper is to frame the key concepts of self-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 self-development, both at the individual and at the collective level, and sketches key application scenarios; Section III 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 self-development, each *per se* challenging; Section IV identifies additional horizontal challenges to be attacked, emphasizing the key role that the self-adaptive and self-organizing research community could play. Finally, Section V concludes by sketching our current and future research work in the area.

II. FRAMEWORK AND APPLICATIONS

The term “self-development” is used to indicate the process carried out by infants during the early stages of their life [1] but, more generally, it can be also associated to the developmental nature of agents that live and interact with a novel environment. The idea of our framework is depicted in Figure 1. At the *individual* level, the first contacts an agent has with a new environment are through *embodiment and perception*: it typically tries to move and interact, in order to test the effect of its own action, so as to acquire a *sense of agency*. Only after these skills have been sufficiently developed, the agent can start behaving in a *goal-oriented* way, by choosing the sequence of actions that can bring to the fulfilment of a goal.

Clearly, the individual level quickly turns into a *collective* one, where the agent has to face other agents, which are not under its control: thus, the agent learns to recognize *self and non-self*, as well as to develop *strategic thinking*, by choosing its own actions by taking into account the behaviour of the other agents. As the complexity grows, the agent will need to

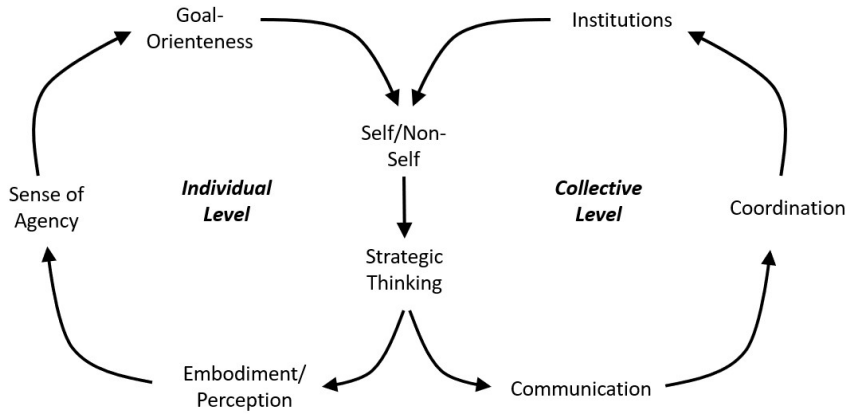


Fig. 1: The conceptual framework of self-development.

understand whether it can *communicate* with others, and with which protocols, as well as *coordinate* in order to jointly act towards a common goal, possibly through the creation of an *institution*.

The whole development, at both the individual and collective level, can be seen as a never-ending, cyclic process, where agents have to continuously adapt to new situations and environments.

A. The Individual Level

At the individual level, an agent \mathcal{X} immersed in an environment can observe (or sense) a set of variables $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$. Internal variables that describe the status of the agent are included in this set as well. The agent can interact with the environment through a set of “actions” $\mathcal{A} = \{a_0, \dots, a_{n-1}, null\}$, including the *null* action.

Embodiment and Perception. Initially, the agent needs to autonomously identify and recognize the components of sets \mathcal{A} and \mathcal{V} : this means that it should get acknowledged of its actuating and sensing skills. Even without resorting to complex AI techniques, methods from the reflective and self-adaptive programming systems can effectively apply in this phase [7] to let the agent dynamically self-inspect its capabilities and start analyzing the observed variables. Still in this phase, the agent can also start acquiring some understanding of the relations between the observed variables over time, as well as some simple prediction capabilities.

Sense of Agency. After the first phases that are mostly dealing with perception, the agent needs to understand what are the effects of each $a_j \in \mathcal{A}$ on \mathcal{V} . This can happen even by chance, with random actions, thus trying to apply actions, without any goal in mind, just to see their effects [2]. Throughout this process, the agent will eventually recognize that, given a current state v^t and the application of an action a_i , the environment will reach (with some probability) a different state v^{t+1} . This mechanism enables the construction of the basic sense of agency [1], and of the sense of causality.

Goal-orientedness. As the agent acquires more sophisticated skills, it can start applying \mathcal{A} with a specific goal in mind. Given the current state v^t and a desired future state v^g (the desired “state of the affairs”), the agent applies the acquired sense of agency by applying the actions that can possibly lead to v^g . This also involves achieving the capability of planning the required sequence of actions to achieve a specific goal.

Self and Non-Self. After an individual agent starts interacting with the environment and testing the effects of its own actions \mathcal{A} , it recognizes that such actions have effect on the environment. As an immediate consequence, it also understands that there are effects that are not under its own control. That is, there are “non-self” entities acting in the environment, too. By learning how to apply \mathcal{A} , the agent also learns the limits of such actions because of non-self entities affecting v^t .

Strategic Thinking. Once the agent has built a world model (how \mathcal{A} affects v^t) and has included the mental models of others (non-self) [8], it can start designing strategies. That is, it can recognise that there are goals that it can possibly (or hopefully) attain only by accounting for the actions of others.

Once again, we remark that self-development is not to be conceived as a “once-and-for-all” process. Rather, it is a life-long process: environments can be dynamic, new variables may become available and thus enable more detailed observations. Also, new actions may become feasible or, the other way round, be no longer be available. This requires the agents to re-tune their learnt sense of agency, and re-think how to achieve goals in isolation and in the presence of non-self entities.

B. The Collective Level

As multiple agents enter the arena, each of them quickly recognizes that there are goals that cannot be achieved in isolation or by simply applying strategic thinking, but they rather need a deep interaction among all the actors. Therefore, as part of their individual self-development, also need to develop some forms of “autonomous social engagement”.

Communication. A first, necessary, step is to identify the way in which agents can communicate and exchange messages. Agents should thus be provided with a specific set of “communication actions”, which could take the form of explicit communication acts (messages) or implicit actions that aim at influencing the others, i.e., by leaving signs in the environment (stigmergy) or by acting in a way that is easily noticeable by others (behavioural implicit communication)[9]. In some cases, the agent has to learn how to receive and send such messages, as a social form of action and perception.

Coordination. When evaluating the possible communication actions, each agent understands the way in which such acts can be exploited to control some environmental variables, and even those that are not (fully) controllable by itself alone. Therefore, such explorative behavior enables the learning of basic forms of coordination, which can be thought of as a social form of learning the sense of agency.

Institution. After exploring coordination protocols, the agents can eventually “institutionalize” their way of interacting. That is, they will learn those acceptable social patterns of coordination, and the set of social norms and social incentives, that enables them to systematically achieve goals together [10].

As in the single-agent setting, a dynamic environment or an evolving agent population may require the above collective process to assume a continuous cyclic nature. We hereby remark that the communication, coordination, and institution stages are not necessary to promote complex goal-oriented collective actions [11]. Yet, whenever communication protocols exist, the self-development process will naturally and gradually learn how to exploit them.

C. Application Scenarios

There are diverse application scenarios that can potentially take advantage of systems capable of self-development.

Robotics is the area which first identified the profitability of building robots capable of self-development [12]. In particular, it is necessary when the robot gets damaged while in operation, and has to develop a novel understand of what it can do according to its residual operational capabilities. At the collective level, 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 [13].

Smart factories, as collective robotic systems, can be seen as an aggregated group of components that act together in order to achieve a production goal. Beside their basic scheme of functioning, defined at design time, if one component of the manufacturing system breaks or has some unexpected behaviour, the manufacturing system should ideally adapt to the new situation, and self-develop capabilities of acting so as to overcome the problem without undermining production [14]. The need for adaptability and flexibility is indeed explicitly recognized as a key challenge in Industry 4.0 initiatives [15].

Smart homes can facilitate our interactions with the environment and increase our safety and comfort. We envision

that once a new home is built, its smart devices could start exploring their own individual and collective capabilities, so as to eventually learn how they can affect the home environment, and apply such capabilities once users will start populating it. This will also require to continuously adapt to habits and preferences of users, accommodate new devices and services, tolerate partial failures. Our preliminary experience suggests the feasibility of the vision [3].

Smart cities as well can potentially take advantage of self-development approaches [16]. However, unlike in a smart home, a smart city is not a system free to explore the effect of its actions and interactions, and eventually become capable to act in a goal-oriented way. Thus, for this scenario (but most likely also for smart factories), simulation-based approaches should probably be exploited: system components will be made self-developing in a simulated environment, before being eventually deployed in the real world [6].

III. RESEARCH APPROACHES

The idea of self-development, at both the individual and collective level, has been widely investigated in areas such as cognitive psychology, neuroscience, philosophy, and ethics [2]. We hereby focus on the computational perspective, and in particular on the most recent approaches that can contribute to realise the self-development vision (Figure 2). Although most of these approaches can play a fundamental role and are already providing precious insights on the problems, they still have to attack several challenges to become practical tools for future self-developing systems.

We do not focus here on the basic levels of individual self-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 [17]) and the capability of controlling their own actuators purposefully.

A. Goal-oriented Learning

The broad area of reinforcement learning shares with our vision the objective of training machines to act in a goal-oriented way in a specific context. However, despite the amazing recent results in the area, in particular with deep Q-learning [18], most current approaches do not aim at building systems with a sense of agency and capable of developing an interpretable world model, but rather at achieving goals based on explicit, domain-based rewards, that are named *extrinsic*. This makes most approaches highly ineffective in scaling up to learning tasks in complex contexts, or across domains, or despite the ever-changing dynamics of the environment.

Curriculum-based approaches to machine learning go somewhat in the direction of gradually developing the capability to act in complex scenarios [19]. The agent is first trained on simple tasks, and the gained knowledge is accumulated and exploited in increasingly complex scenarios, where further skills can thus be effectively learnt. Yet again, most of these approaches do not focus on the development of a world model and of an explicit sense of agency.

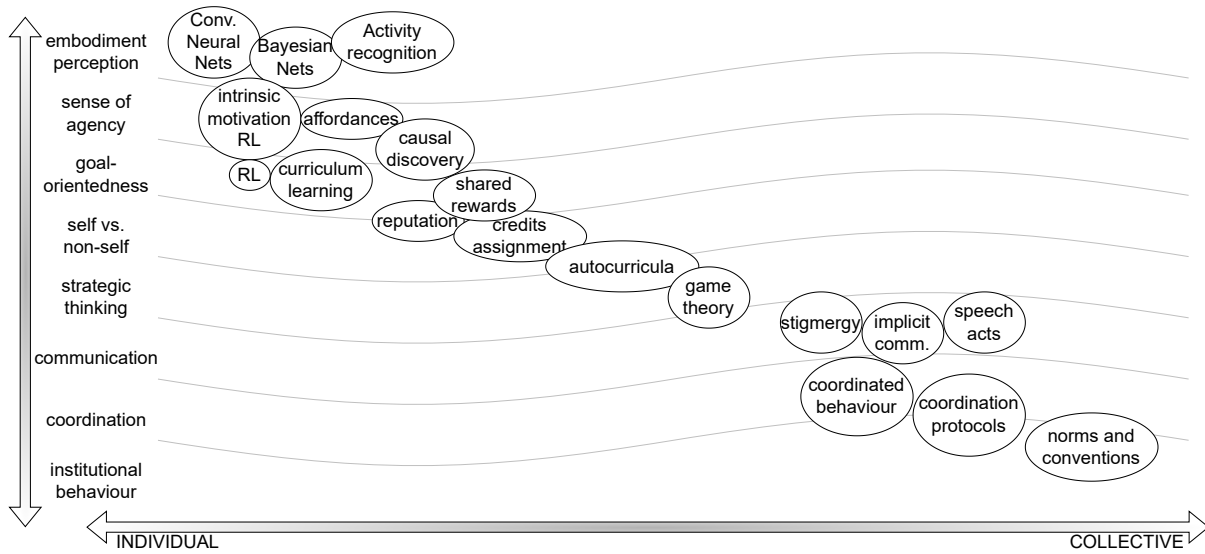


Fig. 2: Key techniques for self-development.

Reinforcement learning approaches based on *intrinsic* rewards [20], instead, more closely exploit the idea of exploring the world to develop a sense of agency. In fact, while extrinsic rewards are typically designed by a “teacher” (e.g., the score in a videogame) intrinsic rewards are developed by the agent itself to satisfy its curiosity (i.e., when it discovers how to achieve specific tasks). For example, in [21] intrinsic rewards are computed as the error in forecasting the consequence of the action performed by the agent given its current state.

Recent approaches based on the theory of affordances [22] 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 sense of agency, i.e., a model of how their actions affect the environment.

In any case, all these approaches face the key challenge of building general conceptual and practical tools to: (i) learn to effectively act in an environment by exploiting the power of model-free sub-symbolic (deep learning) approaches; and, at the same time, (ii) learn incremental and reusable causal models of the world. The latter being increasingly recognized as a key ingredient for intelligence and self-development.

B. Learning causality

Understanding and leveraging *causality* is recognized as a key general challenge for AI in the coming years [23]. Judea Pearl [24] 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. The first level consists in simply detecting causal relations as associations, whereas the second one assumes the possibility to intervene in the environment and observe the effects of the taken actions. Finally, the third level enables reasoning and planning on the basis of counterfactual analysis. Such layers correspond to some of the phases of the self-development

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

Bayesian and causal networks are among the models that are most widely exploited in order to build interpretable causal models of the world [24]. A recent contribution that is in line with the ideas we envision for self-development is the application of curriculum learning to the problem of learning the structure of Bayesian networks [25]. On a pure sub-symbolic level, on the other hand, another recent work proposes to learn causal models in an online setting [26], with the aim to find (and strengthen) causal links between input and output variables.

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

C. Autocurricula

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 have deeply investigated this problem and the decision-making processes behind [27]. In this context, it has also been shown that agents can effectively learn in autonomy to improve their performance in dealing with others [28].

However, when moving from theoretical settings (e.g., the prisoner’s dilemma) to complex and realistic scenarios where agents have complex goals (e.g., hide-and-peek in a building), peculiar phenomena arise. The more one agent learns, the more it challenges others, triggering a continuous increase in complexity of behaviour, ultimately enabling to incrementally learn

more sophisticated means to act. This somewhat resembles the increase of complexity that agents face in curricula approaches to reinforcement learning. The key difference being 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* [11].

Recently, autocurricula-based approaches have produced stunning results in multiagent environments, both cooperative and competitive (e.g., in the hide and seek scenario [29]). And we consider such approaches fundamental towards the self-development of complex agent societies. However, a deep understanding of the process that drives evolution of individual and collective behaviours is still missing, and is a key challenge for the next few years. To this end, providing agents an explicit modelling (possibly in causal terms) of the others' behaviour and of the overall societal behaviour, may be necessary [8]. Also, autocurricula approaches do not currently account for the possibility of interact with other agents, which may indeed be fundamental to improve collective learning.

D. Learning to communicate and coordinate

As already mentioned, agents may communicate and coordinate: by explicit messages, by leaving traces in the environment, or implicitly [9].

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 [30]) and coordinate actions. However, these communication approaches are usually assumed as an *innate* capability 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 self-development vision it should be developed by learning.

For example, with reference to explicit communication acts, [31] proposes a voting game to let agents learn to share a communication language and to develop a strategy to communicate. In [32], it is shown that reinforcement learning can be effectively applied to let agents learn how to communicate in order to achieve a specific effect. In the case of *implicit* communication, instead, forms of implicit behavioural communications have been shown to emerge in simple system components that purposefully move in an environment [33], 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.

In any case, the development of general approaches to let agents develop fully-fledged forms of communication and coordination is still an open challenge, which may call for agents to develop not only a model of the world, but an overall model of the society (i.e., a social sense of agency). as a sort of social sense of agency.

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 [34]. The mechanisms leading to the spontaneous emergence of institutions in human society, there included the mechanisms to promote and sustain altruistic and cooperative behaviour (e.g., reputation and shared rewards), have been widely investigated [35]. However, most approaches to building multiagent systems assume such mechanisms as *explicitly designed* [34].

Yet, some promising studies related to the emergence of institutionalised behaviour in multiagent systems have been undertaken (see [10] for a recent survey). For instance, [36] proposes a collective learning framework where agents learn to adopt norms in repeated coordination, i.e., agents eventually *learn* that a social norm has emerged, and “institutionalize” their behaviour in their (social) decision making processes by complying to the norm. Another interesting work [37] integrates rational thought, reinforcement learning, and social interactions to model norms emergence in a society: agents incrementally develop a social behaviour (a social norm) while *internalising* it within their cognitive model.

However, the development of general models and tools to support the proper learning and evolution of institutionalised mechanisms of coordination in ICT and multiagent systems 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 norms; or how to avoid inconsistencies and misunderstandings in their interpretation.

IV. 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 addition, it is possible to identify several additional “horizontal” challenges, i.e., of a general nature independently of the specific approach.

The specific nature of such challenges, in our opinion, makes them specifically suited for being pursued by the self-adaptive and self-organising research community, i.e., the ACSOS community at large.

Engineering. Many of the presented approaches are grounded in machine learning, a discipline with plenty of years of research behind, but in which good engineering practice is often neglected, 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, dependability, thus missing the flexibility to adopt them across different domains, tasks, datasets [38]. In addition, given that the diverse approaches presented can each contribute important pieces to the overall vision of self-development, sound engineering approaches are needed to try to integrate such a heterogeneous plethora into a coherent whole. These represent multi-faceted and horizontal research challenges that, in our opinion, could and should be profitably attacked by the self-adaptive and self-organising

research community, due to its inherent software engineering endeavour.

Controlling evolution. Self-development raises the issue of somewhat controlling how behaviours evolve, as individual learns new skills and tasks, and as the collective learns new way of coordinating and acting together. 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 behaviours 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 self-development path? Experience in self-adaptive components based on feedback, as well as in the study of emergent behaviours in self-organising systems and definitely help in finding proper technical answers, and – why not – *ethical* ones [39].

Humans in the Loop. The more self-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?) 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 [40] or in AI-based hiring technology. Technical challenges will be meat for the HCI and distributed systems communities (there included the self-organising systems one). 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 into the 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.

Sustainability. Algorithms for self-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 pre-emptible CPU cores and 256 GPUs on GCP [29]: the default model optimised over 1.6 million parameters taking 34 hours to reach the fourth stage over six of agents skills progression. This example is a sort of best-in-class projects; 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 self-development systems sustainable, and affordable by others other than the big technology players.

Explainability. Being able to inspect and explain the decision making process of AI systems is already a hot topic, so much that an entire research field (XAI, from eXplainable AI) has born. We already commented several times how such problems should be compulsory accounted for also for self-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 behaviours, patterns, and configurations emerging from local interactions is mostly still considered the “holy grail”.

V. CONCLUSIONS, CURRENT AND FUTURE WORK

In this paper, we have elaborated upon the vision of self-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, multiagent systems, are already showing its potential feasibility.

From our side, 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 [3]. 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 can 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 to 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. Rather, they cooperate to refine their own local causal models whenever they recognize partial, missing, or wrong information, by organising 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. 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.

Acknowledgements: Work supported by the Italian MUR, PRIN 2017 Project “Fluidware”.

REFERENCES

- [1] P. Rochat, “Self-perception and action in infancy,” *Experimental brain research*, vol. 123, no. 1-2, pp. 102–109, 1998.
- [2] 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, 2001.

- [3] M. Lippi, S. Mariani, and F. Zambonelli, "Developing a sense of agency in IoT systems: Preliminary experiments in a smart home scenario," in *17th CoMoRea workshop at PerCom*. IEEE, 2021.
- [4] S. Jha, M. Schiemer, F. Zambonelli, and J. Ye, "Continual learning in sensor-based human activity recognition: An empirical benchmark analysis," *Inf. Sci.*, vol. 575, pp. 1–21, 2021.
- [5] S. Mariani, G. Cabri, and F. Zambonelli, "Coordination of autonomous vehicles: Taxonomy and survey," *ACM Comput. Surv.*, vol. 54, no. 1, pp. 19:1–19:33, 2021.
- [6] Y. Yang, M. Taylor, J. Luo, Y. Wen, O. Slumbers, D. Graves, H. Bou Ammar, and J. Wang, "Diverse auto-curriculum is critical for successful real-world multiagent learning systems," in *20th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, 2021.
- [7] M. Salehie and L. Tahvildari, "Self-adaptive software: Landscape and research challenges," *ACM transactions on autonomous and adaptive systems*, vol. 4, no. 2, pp. 1–42, 2009.
- [8] B. Subagdja and A.-H. Tan, "Beyond autonomy: The self and life of social agents," in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, 2019, pp. 1654–1658.
- [9] S. Mariani and A. Omicini, "Anticipatory coordination in socio-technical knowledge-intensive environments: Behavioural implicit communication in MoK," in *AI*IA 2015*. Springer, 2015, pp. 102–115.
- [10] A. Morris-Martin, M. De Vos, and J. Padget, "Norm emergence in multiagent systems: a viewpoint paper," *Autonomous Agents and Multi-Agent Systems*, vol. 33, no. 6, pp. 706–749, 2019.
- [11] 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," *arXiv preprint arXiv:1903.00742*, 2019.
- [12] J. Bongard, V. Zykov, and H. Lipson, "Resilient machines through continuous self-modeling," *Science*, vol. 314, no. 5802, pp. 1118–1121, 2006.
- [13] N. Cambier, R. Miletitch, V. Frémont, M. Dorigo, E. Ferrante, and V. Trianni, "Language evolution in swarm robotics: A perspective," *Frontiers in Robotics and AI*, vol. 7, p. 12, 2020.
- [14] M. Martinelli, S. Mariani, M. Lippi, and F. Zambonelli, "Self-development and causality in intelligent environments," in *Workshops at 18th International Conference on Intelligent Environments (IE2022), Biarritz, France, 20-23 June 2022*, ser. Ambient Intelligence and Smart Environments, vol. 31. IOS Press, 2022, pp. 248–257.
- [15] J. Zhang, X. Yao, J. Zhou, J. Jiang, and X. Chen, "Self-organizing manufacturing: Current status and prospect for industry 4.0," in *5th International Conference on Enterprise Systems*, 2017, pp. 319–326.
- [16] M. Saelens, Y. Kinoo, and D. Weyns, "Heyciti: Healthy cycling in a city using self-adaptive internet-of-things," in *IEEE International Conference on Autonomic Computing and Self-Organizing Systems Companion*, 2020, pp. 226–227.
- [17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [18] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [19] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," in *Proceedings of the 26th International Conference on Machine Learning*, 2009, pp. 41–48.
- [20] J. Schmidhuber, "Formal theory of creativity, fun, and intrinsic motivation (1990–2010)," *IEEE Transactions on Autonomous Mental Development*, vol. 2, no. 3, pp. 230–247, 2010.
- [21] Y. Burda, H. Edwards, D. Pathak, A. Storkey, T. Darrell, and A. A. Efros, "Large-scale study of curiosity-driven learning," in *International Conference on Learning Representations*, 2018.
- [22] K. Khetarpal, Z. Ahmed, G. Comanici, D. Abel, and D. Precup, "What can i do here? a theory of affordances in reinforcement learning," in *International Conference on Machine Learning*, 2020, pp. 5243–5253.
- [23] B. Schölkopf, F. Locatello, S. Bauer, N. R. Ke, N. Kalchbrenner, A. Goyal, and Y. Bengio, "Toward causal representation learning," *Proceedings of the IEEE*, 2021.
- [24] J. Pearl and D. Mackenzie, *The book of why: the new science of cause and effect*. Basic Books, 2018.
- [25] Y. Zhao, Y. Chen, K. Tu, and J. Tian, "Learning bayesian network structures under incremental construction curricula," *Neurocomputing*, vol. 258, pp. 30–40, 2017.
- [26] K. Javed, M. White, and Y. Bengio, "Learning causal models online," *arXiv preprint arXiv:2006.07461*, 2020.
- [27] R. B. Myerson, *Game theory*. Harvard university press, 2013.
- [28] A. Nowé, P. Vrancx, and Y.-M. De Hauwere, "Game theory and multi-agent reinforcement learning," in *Reinforcement Learning*. Springer, 2012, pp. 441–470.
- [29] B. Baker, I. Kanitscheider, T. Markov, Y. Wu, G. Powell, B. McGrew, and I. Mordatch, "Emergent tool use from multi-agent autocurricula," *arXiv preprint arXiv:1909.07528*, 2020.
- [30] S. Meganck, S. Maes, B. Manderick, and P. Leray, "Distributed learning of multi-agent causal models," in *IEEE/WIC/ACM International Conference on Intelligent Agent Technology*. IEEE, 2005, pp. 285–288.
- [31] S. Gupta and A. Dukkhipati, "Winning an election: On emergent strategic communication in multi-agent networks," in *International Conference on Autonomous Agents and Multiagent Systems*, 2020, pp. 1861–1863.
- [32] J. N. Foerster, Y. M. Assael, N. De Freitas, and S. Whiteson, "Learning to communicate with deep multi-agent reinforcement learning," *arXiv preprint arXiv:1605.06676*, 2016.
- [33] N. A. Grupen, D. D. Lee, and B. Selman, "Low-bandwidth communication emerges naturally in multi-agent learning systems," *arXiv preprint arXiv:2011.14890*, 2020.
- [34] 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*. Springer, 2001, pp. 126–147.
- [35] M. A. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, 2006.
- [36] 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, 2014.
- [37] R. Beheshti, "Normative agents for real-world scenarios," in *International Conference on Autonomous Agents and Multi-Agent Systems*, 2014, pp. 1749–1750.
- [38] B. Porter and R. Rodrigues Filho, "Distributed emergent software: Assembling, perceiving and learning systems at scale," in *IEEE International Conference on Self-Adaptive and Self-Organizing Systems*, 2019, pp. 127–136.
- [39] A. Yapo and J. Weiss, "Ethical implications of bias in machine learning," in *51st Hawaii International Conference on System Sciences*, 2018.
- [40] 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, 2018.