
A BRIEF SURVEY OF ASSOCIATIONS BETWEEN META-LEARNING AND GENERAL AI

A PREPRINT

 **Huimin Peng***

peng.huimin.pennie@gmail.com

ABSTRACT

This paper briefly reviews the history of meta-learning and describes its contribution to general AI. Meta-learning improves model generalization capacity and devises general algorithms applicable to both in-distribution and out-of-distribution tasks potentially. General AI replaces task-specific models with general algorithmic systems introducing higher level of automation in solving diverse tasks using AI. We summarize main contributions of meta-learning to the developments in general AI, including memory module, meta-learner, coevolution, curiosity, forgetting and AI-generating algorithm. We present connections between meta-learning and general AI and discuss how meta-learning can be used to formulate general AI algorithms.

Keywords Meta-Learning · General AI · Coevolution · Curiosity · Forgetting

1 Introduction

1.1 General AI

Current AI research primarily concentrates attention upon computer vision, natural language processing and automated AI systems, seeking to narrow the performance gap between machine and human beings [1]. General AI (AGI, Artificial General Intelligence) seeks algorithms that are as 'smart' as human beings. For example, computer-based systems such as AlphaGo are created and they can defeat even the best human players. General AI aims to design algorithmic models that are applicable to a wide range of tasks. Usually in meta-learning, deep models can only perform self-improvement to solve similar tasks but not vastly different tasks. In AutoML, spaces of deep models are searched through to identify the optimal network for any task in a highly automated fashion. AutoML is a general AI algorithm that can be applied to a wide range of tasks potentially including vastly different ones. Meta-learning pursues speed and accuracy in the self-improvement of deep models to solve unseen tasks. AutoML is slower and needs to accelerate model search through sparse search spaces using meta-learning tools. In the future, meta-learning may develop more general methods that are capable of adapting deep models to solve vastly different tasks more efficiently.

Strong AI pursues the development of algorithms comparable or superior to human intelligence and suffers criticism of ethics. General AI belongs to weak AI realm and invents algorithms to handle varied tasks that would require human intelligence to settle [2]. There have been ongoing discussions as to whether computer-based algorithms will bring about general AI. As in [1], considering the fast development in AI research and application of AI to almost all areas of human life while exploiting the benefits of big data, general AI will be accomplished in this century. In contrast, [2] believes that computer-based algorithmic models will not realize general AI for the following reasons. First, general AI requires that model account for causality and perform human-like logic reasoning. Second, there is no explicit representation for tacit knowledge from human comprehension, which is the critical piece of information applied in problem solving. Sometimes human beings are not even aware of the tacit knowledge they manipulate when they complete a mission. To the best of our knowledge, tacit knowledge cannot be represented explicitly in algorithmic general AI models. Third, computer-based learners are not positioned in any real social network. Algorithms do not acquire crucial information

*Thank you for all helpful comments! Feel free to leave a message about comments on this manuscript. In case I did not receive email, my personal email is 974630998@qq.com. Thanks to github.com/kourgeorge/arxiv-style for this pdf latex template.

externally through versatile interactions within social networks, where subjective and objective information can be balanced to make sensible optimal decisions.

Current best general AI systems include IBM's **Watson** and Google's **AlphaGo**. Watson excels at natural language processing and can be used as AI doctors, which can communicate with patients through dialogues and write prescriptions. Watson has access to history of medical records, related academic literature, previously prescribed medication, etc. These vital records are thoroughly weighed and processed in Watson to produce recommendations and to assist doctors with routine tasks. Furthermore, AlphaGo outperforms human Go players by training supervised deep reinforcement learning models with real games data. Despite the pronounced success of Watson and AlphaGo, even the best general AI systems are susceptible to failure cases. [2] points out that Watson may not handle vastly different tasks such as unseen out-of-distribution tasks. Furthermore, AlphaGo trained on Go may not be well generalized to other games with different explicit rules. Although deep learning models succeed in solving trained tasks with high accuracy, in reality, we are faced with more challenging situations where we should generalize to dynamic unseen tasks. Moreover, both causality and tacit knowledge should be considered in algorithmic models to realize general AI [2]. On one hand, tacit knowledge may be implicitly represented by deep neural network models where hidden layers process a large amount of high-dimensional information. On the other hand, although algorithmic models can neither be placed within real social networks nor learn from subjective emotions, these elements may be embedded in algorithms to compensate for over-simplified model specifications.

ALE (Arcade Learning Environment) [3] is a software platform devised to evaluate the performance of general AI algorithms. In ALE, algorithms are scored using their performances on Atari 2600 games. A game is seen as a task. Algorithms are trained on several games and tested on others to judge generalization capacity. Game types in the game pool are taken to be as diverse as possible to best measure the generalization capacity of general AI algorithms. The best general AI model shows the best performance on the greatest number of games in ALE. General AI is supposed to score high under all game modes in all games. The proportion of all tasks that can be scored well with the general AI system is its primary performance measure. Meta-learning ingredients are integrated into general AI to achieve better model generalization capacity and to accelerate optimal model search in automated machine learning.

General AI can be realized by learning to reason automatically from natural language using **autoformalization** methods [4]. Autoformalization performs causal reasoning in an automated way through deep learning. It provides an attempted solution to the first hurdle of general AI stressed in [2]. Autoformalization system transforms natural language into uniform representations of mathematical reasoning, from which deductions and inductions can be made using rules of mathematical reasoning. Autoformalization allows general AI algorithms to standardize diverse natural language and to reason mathematically in an automated way. It is assumed that mathematical reasoning is fundamental and that autoformalization technique is applicable to all reasoning tasks [4]. Autoformalization is comprised of the following three components: (1) A dimension-reduction model that transforms standard mathematical reasoning statements into formal feature embedding, (2) A translation model that translates diverse natural language into standard mathematical statements and extracts the corresponding feature embedding, (3) An exploration model that searches for the premise model and the reasoning conversion between premise and current embedding. Automated autoformalization system can be self-improved to generalize to different reasoning tasks. However, searching for the optimal reasoning conversion using open-ended exploration is still challenging in the general AI system. Meta-learner can be used to accumulate reasoning experiences and to guide the search for reasoning conversion in the exploration model.

1.2 Meta-Learning

A comprehensive survey of recent advances in few-shot meta-learning is provided in [5]. Updating deep learning models for adaptation to few-shot unseen tasks has attracted significant attention. Meta-learning improves model generalization capacity of deep neural networks to achieve the objective of 'learning-to-learn'. For in-distribution tasks (similar to trained tasks), meta-learner identifies model training results of the most similar tasks, and directly applies these results to solve in-distribution tasks. For out-of-distribution tasks (different from trained tasks), meta-learner performs model adaptation using few-shot task data through model generalization modules. Meta-learner aggregates prior model training experiences to confirm the most promising direction for current model adaptation. But meta-learning is not constrained in few-shot deep learning and has also a significant influence upon automated general AI. This part was not covered in [5]. To complement for this part, this manuscript is mainly about the contribution of meta-learning to general AI.

AI-GA (**AI-Generating Algorithm**) [6] applies meta-learning architectures to construct general AI algorithms. AI-GA is comprised of three components: (1) A meta-learning architecture that provides the framework of general AI algorithm, (2) A meta-learning algorithm that implements self-improvement for specified machine learning models, (3) An effective generative mechanism that creates increasingly complicated tasks. For reinforcement learning tasks, a generative scheme offers increasingly complex environments which are used as new tasks to train the solver. A general AI algorithm should

be applicable to diverse task settings and should be able to automate any machine learning model. AI-GA framework combines generative algorithms and meta-learning architectures to create a general AI system that can settle a wide variety of tasks. In generative scheme, AI-GA considers task-solver pairs, in which task evolves to be more complex and solver performs self-improvement to solve generated tasks. Meta-learner accumulates model training experiences and guides the self-improvement of learner to solve generated tasks efficiently. Meta-learner is the 'brain' that looks back and makes analysis so that future exploration is more sensible. Meta-learner can automate the self-improvement process of any machine learning algorithm to solve unseen tasks.

AI-GA guides the self-improvement of solver through meta-learner and task generative mechanism [6]. AI-GA trains a meta-learning-based general AI system using generated tasks under the guidance of 'upper-level' meta-learner. In the real 'world', task generative mechanism is often governed by hyperparameters which vary cyclically so that generated tasks are not far off the 'mainstream'. Since open-ended unbounded task generation may not provide useful tasks which best help learner self-update to overcome the most urgent disadvantage, task generation should be controlled and guided by a meta-learner. In addition, careful experimental design may be integrated in the task generative mechanism so that it explores the unknown 'world' more efficiently. AI-GA uses meta-learning architectures to construct general AI algorithms. Any machine learning model can be regarded as the base learner / solver. Meta-learner provides automated adaptation of the machine learning model to solve any task. In the generative mechanism, tasks evolve to be increasingly complex and correspondingly the solver is self-updated to solve generated tasks with high performance.

Meta-learning methods are integrated into general AI algorithms to achieve high automation and extensive applicability. **AutoML** (Automated Machine Learning) searches for the global optimal neural network for any task in an automated way. In this sense, AutoML is also a general AI algorithm since it automates the process of using neural network model to solve any task and to achieve highest prediction accuracy. Meta-learning models are integrated within AutoML where meta-learner aggregates training experiences, predicts performance of each hyperparameter combination and points out the most promising hyperparameter combination to explore in the future. Memory module stores long-term model training experiences so that long-range dependence can be considered in the model. Finding a new network that outperforms the current best neural network is regarded as the current task to fulfill. Network model is self-updated to solve the current task and a better neural network model is found. For any task, meta-learner utilizes previous model training experiences to offer a promising initial model, which is not far away from the global best model for this task. The aforementioned **autoformalization** [4] is a general AI algorithm that conducts automated mathematical reasoning from diverse natural language. Meta-learner can be used to accelerate the premise and conversion search in the exploration model of the automated autoformalization reasoning system.

This paper provides a brief review of meta-learning methods that contribute to the realization of general AI and is organized as follows. Section 2 lists out versatile neural network designs in meta-learning that may be applied in general AI algorithms. Section 3 summarizes coevolution frameworks in meta-learning that can be used to construct general AI methods. Section 4 describes artificial curiosity which is a meta-learning concept that can be applied to avoid local optima in general AI optimization. Section 5 surveys forgetting mechanisms in machine learning models. Section 6 summarizes the contribution of these meta-learning frameworks to the developments in general AI.

2 Neural Network Design

[7] reviews the significant historical breakthroughs in deep learning covering a complete archive of neural network model specifications. Meta-learning does not constitute a large proportion in this survey paper, except that RNN and LSTM are both meta-learning systems to adapt deep learners efficiently to unseen tasks. RNN and LSTM can continually perform self-improvement through recursive cells. RNN and LSTM can be very deep including more than 1,000 layers, and gradients in backpropagation do not explode or vanish. RNN and LSTM can process long time series data and can be used to model long-range dependence. Meta-learning models pursue speed and accuracy and can adapt deep learning models to solve unseen few-shot tasks efficiently. Meta-learning may be exploited as a complement module of deep learning to increase the generalization capacity of deep neural networks and to infuse higher level of automation into deep models.

LSTM can be viewed as a meta-learning system which devises internal memory modules to store model training experiences, recursive cells to perform self-improvement, multiplicative gates to control input and output, and forget gates to discard redundant past information [8]. LSTM is a complete system for meta-learning and can be used as a base learner or meta-learner in which stochastic gradient descent self-update is performed. RNNAI is a meta-learning framework, where both base learner and meta-learner are specified as RNN models [9]. In RNNAI, parameters in base learner and meta-learner are updated alternately. Parameter update in base learner depends upon parameter update in the meta-learner. Conversely, parameter update in meta-learner also depends upon parameter update in the base learner. Communication between base learner and meta-learner makes model training more efficient. Either efficient

base learner or efficient meta-learner should improve training efficiency of the overall meta-learning system. Usually self-update of base learner is posited to be more efficient and meta-learner is updated more slowly to aggregate training experiences.

In highway network [10], different layers in neural network are connected using highway shortcuts allowing information to flow directly from one layer to another. It is often compared with ResNet which has proven to be very effective with thousands of layers and with wide width. ResNet uses a shortcut with identity-mapping between adjacent layers. Identity mapping is critical to guarantee that gradients in backpropagation do not explode or vanish so that neural network can be very deep contributing to its superior predictive performance. Different from ResNet, shortcuts in highway network are controlled by gates, where gate function can be specified as a linear combination of signals. Highway network may contain shortcuts between distant layers which can model long-range dependence. In AutoML, highway shortcuts may be considered as discrete hyperparameters to be optimized. Highway network notably enriches the flexibility in neural network model specifications, which can increase both level of automation and performance of optimal model. Gate functions should be specified in a similar way to the identity mapping used in ResNet. With properly specified gate functions, gradients backpropagation is stable through many layers allowing highway network to be very deep.

3 Coevolution

We know that communication between base learner and meta-learner improves the overall training efficiency of the meta-learning system. Similarly, coevolution between related modules improves the overall training efficiency of the general AI system.

3.1 Association Between Coevolution and Algorithms

Coevolution is common in nature where evolution of one creature depends upon the evolution of another different but related creature. In algorithms, evolution stands for self-improvement or self-update of learner. There are several forms of analogies between coevolution and general AI algorithms. First, coevolution utilizes the communication between self-updates of multiple learners to improve the overall training efficiency of all learners. For example, in meta-learning framework, evolution of base learner depends upon the evolution of meta-learner and evolution of meta-learner also depends upon the evolution of base learner. Second, evolution of one algorithm depends upon the evolution of other algorithms. For instance, after discovering a better feature extraction model, embedding models of former algorithms may be updated with the current best in pursuit of better performance. Third, evolution of solver depends upon the evolution of task. As in AI-GA, generated tasks evolve to be more complex and solver is self-updated correspondingly to settle new tasks. Fourth, evolution of future algorithms depends upon evolution of the current best model. For example, AutoML constantly searches for better neural network models to defeat the current best.

As in [2], a qualified general AI model should be placed within the real 'world' where different learners affect each other and objective or subjective information is obtained to make wise decisions. Coevolution between algorithms is such a way to place general AI models within 'world'. Relationship between coevolved model components may be cooperative or competitive to each other or both. A communication mechanism is devised between different algorithmic models and allows the self-improvement processes of different modules to interact with each other. When algorithms affect each other and evolve collectively, overall all algorithms are trained more efficiently. Meta-learning has memory modules to store training results. Coevolution mechanisms are used to model the self-improvement of multiple algorithms. Meta-learning system integrated with a coevolution module can be used to formulate general AI methods.

Coevolution mechanism in algorithms resembles coevolution in nature. [11] reviews all significant contributions to research on the genetic process in coevolution. [11] also studies the genetic determinants of resistance in host and avirulence in parasite. In nature, coevolution processes are driven by selective forces from internal or external sources [12]. Individuals with the highest fitness scores are retained within each generation and perceived as seeds for further mutations. In AutoML, selective forces can be viewed as the current best model to outperform, and the corresponding fitness score is whether new algorithms outperform the current best model. In meta-learning framework, selective forces can be viewed as the meta-loss function to minimize, and the corresponding fitness score is the overall prediction accuracy of current individual algorithms. In algorithms infused with coevolution mechanisms, selective forces and fitness scores depend upon task objectives and can be flexibly specified in general AI methods.

A differential equation model is composed in [12] to describe the underlying genetic process of coevolution between host and parasite. Parasites seek to infect hosts and the selective force is that parasites evolve to be unharmed (avirulent) to hosts. Empirically, resistance is the dominant feature in hosts and avirulence is the dominant feature in parasites. In nature, selection pressure favors these empirically dominant features, and the corresponding fitness scores hinges upon

observed genetic frequencies of favored features. In algorithms, selection pressure resembles the regularity terms in objective function which alleviates over-fitting and restricts model complexity to be moderate. Genetic processes of coevolution rely upon relation between host and parasite, which can be empirically estimated from metapopulations using genetic frequencies. Genetic process behind coevolution is a linear combination of genetic processes for resistant host and avirulent parasite, resistant host and virulent parasite, susceptible host and avirulent parasite, susceptible host and virulent parasite. Current coevolution process determines the future dominant features.

To see the influence of coevolution concept upon algorithms, we reveal the analogies between natural coevolution and algorithmic models. In meta-learning framework, meta-learner can be seen as 'host' and base learner can be seen as 'parasite'. From coevolution perspective, objective is to maximize fitness score (overall predictive performance) of 'host' and 'parasite' and to maximize avirulence (task-specific performance) of 'parasite'. From algorithmic perspective, objective is to maximize overall predictive performance of 'host' and 'parasite' and to maximize task-specific performance of 'parasite'. In AutoML, the current best model can be seen as 'host' and new algorithms to be tested can be seen as 'parasites'. From coevolution perspective, objective is to maximize resistance (current best model performance) of 'host' and to maximize avirulence (outperform current best model) of 'parasites'. From algorithmic perspective, objective is to maximize current best model performance of 'host' and to maximize the excess over current best model performance of 'parasites'. In AI-GA, task can be seen as 'host' and the corresponding solver can be seen as 'parasite'. From coevolution perspective, objective is to maximize avirulence (task-specific performance) of 'parasites'. From algorithmic perspective, objective is to maximize task-specific performance of 'parasites'. In AI-GA, task evolution may be specified to be open-ended, bounded in cycles or conditional upon solver evolution. Eventually, complex tasks that cannot be solved by training from scratch can be settled by AI-GA.

Emergence of new parasite such as new pathogen can be explained in coevolution models [13]. As mentioned earlier, current coevolution model determines the most probable relation between host and parasite in the future. Whether an unseen pathogen will emerge depends upon the tradeoff between transmission and avirulence of pathogens. Related features include the number of infected hosts, death rate and recovery rate of infected hosts, death rate of hosts subject to other causes, etc. It may take a long time before a new pathogen emerges. Likewise in general AI models, it may take a long time before a valid solution is reached. Analogous to transmissibility and avirulence of parasites, a valid general AI model should settle diverse tasks (transmissibility) with high performance (avirulence).

Victim-exploiter coevolution in nature is summarized to be in four types, as in [14]. First, **predator-prey** is a zero-sum game where predator maximizes number of prey captured and prey minimizes its loss to predator. In game theory, more types of game models may be applied to depict relation between predator and prey. In algorithms, both predator and prey are seen as learners. Through game-theory optimization of reward functions, predator and prey learners achieve the best outcome under the constraint from each other. Second, **host-parasite** is a system where host evolves to be resistant to parasite and parasite evolves to be unharmed to host. Third, **host-pathogen** is a system where host is resistant to pathogen and pathogen is harmful to host. Fourth, **plant-herbivore** is a system where plant is prey and herbivore is predator. There are also other forms of coevolution. An example of coevolution in birds is **brood parasitism** where one species lays eggs in another species' nests [15]. Disguise of eggs affects the detection rate of fake eggs and may result in rejection of eggs by another species. Coevolution dynamics are represented by genetic processes where hyperparameters vary in cycles to limit feature size and model complexification growth. Only restricted model complexification growth exists in nature. In general AI algorithms, hyperparameters may be posited to vary in cycles so that model complexity growth is bounded.

3.2 General AI Algorithms Based Upon Coevolution

Previous section introduces the analogies between general AI algorithms and natural coevolution phenomenon based upon intuitive descriptions. In this section, applications of coevolution in general AI algorithms are summarized in more detail laying out two main frameworks: coevolution between learners, coevolution between task and solver. Generally, too many coevolution modules contain too many rounds of performance evaluations leading to lower algorithmic efficiency. Too much competition or improper cooperation between groups may harm the self-improvement of learners. However, properly specified coevolution modules are encouraged in general AI models.

3.2.1 Coevolution Between Learners

Coevolution mechanism in general AI framework may be in two forms. One is the coevolution between machine learning algorithms [16]. Another is the coevolution between task and the corresponding learner, as in POWERPLAY [17] and AI-GA [6]. Coevolution is not limited to these two forms. This section uses these two coevolution specifications as examples to illustrate the contribution of coevolution to general AI systems. Coevolution models the interaction between related components in general AI methods. Coevolution may be posited to be in different explicit forms based

upon the concrete task. Coevolution makes use of the mutual information shared between coevolved modules and improves the overall training efficiency of the whole general AI machine.

[16] proposes using coevolution between learners to train neural network models and to achieve general AI. Coevolution evaluates every machine learning algorithm in the 'world' (memory module), which consists of previous learners. Relation between learners is established through the following channels: (1) comparison of learner performance on diverse tasks, (2) interaction effect upon each other's neural model evolution (self-update), etc. For example, objective function of one learner may depend upon performances of cooperators and competitors. Self-improvement of one module may be associated with self-updates of other modules. In coevolution, learners form a 'social network' where they compete, cooperate, communicate, coevolve, etc. This 'social network' of learners improves the training efficiency of all learners overall.

In evolutionary multi-objective optimization, evolving neural network models are seen as learners, and testers are trained neural network models used to evaluate learners. The objective is to find a learner that outperforms all testers. A good learner is expected to surpass as many testers as possible. In Hall of Fame (HOF) type of memory archive, the best learners from past generations are stored in Coevolutionary Memory (CM) and are used as testers on future learners. CM module used in [16] is Layered Pareto Coevolution Archive (**LAPCA**), another type of memory archive which contains not only testers but also learners. Learners in LAPCA are ordered through Pareto-dominance. Learner A Pareto-dominates learner B if learner A outperforms not only all testers defeated by learner B but also other testers. Pareto front consists of the best learners never Pareto-dominated. Under proposed ordering of learners, Pareto front collects all best learners from previous generations. LAPCA contains efficient memory size and conducts efficient performance evaluations of generated learners.

NeuroEvolution of Augmenting Topologies (NEAT) is applied to produce efficient evolutions (self-improvement) of neural network models. In coevolution, neural network model complexity is in monotonic growth and performance evaluations of one model depend upon all other models in the memory archive. Testers may be randomly drawn from CM to evaluate performances of existing learners. In LAPCA, learners that never Pareto-dominate others and testers that fail to distinguish any other learners are discarded. Only learners in the top layers under Pareto-dominance ordering are retained in LAPCA. The general AI algorithm formulated in [16] is an evolutionary algorithm combined with a memory archive to store trained learners, such as NEAT with LAPCA. After training under this framework, best neural network models are present in LAPCA. This general AI algorithm is applicable to neural network models and other machine learning models. It provides an automated solution to diverse tasks. In conclusion, this general AI framework consists of a self-improvement algorithm and a corresponding memory module, where learners form a 'social network' and perform self-updates collectively through coevolution mechanism.

3.2.2 Coevolution Between Task and Solver

On the other hand, **POWERPLAY** [17] considers the coevolution between task and solver rather than the coevolution between learners [16]. **POWERPLAY** concentrates attention upon task-solver pairs, where the solver tackles diverse generated tasks with high performance. Trained solver is a general AI algorithm that can settle different tasks efficiently in an automated way. Upon an unseen task, the whole general AI system is self-updated to solve old tasks and the new one. Since the set of unseen tasks is very large, coevolution between task and solver can make the learner to be self-updated too frequent and thus inefficient in solving diverse problems. The threshold for learner self-update may be specified to be higher or adaptive for improved efficiency. **POWERPLAY** consists of three components: task invention, solver modification, and correctness demonstration.

First, **task invention module** generates tasks and finds the simplest task that is unsolvable by old solver. Task invention simulates human inner-driven curiosity to pursue challenging tasks and to achieve self-improvement by continuously solving more difficult tasks. Task invention is the most critical component in **POWERPLAY** since it uses generated tasks to guide the self-improvement of learners. Task invention constitutes a meta-learner which acts as the 'brain' to decide what kinds of further training can best improve the key aspects of current solver. In the end, task invention module identifies failure cases of current solver and offers the simplest failure case as a challenge. In [16], motivation of learner's self-improvement is the external competition or cooperation from other learners in the same 'social network'. Here in **POWERPLAY**, motivation of learner's self-improvement is the internally self-generated failure cases of current solver.

Second, **solver modification module** takes as input the simplest unsolved task from task invention module, performs solver self-improvement, solves all old tasks and the newly composed failure case. Similar to [16], a memory archive and a self-improvement algorithm are equipped in **POWERPLAY**. But here the continually growing memory archive contains task-solver pairs and the self-improvement algorithm is present in solver modification module. Finally, **correctness demonstration module** illustrates that self-updated solver tackles all old tasks and the new challenging task proposed by task invention module with desirable performance. Correctness demonstration can be inefficient since

the archive of all old tasks can be very large and evaluations should be performed on all old tasks every time the solver is self-updated. We may avoid this trouble by not requiring the self-improved learner to retain good performance on all old tasks. Instead we revise the general AI system to provide a light-weight solution to every new task encountered in an automated fashion. The light-weight solver only needs to solve the currently generated task well. General AI system is capable of providing such light-weight solvers to diverse tasks. Though computation of a light-weight solver is more efficient, updating the whole general AI system can still be slow.

As a meta-learning framework, POWERPLAY utilizes former model training experiences to accelerate solver self-improvement. Task similarity can be explicitly formulated so that trained models from the most similar tasks can be utilized directly to propose a good initial model. POWERPLAY accounts for both internally self-generated tasks and external tasks, which are considered as challenges to the current solver. Since task generation guides the self-improvement of solver, external tasks represent motivation from external forces that drive the solver improvement. The concept of AI-GA is analogous to POWERPLAY, where the coevolution between task and solver is considered. The difference is that both coevolution between learners and coevolution between task and solver are utilized in AI-GA.

4 Curiosity

Similar to coevolution, curiosity is another concept from meta-learning that can contribute to general AI. Coevolution creates a joint system for associated modules to perform self-improvement collectively allowing learners to communicate, compete, cooperate, etc and allowing learners to coevolve with self-generated tasks. Appropriately specified coevolution improves training efficiency of the whole general AI system. Curiosity can be applied in self-improvement of learners to circumvent local optimas and reach global optima. Curiosity mechanism can be infused within coevolution modules to improve efficiency of learner self-update.

4.1 Association Between Curiosity and Algorithms

In coevolution-based general AI methods, learner update is influenced by the coevolution between learners or the coevolution between task and solver, which simulates the real world in algorithms and introduces aggregated automation. In addition to coevolution, curiosity is another critical contribution from meta-learning realm to general AI. Curiosity and coevolution can be implemented together to design versatile general AI algorithms for all types of tasks. Curiosity makes the AI algorithm more general by encouraging exploration of more diverse features and more efficient by avoiding over-exploration of close spots [18]. However, diverging too far away from current exploration spot may not be wise when we are already close to global optimum. At the early stage of exploration, encouraging curiosity can effectively contribute to avoiding local optimum traps.

Artificial curiosity (AC) can be defined as the unexpectedness of an event [18]. For a predictable event, any outcome that is distinct from the most probable outcome is termed as unexpected. Therefore the pre-requisite of unexpectedness is the predictability of an event, the outcome distribution of which can be well estimated. Curiosity may be integrated into the objective function where unexpectedness is maximized so that learners explore more diverse features of the search space. For example, [19] adds curiosity or creativity reward in the objective function of reinforcement learning so that robots' behaviors are more diverse. Novel metrics measuring algorithmic predictability may be devised to compute curiosity efficiently in general AI algorithms. During the self-improvement of learners, predictability of learner outcome can be examined along with performance evaluations so that unexpectedness in curiosity is measured timely and accurately. For trained deep neural network, unexpectedness often corresponds to failure cases of deep model. In these failure cases, outcomes cannot be predicted accurately using the trained deep model. Curiosity alleviates over-exploration of explored spots and encourages learners to discover new features ignored by other learners. Since curiosity may turn out to be harmful sending learners to wander purposelessly in the search region, general AI methods should contain adaptive criteria to control curiosity.

Traditionally, evolution is applied to conduct self-improvement of learners. But evolution is often criticized for not being sufficiently efficient. Later in pursuit of higher performance, evolution is useful in complex optimization tasks where evolution can alleviate the influence of local optimas. Evolution helps skip local optimum and reach global optimum. In deep learning, to control over-fitting, regularity terms are included in the objective function so that the deep model does not grow to be too complex. Multiple objective functions may be combined in the optimization to balance several objectives simultaneously in the same deep model. Here the objective function for evolution (learner self-improvement) may be solely based upon the curiosity criterion, which guides the learner to explore as diverse spots as possible all over the search space [20]. Performance is evaluated across all learners explored through curiosity mechanism and the best learner is identified.

Objective functions can be defined with or without curiosity terms. Traditionally the objective function is defined to be prediction error, reward, etc. The benefits of using curiosity-based objective function are as follows. First, for different tasks, performance-based objective functions are different but curiosity-based objective functions can be defined in the same way by maximizing the distance between exploration spots. For the same task, the objective function is not unique and may be defined in multiple ways. Curiosity-based objective functions make general AI algorithms applicable to a wider range of tasks. Second, most real-life information is obtained and analyzed by human without a clear objective in mind. For example, human sees birds in a forest, recognizes different types of birds and saves information in memory. This process is not driven by any clear reward or objective. Curiosity-based objective functions simulate such unintentional human learning activities which are purely out of curiosity. Third, sometimes objective functions can be short-sighted and may misguide learners to be trapped in local optimum [20]. For instance, there is a wall intercepted right at the shortest path from start to end. How do learners detour based upon the shortest path motivated by the objective function? In this case, objective functions help trap learners in this local optimum unable to identify a better path. Curiosity-based objective function excels at avoiding local optimum and has a better chance of reaching a desirable solution. Fourth, curiosity-based objective forces learners to explore diverse features of search space so that the algorithm has a deeper insight into the overall conditions around the space. Learners accumulates critical knowledge and skills which can be referred to later in other tasks. With curiosity-based objective, memory collects more information and skills which can facilitate model generalization to a wider range of tasks. Fifth, similar to POWERPLAY, curiosity-based objective is applied to several self-generated challenging tasks and model training experiences are saved in meta-learner. After long exploration, meta-learner accumulates sufficient amount of model training experiences which can be used to solve more complex tasks.

However, in practice, merely rewarding novelty metrics is not adequate especially in reward-driven cases where maximizing profits is required. In such cases, curiosity-based objective function makes learners wander in the search space and cannot derive an efficient search algorithm for the global optimum. In general AI, performance-based and curiosity-based objectives may be combined to achieve better performance and efficiency. Performance-based and curiosity-based optimization can be combined in the following ways. First, the objective function may be specified as a linear combination of performance-based and curiosity-based terms. In optimization, the objective function is maximized where performance-based and curiosity-based criteria are considered simultaneously to update learners. Second, a mixture procedure may be applied with steps alternating between performance-based optimization and curiosity-based optimization. One performance-based learner update is followed by another curiosity-based learner update. Proportion of performance-based steps is adaptive and depends upon the profit-seeking degree of the given task. Curiosity-based steps are primarily at the early stage of exploration and can help avoid local optimum. Third, curiosity-based objective acts as a generative model of proposed exploration spots, and performance-based objective is the 'meta-learner' that screens these proposed spots to identify the most promising one. Combination of performance-based and curiosity-based objectives may occur in other forms than these three forms. Infusion of curiosity in performance-based optimization can be much more versatile.

Generally, it is not efficient to focus upon either performance or curiosity alone. General AI methods should include both performance-based and curiosity-based ingredients. For reward-driven tasks, performance-based modules dominate curiosity-based ones. For exploring activities purely out of curiosity, curiosity-based modules dominate performance-based ones. In the previous section, we summarize that coevolution is an indispensable component in general AI. Coevolution links related modules and performs self-improvement of these parts collectively to improve overall training efficiency. Curiosity is often applied jointly with coevolution such that self-improvement of learners can avoid local optimum and reach global optimum improving learner performance. However, open-ended evolution and unbounded curiosity may adversely affect the training efficiency of general AI algorithms. In curiosity-based modules, novelty metrics should be carefully devised to pursue preferred novelty behaviors of specific tasks, which depend upon tacit knowledge obtained through human intervention [20]. Though well-defined novelty metrics may contribute significantly to improved performance, seeking global optimum is still time-consuming especially for large unexplored search space and sparse solutions. To our knowledge, if all local optimas are found, then the optima of all local optimas is the global optima. The challenge is that learners can identify one or several local optimas, but not all local optimas. However, as we find more local optimas, we are closer to the global optima.

4.2 General AI Algorithms Based Upon Curiosity

Under meta-learning framework, meta-learner contains a curiosity-based generator of exploration spots. Meta-learner aggregates model training experiences, provides performance-based evaluations of proposed exploration spots, and pinpoints the most promising future exploration spots. Meta-learner increases the efficiency of learner self-improvement and reduces the risk of local optimum. Performance-based objective is often trained with stochastic gradient descent, and curiosity-based objective is usually applied in evolution or coevolution based self-improvement. Therefore, in combination of performance-based and curiosity-based objectives, learner self-update hinges upon evolution or

coevolution. It is known that the optima of all local optimas is the global optima. The more local optimas we manage to identify, the closer we are to the global optima. In the search space, local optimas and peculiar points (boundaries, non-differentiable spots, etc) are potential candidates of the global optima, and they deserve special attention in the search algorithm.

For exploration spots proposed by the curiosity-based objective, performance-based objective is evaluated directly, or performance prediction using meta-learner is conducted. Only spots with high performance evaluations or predictions are explored in the future. In novelty-driven search, base learner keeps exploring more spots and meta-learner accumulates performance evaluations of explored spots. In maze problem, performance-based objective is minimization of distance travelled from start to end, and curiosity-based objective is maximization of distance between explored spots and spots to be explored [21]. Since both performance-based and curiosity-based criteria are based upon the same distance measure, outcomes from these two criteria are similar. Novelty metric in maze tasks may be defined in the following ways: (1) area covered by explored spots [21], (2) distance covered by explored spots [21], (3) density of explored spots all over the search space [21], (4) distance between endpoints of explored trajectories [22], (5) distance between explored trajectories [22], etc. Curiosity-based criteria should be devised carefully in the most appropriate way to improve search efficiency of the global optimum. The most appropriate novelty metric is defined based upon tacit knowledge which relies upon human interpretation of task information and cannot be utilized in an automated way to construct general AI methods.

There are two types of curiosity-based search: **novelty search** and **curiosity search** [22]. Novelty search encourages individual learners to explore different spots from others. Novelty search trains learners to solve a particular type of tasks well and to be specialists in an area. Curiosity search encourages each individual to be versatile and to learn as many different exploration skills as possible. Curiosity search trains learners to acquire more types of skills that can be generalized to a wider range of tasks. Learners trained with curiosity search are more capable and can tackle challenges much more efficiently. Novelty search can be combined with curiosity search to produce learners with maximum number of different skills and maximum behavioral diversity for each skill. For example, in maze tasks, novelty search guides learners to explore spots far away and curiosity search rewards learners for acquiring different exploration patterns. There are two components in curiosity search [22]: intra-life novelty score and fitness function. Intra-life novelty mechanism assigns scores to different exploration patterns of each individual. Fitness function is an objective that maximizes the diversity of individual exploration patterns. Exploration patterns depend upon the following aspects: (1) performance-based objectives, (2) curiosity-based objectives with several novelty metrics, (3) coping with local traps or mechanical emergencies, etc. For instance, in maze tasks, exploration patterns contain opening different types of doors, grabbing objects, avoiding stepping on stones, walking, running, jumping, etc.

Performance-based objectives can be integrated into novelty search and curiosity search so that trained learners acquire diverse skills and perform well at the same time. Intra-life novelty identifies behaviors that can substantially improve understanding of tacit knowledge within the specific task and increase the survival probability of learners. Searches purely relying upon curiosity seek to accumulate more skill sets in memory, and often ignore performance evaluations. Curiosity-based search is exploratory, seeks self-generated or external challenges, intends to build more capable and more general learners. With novelty search, curiosity search and performance-based evaluations on diverse tasks, general AI system can adapt deep models to solve vastly different tasks. Since trained deep learners already possess multiple necessary skills to solve a wide range of tasks, model adaptation will be fast and efficient.

In learner optimization, performance-based and curiosity-based objectives can be combined to achieve the best search efficiency and to avoid local optimas. Model generalization capability is closely associated with learner optimization technique. In practice, generalization capabilities of curiosity-based approaches are higher. The reasons are as follows. Under performance-based objective, model generalization should also be conducted using performance-based methods. Under curiosity-based objective, model generalization should be executed using curiosity-based adaptation algorithms. Since curiosity-based approaches accumulate more model training experiences including both exploration patterns and diverse behaviors, it is probable that curiosity-based model adaptation can solve more diverse tasks including vastly different ones.

On the other hand, model generalization capacity is related to complexity measures of deep neural network models. For deep models, [23] summarizes the upper bound of generalization capacity based upon model complexity. As neural network complexity increases, the upper bound of model generalization capacity also rises. Neural network complexity is positively associated with the number of weight parameters and the number of layers. When we expect general AI systems to solve vastly different tasks, we may choose deeper and wider neural network models. It is pointed out that the generalization capacity of unsupervised learning models using unlabelled data will generally be higher than supervised learning models [23]. Similarly, the generalization capacity of self-supervised learning models using unlabelled data will also be higher overall. Moreover, the generalization capacity of learners using randomly generated data labels will also be higher than learners using true labels. Taking advantage of all these phenomena, we may continually improve

the generalization capacity of a deep model until it approaches the generality of a general AI method and eventually becomes a general AI algorithm.

5 Forgetting

We have summarized coevolution and curiosity which contribute to the construction of general AI methods. This section concerns forgetting which is also an indispensable component in general AI.

Why do we need forgetting mechanisms? First, meta-learner continually gathers more critical training experiences and forgetting helps avoid memory explosion. Forgetting discards redundant and obsolete memory items and keeps memory size efficient. With forgetting, we concentrate our attention upon the most relevant and critical pieces of information in memory. Second, forgetting is necessary especially for certain reward-driven optimization techniques. In some reward-driven algorithms, every step has a separate goal of maximizing profits received at that step. But several steps combined may not have maximized the total profits received. Forgetting reverts actions at previous steps to increase combined profits. Third, forgetting mechanisms remove training experiences at exploration spots with high regret or high boredom from memory. Regret of any spot is the distance between any exploration spot and the global optima. When regret is high, these exploration spots are far away from global optima and forgetting mechanisms may discard the corresponding training experiences. Boredom of any spot is the similarity between this spot and spots already explored. When boredom is high, meta-learner has already accumulated too many similar training experiences at similar spots and exploring this spot is no longer intriguing.

In HOF or LAPCA type of memory archives within coevolution modules, forgetting mechanisms remove obsolete or dominated learners to maintain high-level competition between learners [24]. In meta-learning, forgetting mechanisms are installed in memory modules to avoid memory explosion and to improve search efficiency. Forgetting mechanisms adaptively remove memory items by forecasting which components in memory have the lowest probability of ever being used again. Forgetting mechanisms can also discard the oldest memory when long-range dependence is not considered in the model. When too many similar items are saved, forgetting mechanisms can delete redundant information without harming the representativeness of memory. But we should be careful that forgetting mechanisms would not underestimate any important information. Forgetting mechanisms perform feature extraction and information distillation on memory items. Saving memory size makes it faster to search for critical information from memory. Forgetting mechanisms can also use dimension reduction techniques to make feature size more efficient. Forgetting mechanisms are indispensable components in creating general AI algorithms.

Forget gates in LSTM are very effective and are major contributions to its superior model performance [25]. Curiosity can be seen as another form of forgetting since it minimizes (forgets) the influence of prior exploration spots upon current spots. Lars and Lasso are both solutions to optimization of the objective function with L1 regularization. Lasso includes a forgetting mechanism which deletes previously added feature variables for improved model performance. But Lars has no forgetting mechanism and keeps all added features in the selected model. In practice, Lars is better than Lasso in some datasets and Lasso can also be better than Lars in others. From this perspective, including forgetting mechanisms in algorithms may not be better necessarily. Theoretically forgetting mechanisms bring learners closer to global optima by showing regret on former actions which have not turned out to be beneficial. By re-visiting previous decisions, future information is applied to re-judge former actions and to correct sub-optimal choices.

Decomposition of knowledge representation can be used to construct concrete steps in specific tasks. Correspondingly decomposition of global objective can be applied to formulate sub-goals in each concrete step. At each step, sub-parts of knowledge are exploited to realize sub-goals through optimization. Under appropriate decomposition scheme, fulfilling all sub-goals should reach the global optima of the final objective. However, in practice, task decomposition may not provide steps that lead to global optima. The reasons are as follows. For computation efficiency, task decomposition may be approximated with simpler mechanisms. Sub-parts of knowledge at each step may be insufficient, biased or misleading. Coevolution modules within decomposed steps may be improperly devised such that computation error at each step may accumulate and explode making algorithms unstable. Following decomposed steps may reach other optimas that are approximately close to global optima. Embedding forgetting mechanism allows re-visiting previous steps after future performance is observed, and making convenient adjustments for better ultimate performance. Apart from forgetting mechanisms, several decomposition schemes may be combined and optimization at each step may communicate with each other for improved performance.

6 Conclusion and Discussion

This paper briefly summarizes connections between meta-learning and general AI, covering concepts such as coevolution, curiosity and forgetting. Meta-learning is applicable not only to few-shot deep learning, but also to general AI. Generally,

there are many existing paths to realize general AI and meta-learning makes important contributions by supplying useful tools such as memory module, meta-learner, coevolution, curiosity, forgetting, etc. Though algorithms are not as smart as human intelligence, at least general AI methods strive to approach human wisdom. Hopefully general AI can be realized bringing automation into solving simple, time-consuming, dangerous, laborious or repeated tasks.

In general, flexible neural network designs offer diverse options to process information flow and evolutionary algorithms realize stable efficient self-improvement of learners. Curiosity search encourages learners to avoid local optima and to acquire diverse skill sets. Forgetting improves the efficiency of memory module and corrects previously made sub-optimal decisions. Coevolution positions self-improvement of learners in a 'social network' of learners and considers competition, cooperation, communication, etc among learners. Curiosity metrics, objectives and forgetting mechanisms are often best determined using tacit knowledge. Human interpretation is required to build tacit knowledge within specific tasks into general AI models. The real world is sophisticated and impossible to simulate. General AI algorithms cannot be close to human performance in all kinds of tasks. But hopefully with tacit knowledge built into algorithms, general AI can at least solve a wide range of tasks in an automated way.

References

- [1] Naveen Joshi. How Far Are We From Achieving Artificial General Intelligence?, 2019.
- [2] Ragnar Fjelland. Why General Artificial Intelligence will not be Realized. *Humanities and Social Sciences Communications*, 7(1):1–9, 2020.
- [3] Marc G. Bellemare, Yavar Naddaf, Joel Veness, and Michael Bowling. The Arcade Learning Environment: An Evaluation Platform for General Agents. *IJCAI International Joint Conference on Artificial Intelligence*, 2015-January:4148–4152, 2015.
- [4] Christian Szegedy. *A Promising Path towards Autoformalization and General Artificial Intelligence*, volume 12236 LNAI. Springer International Publishing, 2020.
- [5] Huimin Peng. A Comprehensive Overview and Survey of Recent Advances in Meta-Learning. 2020.
- [6] Jeff Clune. AI-GAs: AI-Generating Algorithms, an Alternate Paradigm for Producing General Artificial Intelligence. *arXiv*, 2019.
- [7] Jürgen Schmidhuber. Deep Learning in Neural Networks: An Overview. *Neural Networks*, 61:85–117, 2015.
- [8] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780, nov 1997.
- [9] Jürgen Schmidhuber. On Learning to Think: Algorithmic Information Theory for Novel Combinations of Reinforcement Learning Controllers and Recurrent Neural World Models. pages 1–36, 2015.
- [10] Rupesh Kumar Srivastava, Klaus Greff, and Jürgen Schmidhuber. Highway Networks. 2015.
- [11] J.J. Burdon and J.N. Thompson. Gene-for-Gene Coevolution between Plants and Parasites. *Nature*, 360:121–125, 1992.
- [12] R. M. Anderson and R. M. May. Coevolution of Hosts and Parasites. *Parasitology*, 85(2):411–426, 1982.
- [13] R. M. Anderson. Parasite–Host Coevolution. *Parasitology*, 100(S1):S89–S101, 1990.
- [14] J. Bergelson, G. Dwyer, and J. J. Emerson. Models and Data on Plant-Enemy Coevolution. *Annual Review of Genetics*, 35:469–499, 2001.
- [15] Stephen I. Rothstein. A Model System for Coevolution: Avian Brood Parasitism. *Annual Review of Ecology and Systematics*, 21(1):481–508, 1990.
- [16] German A. Monroy, Kenneth O. Stanley, and Risto Miikkulainen. Coevolution of Neural Networks using a Layered Pareto Archive. *GECCO 2006 - Genetic and Evolutionary Computation Conference*, 1:329–336, 2006.
- [17] Jürgen Schmidhuber. PowerPlay: Training an Increasingly General Problem Solver by Continually Searching for the Simplest Still Unsolvable Problem. *Frontiers in Psychology*, 4:1–21, 2013.
- [18] Jürgen Schmidhuber. Artificial Curiosity based on Discovering Novel Algorithmic Predictability through Coevolution. *Proceedings of the 1999 Congress on Evolutionary Computation, CEC 1999*, 3:1612–1618, 1999.
- [19] Jürgen Schmidhuber. Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990-2010). *IEEE Transactions on Autonomous Mental Development*, 2(3):230–247, 2010.
- [20] Joel Lehman and Kenneth O. Stanley. Abandoning Objectives: Evolution Through the Search for Novelty Alone. *Evolutionary Computation*, 19(2):189–222, 2011.

-
- [21] Roby Velez and Jeff Clune. Novelty Search Creates Robots with General Skills for Exploration. *GECCO 2014 - Proceedings of the 2014 Genetic and Evolutionary Computation Conference*, pages 737–744, 2014.
 - [22] Christopher Stanton and Jeff Clune. Curiosity Search: Producing Generalists by Encouraging Individuals to Continually Explore and Acquire Skills Throughout Their Lifetime. *PLoS ONE*, 11(9):1–20, 2016.
 - [23] Behnam Neyshabur, Srinadh Bhojanapalli, David McAllester, and Nathan Srebro. Exploring Generalization in Deep Learning. *31st Conference on Neural Information Processing Systems*, 2017.
 - [24] Sevan G Ficici and Jordan B Pollack. A Game-Theoretic Memory Mechanism for Coevolution. *Proceedings of the 2003 Genetic and Evolutionary Computation Conference*, pages 286–297, 2003.
 - [25] Jos van der Westhuizen and Joan Lasenby. The Unreasonable Effectiveness of the Forget Gate. pages 1–15, 2018.
 - [26] B Efron, T Hastie, I Johnstone, and R Tibshirani. Least Angle Regression. *The Annals of Statistics*, 32(2):407–499, 2004.