Towards AI Research Agents in the Experimental Sciences

Ofer M. Shir*

Abstract

An underlying problem shared by experimental scientists is to achieve optimal behavior of their systems and arrive at new discoveries while searching over an array of controls. Accordingly, every scientific discovery may be reduced to solving a Combinatorial Optimization problem upon formulating the characteristic array of decision variables. While AI systems already excel in navigating the landscape of possible experiments, we argue in this perspective paper that they will be able to drive the entire process of scientific experimental research. Especially, in response to the so-called Nobel Turing Challenge, regarding which Kitano envisioned AI Scientists [1], the goal of this paper is to provide a pragmatic roadmap to obtain AI Research Agents in the experimental sciences. We begin by reviewing the existing integration of Computational Intelligence into experimental systems, which already benefit from solving discovery/optimization problems. We mention recent discoveries in the domains of Enzymes' Design, Material Science, Quantum Mechanics, and Postharvest, in which AI systems played active roles in attaining some ground-breaking results – thanks to being conception-free and unbiased by flawed intuition. We then devise a concrete work plan to train agents to formulate hypotheses by Deep Symbolic Reinforcement Learning, using knowledge representations based on processed scientific textbooks. We focus on experimental Physical Sciences with underlying stationary Knowledge Graphs, and propose how to obtain an independent AI system in Chemistry at the graduate student level.

Keywords: Artificial General Intelligence, Scientific Discoveries, Hypothesis Formulation, Reasoning

1 Introduction

The majority of experimental sciences share the common basis of physical observables that may play the role of objective functions to be optimized. Scientists aim at optimal behavior of their systems and arriving at new discoveries while navigating the landscape of possible experiments. Furthermore, it was shown that this perspective enables the **reduction of every scientific discovery to a Combinatorial Optimization problem [2]**, even if the underlying landscape is unveiled in retrospect [3]. Indeed, algorithms and metaheuristics have been successfully applied to navigate in such landscapes of possible scientific experiments [4], with quite a few success stories that may be considered as scientific discoveries. Examples include but are not limited to combinatorial drug discovery [5], enzymes production [6], as well as recent successes in discovering protocols for protein expression [7] and for Postharvest treatment of fresh cucumbers [8]. Although this AI-driven sequential experimentation perspective is not novel from the algorithmic point of view [4], it is of high importance when mounting AI capabilities toward conducting scientific research. In practice, it shapes an AI tool as a facilitator for effective navigation towards discoveries, as long as the experimental system's controls were correctly set up *a priori*. Capitalizing on these capabilities, it is expected that more experiments will be algorithmically-guided in the future, and accordingly, the roles of the human scientists/engineers will shift from locating solutions/designs to explaining the nature of the attained results while aiming for mechanistic understanding. The rise of ML capabilities, especially since the Deep Learning revolution [9], has positioned the machine in the center of scientific research. With data mining and pattern recognizing becoming traditional tasks, ML models are being increasingly integrated with human knowledge, a trend which altogether boosts their capabilities in aiding scientific research [10]. Some overwhelming systems – being the outcome of long research, rich datasets and clever integration – have been recently announced in the forms of a *reproducer* of scientifically published results [11], an accurate proteins' structure *predictor* [12], as well as an effective molecular de novo *designer* [13]. We thus question whether more decisions in scientific experiments may be driven by the machine, even to the extent of an AI system that independently leads scientific research.

*Tel-Hai College and Migal Institute, Upper Galilee, Israel: ofersh@telhai.ac.il.

Contribution and Paper Organization

While this challenge of constructing a *research agent* resides under the broad umbrella of Artificial General Intelligence [14], with a generalized goal to obtain a machine that thinks, learns and infers like humans [15, 16], we are interested in the specific capabilities of drawing scientific hypotheses and conducting experimentation to approve/disapprove them. Indeed, the concept of an AI Scientist has already been put forward by Kitano [1], who also formulated the so-called Nobel Turing Challenge. In his comprehensive article, Kitano envisioned a machine capable of reaching ground-breaking scientific discoveries by generating hypotheses' spaces and searching them as open-ended problems. His vision was oriented towards the Life Sciences, where knowledge bases are ever evolving and often suffer from inaccuracies and/or "descriptions that are not well-defined" (i.e., resulting in uncertain and non-stationary knowledge graphs). Until that dream is brought to fruition, we would like to propose a pragmatic roadmap to obtain AI research agents in the near future within the Physical Sciences. Hence, the contribution of this short perspective paper is to devise a concrete work plan, while focusing on disciplines with stationary knowledge graphs.[1] We capitalize on established knowledge representation frameworks [17] and on Deep Symbolic Reinforcement Learning [18, 19]. The concrete contributions of this paper are the following:

- We outline recent AI-driven experimental discoveries in the domains of Enzymes' Design, Material Science, Quantum Mechanics, and Postharvest. We elaborate on the recent Postharvest discovery.
- We propose a pathway toward an AI framework that serves as a member of a research group - that is, drawing scientific hypotheses from stationary knowledge graphs, and then testing them by means of existing capabilities of algorithmically-guided experimentation. We define a roadmap to accomplish such an AI system in Chemistry.

The proposed evolution of AI within *traditional*↦*modern*↦*future* scientific experimental research is illustrated by a diagram depicted in Figure 1. Importantly, we are not considering at any stage machines capable of discovering new Laws of Nature.

The remainder of this paper is organized as follows. Next, Section 2 describes AI-driven scientific research. Then, Section 3 presents a proposed pathway toward machines which entirely play the researchers' roles.

2 Modern AI-Driven Scientific Research

Scientists aim at attaining optimal behavior of their systems, as measured by a well-defined *assay*, while searching over an array of parameters (decision variables) [2]. The underpinning is thus the existence of strictly experimental objective functions (observables), while the search/learning algorithms play the role of computational tools for addressing the scientific research questions at hand. In the broad Computational Intelligence context, *experimental optimization* is linked to the concept of Embodied Artificial Evolution [21], which brands *"evolution of things"* as a process driven by real-world operations of variation and selection [22]. Statistics-based approaches are considered the gold standard in the domain of experimentation, particularly the family of Optimal Design procedures and the Design of Experiments (DoE) methodology [23]. In fact, DoE modules are often hard-coded into standard laboratory robotics as the default automation procedure, and are practically associated with *experimental optimization*, although their link to global optimization (or, within the current context, to navigation in the landscape of experiments), has not been established. At the same time, the applicability of Computational/Artificial Intelligence to experimental optimization has already been demonstrated in the early days of the Evolutionary Computation field [24]. Ever since, Evolutionary Algorithms have been continuously utilized in a number of studies on experimental optimization within the Natural Sciences [4]. Figure 2 provides an illustration for algorithmically-guided experimentation in the domain of Quantum Mechanics [20]. It should also be noted that Bayesian Optimization models have also been successfully applied to such experimental systems (see, e.g., [25]). Next, we present multiple use-cases of scientific discoveries, in which the machine played active roles, while distancing from prior knowledge and human intuition.

Recent AI-Based Scientific Discoveries

In what follows, we mention several recent scientific discoveries that were enabled by the advent of modern AI. We particularly elaborate on a recent discovery in the domain of Postharvest.

[1] Our consideration is focused on the Physical Sciences, and therefore, in Kitano's terms and analogies, it is oriented towards sciences with solid and well-defined descriptions that may allow treatment similar to AI in Games (i.e., enjoying a more restrictive hypotheses' space, and being able to define so-called *game states*.

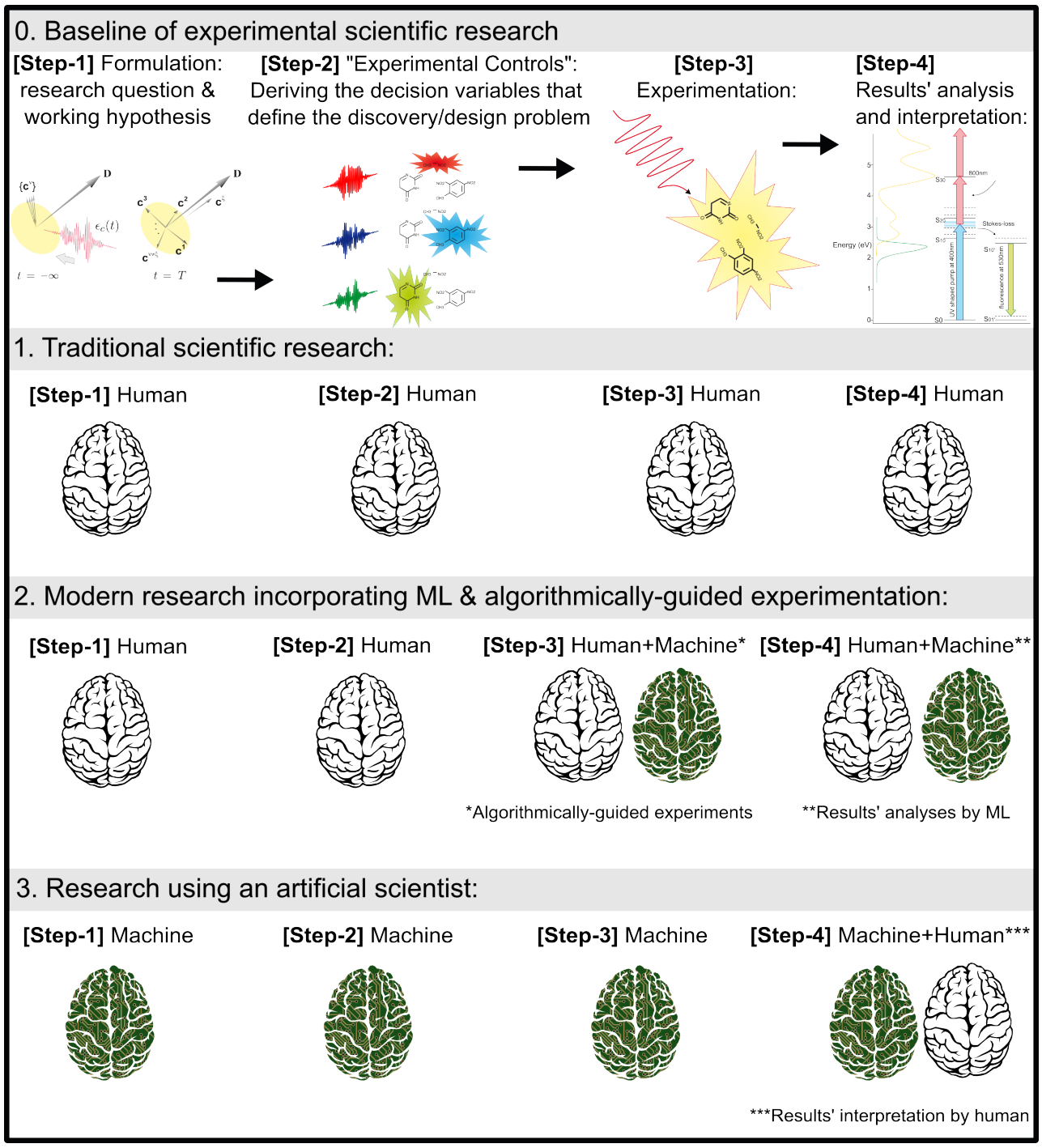

Figure 1: A visual summary of the proposed AI's evolution within traditional↦modern↦future scientific experimental research using a process abstraction of four steps. The steps are presented in the top row (0) using illustrations from quantum experiments [20]; they are followed downward by (1) traditional, all-human, research; (2) modern, machine-aided research, and (3) the envisioned AI research agent. The human role is depicted by a natural brain (black-and-white), while the AI's role is depicted by a digitized brain cartoon (green).

Enzymes' Design Following developments in computational biology and the ML breakthrough in AlphaFold's ability to determine 3D shapes of proteins given their genetic sequences [12], de novo design of functional proteins is nowadays enabled to an extent that was unattainable heretofore. The creation of highly active and specific biocatalysts from scratch with broad applications in biomedicine is a key milestone for computational enzyme design. Indeed, a recent discovery of a novel enzyme was reported as a joint process of a team of human researchers

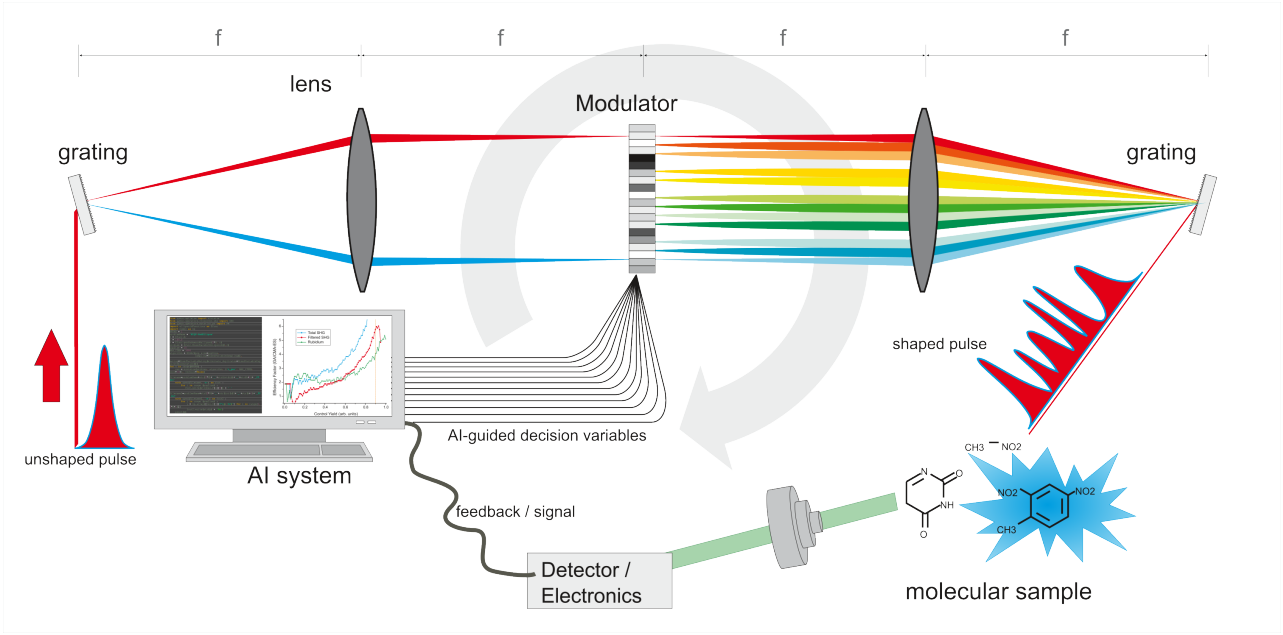

Figure 2: An illustration for algorithmically-guided quantum experiments in modern research (corresponding to [Step-3] in "2. Modern research" in Figure 1). An unshaped laser pulse, approximated by a Gaussian, is shaped by a pixel-based modulator and applied to a molecular sample. The decision variables ("experimental controls"), addressed in the laboratory by an Evolutionary Algorithm, are the individual pixels that determine the pulse shape. The measured signal reflecting the molecular response constitutes the feedback for optimization.

hand-in-hand with deep learning tools [13]. The human researchers first selected chemicals (entitled *luciferins*) that they wanted the proteins to react with, and then used an AI system to design new luciferases that could catalyze the reaction between the luciferins and oxygen.

Synthetic Chemistry The usage of an unsupervised learning algorithm enabled researchers in the domain of Material Science to discover novel composites [26]. Here, addressing the challenge of selecting an effective subset of chemical elements constituted the game-changer in the discovery of novel synthetic compounds, wherein the lack of prior knowledge and human intuition was again an advantage.

Quantum Mechanics In the domain of Quantum Mechanics, the so-called MELVIN algorithm [27] was devised to prescribe optical setups (i.e., specify laboratory implementations for concrete experimentation) on which quantum experiments are likely to locate and manipulate complex quantum states (e.g., to discover entanglement states; see also [28] for its extension). The encoding of the underlying scientific knowledge was obtained by means of symbolic algebra, and the experimental building blocks were manually encoded as elements within a toolbox. Then, a dedicated search algorithm explored the toolbox, proposed setups using the elements, calculated the expected states, and assessed their validity using predefined criteria. This approach was successafully corroborated by means of laboratory experiments [27, 28].

Postharvest Postharvest refers to the collection of practices for handling crops immediately following their harvest, with the explicit goal of maintaining their quality, while boosting their shelf-life [29]. In what follows, we elaborate on a recent scientific discovery, where an AI system devised Postharvest protocols for cucumbers (Cucumis Sativus L.) [8]. Explicitly, given n_t postharvest *treatments*, as well as a set of postharvest *packages*, the generalized Combinatorial Optimization Postharvest problem is defined by $(n_t + 1)$-dimensional decision variable vectors, denoted as $\vec{\tau}$, representing candidate treatment protocols. These vectors encompass integer variables, and constitute permutations over combinations that subscribe to a discrete (mixed nominal and categorical) space $\mathcal{T}$: $\vec{\tau} \in \pi \circ \mathcal{T}, \quad \pi \in P_{\pi}^{(n_t)}, \quad \mathcal{T} = \mathcal{T}_1 \times \mathcal{T}_2 \times \cdots \times \mathcal{T}_{n_t} \times \mathcal{P}$, where $P_{\pi}^{(n_t)}$ denotes the set of permutations of length n_t. $\mathcal{T}_j$ lists the levels/categories of the j^{th} variable/treatment, i.e., $\forall j = 1, \ldots n_t \; \tau_j \in \mathcal{T}_j$, with each variable having an independent cardinality $|\mathcal{T}_j|$. The primary goal was to maintain the fresh product's qualities as recorded at the

harvest point in time (i.e., day 0), denoted as state i, when compared to their assay after a predefined period (e.g., 4 weeks), denoted as state f. Upon quantification of the assays, the desired outcome is the minimal deviation of all quantities, $\mathcal{L}_{i\to f}\left(\vec{\tau}\right)$. In particular, the aggregated scores of the fruit's color, stiffness, mass and aesthetics. The setup dictates the usage of only 2 treatments per protocol, so the targeted discovery/optimization problem may be formulated as follows – given a combinatorial search-space of possible postharvest treatments and packages $\mathcal{T}$, obtain a protocol $\vec{\tau}^*$ of 2 treatments and a package that minimizes the following loss function as long as the product is not rotten (there is no observed decay):[2]

$$\begin{array}{ll} \text{minimize}_{\vec{\tau}\in\pi\circ\mathcal{T}} & \mathcal{L}_{i\to f}\left(\vec{\tau}\right) \\ \text{subject to:} & \#\left\{j:\tau_j\neq 0,\ j=1,\ldots,n_t\right\}==2, \\ & \left\{\text{decay}<\epsilon\right\}. \end{array} \tag{1}$$

The results reported in [8] indicate a breakthrough in the postharvest quality thanks to this approach.

Aftermath

Evidently, the AI system obtained exciting Postharvest results by discovering a diverse set of protocols, which outperform the best known practices, including some protocols with a surprising nature (that will necessitate fundamental Postharvest research for obtaining mechanistic understanding). Beyond the fine Postharvest accomplishment and its impact on Global Food Security, there exists a broader picture of AI systems that contribute to scientific research. This reported discovery constitutes another indication that algorithmically-guided solutions have the potential to break new grounds, especially when the machine is not injected with expertise and knowledge. In other words, conception-free discoveries are attainable by machines, which lack fixation on standard practice and line-of-thinking, and rather freely navigate in the vast search-space of possibilities.

This was also the case for the quantum experiments – which are known to be counter-intuitive – as well as for the synthetic chemistry domain, whose discoveries are guided by human knowledge and conceptions. Overall, the reported use-cases are indicative that distancing from prior knowledge and/or flawed intuition is often an advantage that may enable discoveries.

3 Next Step: Scientific Hypotheses Formulation

Formal representations of knowledge have long been pursued by several disciplines, e.g., Philosophy and Logic. Today, in the digital era, there are well-established knowledge representation frameworks [17], of which we mention the following two:

1. Ontologies or Knowledge Graphs [30, 31] – information objects that capture properties, data, relations and taxonomy.
2. Logic programming – in particular, the `Prolog` programming language.

Therefore, within the context of the current paper, all the existing knowledge within a scientific domain is encodable, and particularly within **the Physical Sciences, their Postulates, Laws of Nature, hypotheses and theorems are all well-defined and encodable**.

Knowledge Representation In Practice Representations of knowledge are commonly encoded into machines nowadays, having the so-called Google Knowledge Graph as the most renowned. However, formal knowledge of scientific research has also been encoded throughout the years. For instance, a mouse embryo anatomy has been encoded in `Prolog` already in 2004 [32, 33]. In principle, every domain's knowledge may be encoded by enumerating its scientific corpus.

[2]The overall search-space cardinality is reduced here to $10\cdot 9\cdot 3\cdot\left|\mathcal{T}\right|_{1:n_t}\cdot\left|\mathcal{T}\right|_{2:n_t}\approx 10^6$, wherein the variables' cardinalities are sorted in a descending order. The cardinality is reduced here from $\sim 10^{17}$ in the generalized problem,

$$\left|\pi\circ\mathcal{T}\right|=n_t!\cdot\left[\prod_{j=1}^{n_t}\left(\mathcal{T}_j^{\max}-\mathcal{T}_j^{\min}\right)\right]\cdot\left|\mathcal{P}\right|.$$

Ontologies-Driven: BSc-Level Machine AI originally started with so-called *experts' systems*, that is, by "injecting" expertise and knowledge into the machine. However, the usage of those systems was not concerned with further activiely pursuing scientific research. Here, we assume that the existing knowledge of a certain scientific field is entirely encoded (say, by means of ontologies), and we argue that the process of teaching the machine to propose novel yet meaningful hypotheses is attainable. A possible learning scheme could be simple form of Reinforcement Learning [34], having a non-trivial hypothesis being defined as a *goal state*. The outcome/feedback will be provided by the human scientist (i.e., the drawn hypothesis is valid/novel/meaningful).

Data-Driven: PhD-Level Machine Inductive learning has become the dominant concept in AI-related tasks, that is, inferring from (possibly big) data [35]. Having machines obtaining *causal inference* concerning scientific observations and propose hypotheses would be the outcome of this direction. Such computational capabilities have already been demonstrated a decade ago [36], and are most likely improved nowadays by far. Meeting-in-the-middle with a hybrid approach (Data plus Ontologies) is probably the royal road, with a long-term goal to accomplish **deductive learning**. To the best of our understanding, this is what Kitano envisioned when he published his Nobel Turing Challenge [1].

Proposed Pathway: Obtaining an AI Research Agent in Chemistry

We outline the necessary steps to obtain an AI system capable of serving as an assistant (at the first-year graduate student level, i.e., "BSc-Level Machine") in an experimental Chemistry research group:

(1) Represent the undergraduate knowledge base in General Chemistry by selecting, say, 10 textbooks; encode them by means of a Knowledge Graph denoted as $\mathcal{G}$. The books' selection is assumed to have a good coverage of the knowledge space at the undergraduate level.

(2) Given $\mathcal{G}$, train and test an agent $\mathcal{A}$ to be able to propose novel hypotheses with respect to this graph. A valid hypothesis is to be defined as a *goal state*, and the information injection may partially adhere to the classical *declarative paradigm* [37]. During this learning process, which may be facilitated by a Deep Symbolic Reinforcement Learning algorithm [18], $\mathcal{G}$ will be enriched and updated with validated hypotheses (that is, knowledge that had not been originally encoded in $\mathcal{G}$, but has been suggested by $\mathcal{A}$ and then validated during this process).

(3) Select a physical system $\mathcal{S}$, whose controls and observables are accurately mapped. For instance, consider the space of organic syntheses.

(4) Let $\mathcal{A}$ propose a scientific hypothesis over $\mathcal{S}$. Once validated by the human Principal Investigator, reinforce $\mathcal{A}$ and let it translate the hypothesis into a CO problem (that is, map the hypothesis onto a physical observable and an array of decision variables).

(5) Let $\mathcal{A}$ treat the formulated CO as an experimental optimization challenge, say, by an evolutionary search heuristic. If convergence is obtained, it should report the solutions to the Principal Investigator. Otherwise, proceed with the optimization problem-solving subject to the experimental budget and constraints.

4 Discussion

We argue that the proposed pathway to construct an AI research agent constitutes a viable roadmap, although it may require an enormous effort. We will discuss in what follows the feasibility of realizing this ambition, and then elaborate on two additional aspects of the proposed framework.

Feasibility The effort in obtaining step (1) in the aforementioned pathway is expected to be labor-intensive and likely to demand tremendous resources. Currently, the task of encoding several scientific textbooks into the desired formal representation requires large teams of human domain experts who are also trained in knowledge representation. In some sense, the scale of effort may be comparable to attaining solid Large Language Models (see, e.g., [38]), a dream that is being brought to fruition only with the involvement of the resourceful Big Tech companies. In the future, such models may be able to facilitate knowledge representation in an automated manner.

Explainability Throughout the AI revolution, humans have considered the reasoning behind the output of AI systems as a prerequisite to gain trust in their operation. Indeed, the capacity to interpret the outcome of AI's decisions has been successfully addressed by the forming sub-field of Explainable AI (XAI; see, e.g., [39, 40, 41]). Nevertheless, the AI decisions in the proposed framework herein constitute scientific hypotheses that do not necessarily require explanations. Equivalently, when scientists formulate hypotheses, they must not always express why they reached them. For instance, a hypothesis stating that a "certain catalysis could be boosted if the combination of substances and operations be searched over" needs no further explanation (other than the trivial reasoning that previous observations in Organic Chemistry indicated that Combinatorial Optimization of catalyses' protocols resulted in improved yields). At the same time, other types of hypotheses may benefit from interpretation (e.g., their context within an ongoing investigation of a natural phenomenon). Toward this end, XAI may be developed to facilitate it and generate appropriate explanations. Recently, the notion of Scientific Understanding has been put forward by Krenn and co-workers [42].

Creativity and Mindset The creativity of AI-generated output had been questioned already in earlier stages of AI research and development in the 20th century [43], when it was recognized that AI can demonstrate creativity in several ways: by producing novel combinations of familiar ideas, exploring the potential of conceptual spaces, and making transformations that enable the generation of previously impossible ideas. The question has been put forward with much emphasis during the current revolution [44], especially upon the release of art generating models [45], and qualms concerning emotions and inspiration.

While there is no clear consensus on how to rigorously define creativity (especially versus novelty), the AI output in the current context need not be creative in the artistic sense, rather ground-breaking and conception-free in the scientific sense. In other words, the merit of the AI-based scientific work will be evaluated in light of laying novel hypotheses that were not considered heretofore, in dropping widely used conceptual directions, and even in shifting paradigms. In that sense, the contribution of an Artificial Scientist may not underlie the inspirational aspect of the ideas, but rather in their capacity to enhance progress. The potential in accomplishing this goal is enormous and primarily rooted in the aforementioned factors (being conception-free and unbiased by flawless intuition, as demonstrated in some of the case-studies in Section 2), but also in lacking mental/cognitive fear and shame, unlike the typical human mindset, which consequently avoids failures and generally makes risk-averse decisions. Finally, when re-considering the "*discoveries as Combinatorial Optimization problem-solving*" perspective, it is mathematically understood that escaping local optima/traps is necessary even at the cost of consecutive (temporary) failures. Yet, such maneuvers are not likely to be made by human scientists, since frequent failures are usually not tolerated by the human mindset. Therefore, being fearless/shameless, so to say, may constitute an advantage of the AI system when freely navigating the broad search-space of possible discoveries.

Acknowledgements

We thank Herschel Rabitz and Dror Noy for the insightful discussions.

References

[1] H. Kitano, "Nobel Turing Challenge: creating the engine for scientific discovery," *npj Systems Biology and Applications*, vol. 7, p. 29, 2021.

[2] D. B. Kell, "Scientific discovery as a combinatorial optimisation problem: How best to navigate the landscape of possible experiments?," *BioEssays*, vol. 34, no. 3, pp. 236–244, 2012.

[3] H. Rabitz, R.-B. Wu, T.-S. Ho, K. M. Tibbetts, and X. Feng, *Fundamental Principles of Control Landscapes with Applications to Quantum Mechanics, Chemistry and Evolution*, pp. 33–70. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014.

[4] O. M. Shir and T. Bäck, "Sequential experimentation by evolutionary algorithms," in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, GECCO '22, (New York, NY, USA), p. 1450–1468, Association for Computing Machinery, 2022.

[5] D. Calzolari, S. Bruschi, L. Coquin, J. Schofield, J. D. Feala, J. C. Reed, A. D. McCulloch, and G. Paternostro, "Search algorithms as a framework for the optimization of drug combinations," *PLoS Computational Biology*, vol. 4, p. e1000249, Dec. 2008.

[6] D. Weuster-Botz and C. Wandrey, "Medium optimization by genetic algorithm for continuous production of formate dehydrogenase," *Process Biochemistry*, vol. 30, no. 6, pp. 563–571, 1995.

[7] C. Erlich, "Experimental combinatorial optimization of phycobiliproteins' expression in E. Coli," Master's thesis, Tel-Hai College, 2019.

[8] O. M. Shir, B. Yazmir, A. Israeli, and D. Gamrasni, "Algorithmically-guided postharvest protocols by experimental combinatorial optimization," in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, GECCO '22, (New York, NY, USA), p. 2027–2035, Association for Computing Machinery, 2022.

[9] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.

[10] C. Deng, X. Ji, C. Rainey, J. Zhang, and W. Lu, "Integrating machine learning with human knowledge," *Iscience*, vol. 23, no. 11, p. 101656, 2020.

[11] S. H. M. Mehr, M. Craven, A. I. Leonov, G. Keenan, and L. Cronin, "A universal system for digitization and automatic execution of the chemical synthesis literature," *Science*, vol. 370, no. 6512, pp. 101–108, 2020.

[12] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, pp. 583–589, 2021.

[13] A. H.-W. Yeh, C. Norn, Y. Kipnis, D. Tischer, S. J. Pellock, D. Evans, P. Ma, G. R. Lee, J. Z. Zhang, I. Anishchenko, B. Coventry, L. Cao, J. Dauparas, S. Halabiya, M. DeWitt, L. Carter, K. N. Houk, and D. Baker, "De novo design of luciferases using deep learning," *Nature*, vol. 614, pp. 774–780, 2023.

[14] B. Goertzel, "Artificial general intelligence: Concept, state of the art, and future prospects," *Journal of Artificial General Intelligence*, vol. 5, no. 1, pp. 1–48, 2014.

[15] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman, "Building machines that learn and think like people," *Behavioral and brain sciences.*, vol. 40, 2017.

[16] B. J. Lansdell and K. P. Kording, "Towards learning-to-learn," *Current Opinion in Behavioral Sciences*, vol. 29, pp. 45–50, 2019. Artificial Intelligence.

[17] A. Hunter and W. Liu, "A survey of formalisms for representing and reasoning with scientific knowledge," *Knowl. Eng. Rev.*, vol. 25, no. 2, pp. 199–222, 2010.

[18] M. Garnelo, K. Arulkumaran, and M. Shanahan, "Towards deep symbolic reinforcement learning," 2016.

[19] M. Garnelo and M. Shanahan, "Reconciling deep learning with symbolic artificial intelligence: representing objects and relations," *Current Opinion in Behavioral Sciences*, vol. 29, pp. 17–23, 2019.

[20] J. Roslund, O. M. Shir, T. Bäck, and H. Rabitz, "Accelerated Optimization and Automated Discovery with Covariance Matrix Adaptation for Experimental Quantum Control," *Physical Review A (Atomic, Molecular, and Optical Physics)*, vol. 80, no. 4, p. 043415, 2009.

[21] A. E. Eiben, S. Kernbach, and E. Haasdijk, "Embodied artificial evolution: Artificial evolutionary systems in the 21st century," *Evolutionary Intelligence*, vol. 5, pp. 261–272, 2012.

[22] A. E. Eiben and J. Smith, "From evolutionary computation to the evolution of things," *Nature*, vol. 521, pp. 476–482, 2015.

[23] G. E. Box, J. S. Hunter, and W. G. Hunter, *Statistics for Experimenters: Design, Innovation and Discovery*. Hoboken, NJ, USA: John Wiley and Sons, second ed., 2005.

[24] H.-P. Schwefel, *Evolution and Optimum Seeking*. New York, NY, USA: John Wiley & Sons, Inc., 1995.

[25] A. Deshwal, S. Belakaria, and J. R. Doppa, "Mercer features for efficient combinatorial bayesian optimization," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, pp. 7210–7218, May 2021.

[26] A. Vasylenko, J. Gamon, B. B. Duff, V. V. Gusev, L. M. Daniels, M. Zanella, J. F. Shin, P. M. Sharp, A. Morscher, R. Chen, A. R. Neale, L. J. Hardwick, J. B. Claridge, F. Blanc, M. W. Gaultois, M. S. Dyer, and M. J. Rosseinsky, "Element selection for crystalline inorganic solid discovery guided by unsupervised machine learning of experimentally explored chemistry," *Nature Communications*, vol. 12, 2021.

[27] M. Krenn, M. Malik, R. Fickler, R. Lapkiewicz, and A. Zeilinger, "Automated search for new quantum experiments," *Phys. Rev. Lett.*, vol. 116, p. 090405, Mar 2016.

[28] A. A. Melnikov, H. P. Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger, and H. J. Briegel, "Active learning machine learns to create new quantum experiments," *Proceedings of the National Academy of Sciences*, vol. 115, no. 6, pp. 1221–1226, 2018.

[29] P. V. Mahajan, O. J. Caleb, Z. Singh, C. B. Watkins, and M. Geyer, "Postharvest treatments of fresh produce," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 372, no. 2017, p. 20130309, 2014.

[30] H. Paulheim and P. Cimiano, "Knowledge graph refinement: A survey of approaches and evaluation methods," *Semantic Web*, vol. 8, no. 3, pp. 489–508, 2016-12-06.

[31] A. Hogan, E. Blomqvist, M. Cochez, C. D'amato, G. D. Melo, C. Gutierrez, S. Kirrane, J. E. L. Gayo, R. Navigli, S. Neumaier, A.-C. N. Ngomo, A. Polleres, S. M. Rashid, A. Rula, L. Schmelzeisen, J. Sequeda, S. Staab, and A. Zimmermann, "Knowledge graphs," *ACM Comput. Surv.*, vol. 54, jul 2021.

[32] A. Burger, D. Davidson, and R. Baldock, "Formalization of mouse embryo anatomy," *Bioinformatics*, vol. 20, 2004.

[33] A. Burger, D. Davidson, and R. Baldock, *Anatomy Ontologies for Bioinformatics: Principles and Practice*. Computational Biology, Springer London, 2007.

[34] R. S. Sutton and A. G. Barto, *Reinforcement Learning, second edition: An Introduction*. Adaptive Computation and Machine Learning series, MIT Press, second ed., 2018.

[35] J. Pearl, "The seven tools of causal inference, with reflections on machine learning," *Commun. ACM*, vol. 62, p. 54–60, feb 2019.

[36] M. Schmidt and H. Lipson, "Distilling free-form natural laws from experimental data," *Science*, vol. 324, no. 5923, pp. 81–85, 2009.

[37] M. Flasiński, *Symbolic Artificial Intelligence*, pp. 15–22. Cham: Springer International Publishing, 2016.

[38] E. A. M. van Dis, J. Bollen, W. Zuidema, R. van Rooij, and C. L. Bockting, "ChatGPT: five priorities for research," *Nature*, vol. 614, pp. 224–226, 2023.

[39] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai," *Information Fusion*, vol. 58, pp. 82–115, 2020.

[40] P. Gohel, P. Singh, and M. Mohanty, "Explainable AI: current status and future directions," 2021.

[41] A. Holzinger, A. Saranti, C. Molnar, P. Biecek, and W. Samek, *Explainable AI Methods - A Brief Overview*, pp. 13–38. Cham: Springer International Publishing, 2022.

[42] M. Krenn, R. Pollice, S. Y. Guo, M. Aldeghi, A. Cervera-Lierta, P. Friederich, G. dos Passos Gomes, F. Haese, A. Jinich, A. Nigam, Z. Yao, and A. Aspuru-Guzik, "On scientific understanding with artificial intelligence," *Nature Reviews Physics*, vol. 4, pp. 761—-769, 2022.

[43] M. A. Boden, "Creativity and artificial intelligence," *Artificial Intelligence*, vol. 103, no. 1, pp. 347–356, 1998. Artificial Intelligence 40 years later.

[44] K. Kirkpatrick, "Can AI Demonstrate Creativity?," *Commun. ACM*, vol. 66, p. 21–23, jan 2023.

[45] G. Marcus, E. Davis, and S. Aaronson, "A very preliminary analysis of DALL-E 2," 2022.