Artificial General Intelligence: A New Perspective, with Application to Scientific Discovery

A. M. Khalili[1*]
[1] Staffordshire University, United Kingdom.
*Correspondence: a.m.khalili@outlook.com

Abstract The dream of building machines that have human-level intelligence has inspired scientists for decades. Remarkable advances have been made recently; however, we are still far from achieving this goal. In this paper, I propose an alternative perspective on how these machines might be built focusing on the scientific discovery process which represents one of our highest abilities that requires a high level of reasoning and remarkable problem-solving ability. By trying to replicate the procedures followed by many scientists, the basic idea of the proposed approach is to use a set of principles to solve problems and discover new knowledge. These principles are extracted from different historical examples of scientific discoveries. Building machines that fully incorporate these principles in an automated way might open the doors for many advancements.

Keywords *AI, AGI, Artificial Intelligence, Artificial General Intelligence, Human-level Machine Intelligence, Scientific Discovery, Science Automation.*

Introduction

In [1-2] Penrose talked about the existence of three different worlds: the mental world, the physical world, and the mathematical world. The physical world is governed by laws that reside in the world of mathematics, our minds emerge from the physical world, and those minds are able to access the mathematical world by discovering mathematics, which is within the scope of reason.

Plato believed that ideas or forms exist in some ideal world outside the physical world, which became later known as the 'Platonic world of forms' [3]. If Plato's realm exists, it is very unlikely that different parts of such realm are disconnected and do not have links with each other, they would be beautifully connected and one can navigate between different parts of that realm, and discover new hidden structures. In the mathematical world, Bourbaki [5] likened mathematics to a city, where the outlying districts and suburbs expand in a chaotic manner on the surrounding country, while the center of the city is rebuilt from time to time, and each time in accordance with a more clear plan and more majestic order. Langland program is a recent attempt to build connections between different parts of the mathematical world.

Although on rare occasions, the intellect might break through into those worlds and get a limited glimpse of those realms as described by Penrose [1], and illustrated through many examples by Hadamard [6], still in most times we follow certain procedures and principles to reconstruct those realms. Similar to the mathematical and physical worlds, a curtail aspect of the mental world and hence Artificial General Intelligence (AGI) is to build the maps that represent other realms by using a set of principles to reconstruct these original worlds and discover new knowledge. The landscape of AGI is extremely vast, in this paper I will focus on the scientific discovery process [6-8, 33-37, 41-42, 54-55, 77, 81, 85, 87-91], see [91] for a recent survey on different scientific discovery systems.

Related Work

In a recent survey [9] on when human-level machine intelligence will be achieved, the view of over 50% of a number of experts was that it would be around by 2040-2050 and over 90% said by 2075. The last few years have seen exceptional progress, much of this progress has come from recent advances in deep learning. Deep learning has achieved remarkable results in many domains such as image classification, speech recognition, and gaming. Despite the success of many AI systems, these systems suffer from many limitations [80, 83, 86] making them unlikely candidates to achieve the goal of AGI.

Domingo [10] has beautifully summarized the main used approaches in machine learning. Symbolists use logic and related representations to capture knowledge and relationships about the world. Connectionists take inspiration from neuroscience and seek to reverse engineer the brain. Evolutionaries take inspiration from genetics and evolutionary biology and seek to simulate evolution. Bayesians take inspiration from statistics, they use probabilistic representations to capture uncertainty. Finally, analogizers take inspiration from psychology and learn by finding similarities. Domingo then argued for the need for a master algorithm that combines key features from all existing approaches to achieve the goal of AGI.

Connectionists approaches such as deep learning are very effective in pattern recognition, but they still have many limitations in high-level functions such as reasoning. Furthermore, they do not have the flexibility to generalize to new tasks, they are also vulnerable to adversarial samples [11-13, 80, 83, 86]. Yuan et al. [12] proposed an architecture that can produce images correctly classified by human subjects but misclassified by a deep network with a 97% adversarial success rate by only changing 4 % of the image on average. Su et al. [13] showed that modifying one pixel only could lead up to 73% adversarial success rate depending on the used images. Recently, there is a growing interest in building neural networks that can learn to reason [76-79]. Saxon et al. [77] demonstrated that current state-of-the-art neural networks show moderate performance in solving basic mathematical problems, the performance deteriorates for questions that require the computation of intermediate values. The model was able to solve only 14/40 questions from maths exams for 16 year old schoolchildren in the UK. In [78-79] the researchers tested neural networks ability in structural, relational, and analogical reasoning by trying to solve IQ-like visual questions. In particular, they tested the models on the Raven's Progressive Matrices (RPM) dataset, which is correlated with many aspects of reasoning. The results show that there is still a clear gap between machine algorithms and humans even when the machines have intensive training. The strongest objection against the connectionists approach is that the brain is unlikely to achieve general intelligence via self-organizing networks of neurons, what is important is the software or the information processing architecture, not the low-level hardware by which the architecture is implemented.

Logic-based approaches have remarkable representational power, logic is also crucial to achieve high-level functions such as reasoning. However, these approaches tend to be limited in learning and creativity, they are also limited in handling noise and uncertainty present in many applications. To handle the uncertainty and complexity present in many real-world problems, Domingo et al. [14] proposed an approach that combines both logical and statistical AI, by combining first-order logic and graphical models. Where many applications require the robustness of probability and the expressiveness of first-order logic. Neural-symbolic computation is an attempt to provide a coherent unification of symbolic AI and connectionism, it aims to integrate the power of both neural networks and reasoning. Neural networks can be used to provide robust learning, while logic can provide the necessary explanation. See [15, 92] for recent advances and different attempts to combine symbolic AI and neural networks.

Recently there is a growing interest in studying human-level artificial general intelligence, Adams et al. [16] presented a broad outline of a roadmap toward AGI. Seven scenarios were presented as milestones to AGI along with many directions for future research. Goertzel [17] presented a survey on recent progress toward AGI where different approaches to AGI were reviewed. Different metrics for general intelligence were evaluated, the conclusion was that assessing partial progress is more controversial compared to assessing the achievement of human-level intelligence which is more straightforward. Clune and othrs [85] argued for an evolution-inspired path to AGI. The basic idea is to create an AI-generating algorithm, which can automatically learn how to produce general AI. This approach could also shed some light on the origin of our own intelligence.

Lake et al. [18] reviewed recent progress in cognitive science, they suggested that human-like thinking and learning machines have to go beyond current trends in both how they learn and what they learn. They argued that these machines should (a) be able to build causal models of the world to support understanding and explanation, rather than only solving pattern recognition problems. (b) Have ground learning in theories of psychology and physics. (c) Incorporate the learning-to-learn approach to acquire and generalize knowledge to new situations and tasks. They proposed some challenges and paths towards these goals such as integrating the power of recent advances in deep learning with more structured cognitive models.

Proposed Framework

Many researchers [19-21] believe that our abilities to construct concepts, act as a basic building block of understanding and reasoning. In neuroscience, there are many recent suggestions that map-like representations may be a mechanism capable of organizing knowledge of all kinds [22-23].

The notion of a centralized control system that guides thought (a Central Executive) is common in information processing theories of cognition [24-26]. The Central Executive is hypothesized to direct planning, perceptual, sensory, and motor systems. The Central Executive would be involved in exploration and search for different strategies to achieve a goal.

The Operating System (OS) in computer systems plays an important role in storing and retrieving information. Another important role of the OS is its ability to execute external programs, facilitates their functionality, and gives them access to different resources such as memory, processing units, etc. Similar to the role of the OS, in AGI there is a crucial need to an algorithm that has minimum functionalities such as information storing and retrieval. The algorithm should be able to build concept maps that link different concepts together, these concept maps represent existing knowledge. It should be also able to perform store, retrieve, and search operations on the maps. The other important role is to apply a set of principles or programs (in a similar way the OS executes external programs) to solve problems and discover new knowledge. These principles should seek to expand the maps by discovering new knowledge, they should also reveal new connections that link different concepts in the maps. In addition to logic, which plays an important role in the thinking process, in reality logic alone is not enough, we usually use more sophisticated principles and structures and apply them in the same way logic operates on concept maps. In the literature, there is a focus on two main principles, concepts combination, and analogies. However, other principles should be taken into account to build a comprehensive framework. These principles can be summarized as follow

1. Mathematization: The ability of mathematics in describing the natural world [30-32] never ceases to amaze scientists. Mathematics today is very effective in studying fields as diverse as physics, computer science, finance, and biology. Mathematics is not only able to describe the natural world, but this description on many occasions led us to predict and discover new aspects of the studied phenomena. On many occasions, testing the mathematical description in new extreme conditions led to new insights and sometimes to new theories. The other remarkable thing is that on many occasions, the mathematics we use to describe the natural world is already discovered by mathematician tens to hundreds years earlier. For instance, using the imaginary unit i and Euler's number e has helped in describing the wave equation, complex numbers were also crucial in describing quantum physics. In 1915 for instance, General Relativity (GR) was at the frontier of the map of physics, many physicists used the mathematization principle to derive new knowledge from the GR equation, they were able to predict gravitational waves, and black holes as solutions to the GR equation, both of these phenomena were confirmed experimentally in the few recent years.

   In AI, there are many attempts to build symbolic regression algorithms, which are automated tools to find the mathematical equation that fits the experimental data [33]. Udrescu and Tegmark [34] developed an algorithm that combines neural network fitting with a set of physics-inspired techniques. They applied it to 100 equations from the Feynman lectures on physics. It was able to discover all of them, the state of the art algorithm was only able to discover 71. For a more difficult test set, the state of the art success rate was improved from 15% to 90%. Many researchers recently [35-37] started to use recent advances in deep learning such as generative adversarial networks to discover physical concepts from experimental data without being provided with any additional prior knowledge and then use the discovered representation to answer questions about the physical system. The main purpose of the algorithm that encapsulates the mathematization principle would be to find the equations that describe the experimental data.

2. Optimization: optimization is one of the most powerful principles, it is one of the most used principles in everyday life, we constantly try to minimize energy, cost, distance, time, etc. Some notable uses of this principle in science include minimizing the energy and time that are required to distribute fuels to the cells, gives rise to the circulatory system networks [27]. Optimizing the balance between the input and output energy gives rise to bird migration patterns [28]. Increasing entropy derives matter to acquire lifelike physical properties [29]. The main purpose of the algorithm that encapsulates the optimization principle would be to find the optimization criteria and constraints that describe the studied problem.

3. Analogies: many prominent cognitive scientists [38] consider analogy to be one of the main building blocks of human cognition. There are many examples where analogy has played a crucial role in discovering new scientific concepts. Polya [39] observed that analogy has a share in all mathematical discoveries. He provided many historical examples where analogy played the main role. See [40] for a long list of the use of analogy in scientific discovery. Nersessian [41-42] also gave a list of examples such as Newton's analogy between projectiles and the moon which gave rise to universal gravitation, Darwin's analogy between selective breeding and reproduction in nature which gave rise to natural selection, and the Rutherford-Bohr analogy between the

structure of the solar system and the configuration of subatomic particles. Many algorithms in computer science have been inspired from biology to solve different problems such as the traveling salesman problem [43], [44]. They took inspirations from ants, which are capable of finding the shortest path from the nest to a food source [45], [46], by using a chemical substance called pheromone. Other notable examples include genetic algorithms, see [47-49] for a list of bio-inspired algorithms. See [93-96] for different theoretical and computational frameworks for the analogy principle. The main purpose of the algorithm that encapsulates the analogy principle would be to find matching between the studied problem and similar problems.

4. Concepts Combination: this is a fundamental cognitive principle underlying much of our thinking [50]. Creativity results from a combination of different ideas, has been proposed by many researchers [51-52]. Many scientists such as Einstein and Poincare described their insights to be the result of concepts combination [53]. Many scientific discoveries could be understood as instances of conceptual combination, where new concepts arise by combining old ones [54-55]. One famous example is the wave theory of sound, which required the development of a new concept of a sound wave. The concepts of wave and sound are part of known phenomena. The ancient Greek Chrysippus combines them together to create the new concept of a sound wave that can explain many characteristics of sound such as reflection and propagation. Concepts combination is also one of the main used themes in theoretical physics. In 1973 for instance, both general relativity and quantum mechanics were at the frontier of the map of physics, by combining ideas from these two fields, Hawking proposed that black holes emit thermal radiation. Moreover, by combining ideas from quantum mechanics and statistical mechanics, Bekenstein and Hawking proposed the formula that describes the black hole entropy, which later led to the holographic principle. Martinez et al. [56] used a theory-based algorithmic blending of mathematical concepts as a basis for concept invention. A related principle is concepts blending (see [57] for different theoretical and computational frameworks).

5. Emergence: emergence is a powerful approach to explain complex behaviors by simple underlying rules. One notable example is birds flocking, some birds fly in coordinated flocks that show remarkable synchronization in movements. Heppner [60] showed that the coordinated movements could be the result of simple movement rules followed by each bird individually. Another example is the Game of Life [61], a two-dimensional cellular automaton with rules that avoid the formation of structures that grow freely or quickly disappear. Remarkable behaviors have been observed such as the glider, a small group of cells that moves like an independent emergent entity. Wolfram [62] used a cellular automaton with simple initial conditions and simple rules to produce highly complex behaviors. The main purpose of the algorithm that encapsulates the emergence principle would be to find the set of rules that gives rise to the emergent behavior.

6. Computability: computation is a new paradigm that has revolutionized science and engineering [63, 82], it has derived many advancements in science and changed the way it is done. Many biologists would agree that biology is an information science. One of the most notable examples is the DNA, which gives rise to the whole biological system. A growing number of physicists would also agree that the interactions between physical systems are information processing, and the universe is nothing but a giant computation [64-65]. Zenil et al. [81] proposed a universal unsupervised and parameter-free model-oriented approach based on the concept of algorithmic

probability to decompose an observation into its most likely algorithmic generative models. The approach uses a perturbation-based causal calculus and principles drawn from algorithmic complexity to infer model representations. They demonstrated the ability of the approach to deconvolve interacting mechanisms regardless of whether the resulted objects are bit strings, images, or networks. The main purpose of the algorithm that encapsulates the computability principle would be to find the program that gives rise to the observed phenomenon.

7. Beauty: aesthetic judgments play a guiding role in scientific discovery [66-69]. Scientists often evaluate models and theories based on their aesthetic appeal. Some scientists have even suggested that the goal of science is to find beauty in nature. Herman Weyl famously said that he would always try to unite the true with the beautiful in his work, but would choose the beautiful if he has to choose between the two. Dirac argued scientists to strive mainly for mathematical beauty when they want to express the fundamental laws of nature in mathematical forms. He argued scientists to have confidence in a beautiful theory independently of its empirical adequacy [70].

   The role of beauty in science has found some skepticism because we still do not have a satisfactory theory that can exactly test the claims made by scientists about the beauty of a theory [71]. A recent interesting study about the nature of aesthetic in science by Zeki et al. [72] demonstrated that the aesthetic appreciation of mathematical equations corresponds to the same brain activity that corresponds to the appreciation of music and art. Dirac argued that while aesthetic appreciation of art might be subjective, beauty in mathematics is objective and universal [70]. Zee [73] and Thuan [74] also argued that beauty's attributes such as simplicity, symmetry, and elegance have universal values and that they should not be subject to revision in science.

8. Universality: universality means that a similar mathematical formulation can describe different phenomena across multiple fields. The spectral measurements of composite materials, such as sea ice and human bones, the time between the buses' arrival in the city of Cuernavaca in Mexico, the zeros of the Riemann zeta function, and many other phenomena have shown to have the same statistical distribution [58]. Power laws are another example of universal laws that have been observed in a wide range of phenomena in fields as diverse as physics, biology, and computer science [59].

9. Unification: unification has played a key role in physics since Newton who unified celestial and terrestrial mechanics, Maxwell who unified electricity and magnetism, then the unification of the weak and the electromagnetic forces, and most recently the attempts to unify all the four fundamental forces. Unification has also played an important role in biology.

10. Symmetry: symmetry has played an important role in physics [75] from Newton's laws to Maxwell's equations, and general relativity. Symmetry has also played a fundamental role in the development of quantum mechanics. Today, it is one of the most used principles in searching for the fundamental laws of physics and further unification.

There are many other domain-specific principles that are specific to certain fields. Finding new principles could be crucial to make new discoveries and revolutionize our understanding, for instance, the use of the symmetry and mathematization principles has revolutionized modern physics, and maybe we need new principles to see in new perspectives and solve current challenges.

The use of the above principles varies from field to field, some principles are still not used, and others are used in a limited fashion. Incorporating these principles fully in an automated scientific discovery framework might open the doors for many advancements, for instant, using the computability principle is still very limited in physics, the use of mathematization principle is still very limited in social sciences, and the use of the beauty principle is more dominant in physics and mathematics than in biology.

The main challenge of the proposed framework is to build the algorithms that encapsulate the set of principles. Deep learning could be a very effective tool to implement some of these principles, it has shown promising results for the mathematization principle. However, it might be limited for other principles.

Discussion and Conclusion

Despite the recent success of deep learning, its many limitations makes it unlikely candidates to achieve the goal of AGI. Combining symbolic AI with deep learning is one of the most promising directions in machine learning, it could improve the reasoning ability of these systems. However, this is not enough, more sophisticated principles and high level structures should be incorporated. This paper has proposed a general framework that encapsulates different principles used in science, these principles are extracted from different historical examples of how different scientists made their discoveries. The proposed approach might help in giving an alternative perspective to the artificial general intelligence problem by investigating the scientific discovery process, which requires a high ability in reasoning and problem-solving. Building machines that fully incorporate these principles in an automated way might also open the doors for many advancements.

References

1. Penrose, R., & Mermin, N. D. (1990). The emperor's new mind: Concerning computers, minds, and the laws of physics.
2. Penrose, R. (2004). The road to reality: A complete guide to the physical universe. Jonathan Cape.
3. Ross, D. (1953). Plato's theory of ideas.
4. Aigner, M., Ziegler, G. M., Hofmann, K. H., & Erdos, P. (2010). Proofs from the Book (Vol. 274). Berlin: Springer.
5. Bourbaki, N. (1950). The architecture of mathematics. The American Mathematical Monthly, 57(4), 221-232.
6. Hadamard, J. (1996). The mathematician's mind: The psychology of invention in the mathematical field.
7. Dunbar, K., & Fugelsang, J. (2005). Scientific thinking and reasoning. The Cambridge handbook of thinking and reasoning, 705-725.
8. Sparkes, A., Aubrey, W., Byrne, E., Clare, A., Khan, M. N., Liakata, M., ... & Young, M. (2010). Towards Robot Scientists for autonomous scientific discovery. Automated Experimentation, 2(1), 1.
9. Müller, V. C., & Bostrom, N. (2016). Future progress in artificial intelligence: A survey of expert opinion. In Fundamental issues of artificial intelligence (pp. 555-572). Springer, Cham.
10. Domingos, P. (2015). The master algorithm: How the quest for the ultimate learning machine will remake our world. Basic Books.

11. Papernot, N., McDaniel, P., Jha, S., Fredrikson, M., Celik, Z. B., & Swami, A. (2016, March). The limitations of deep learning in adversarial settings. In 2016 IEEE European Symposium on Security and Privacy (EuroS&P) (pp. 372-387). IEEE.
12. Yuan, X., He, P., Zhu, Q., & Li, X. (2019). Adversarial examples: Attacks and defenses for deep learning. IEEE transactions on neural networks and learning systems.
13. Su, J., Vargas, D. V., & Sakurai, K. (2019). One pixel attack for fooling deep neural networks. IEEE Transactions on Evolutionary Computation.
14. Domingos, P., & Lowd, D. (2019). Unifying logical and statistical AI with Markov logic. Communications of the ACM, 62(7), 74-83.
15. Besold, T. R., Garcez, A. D. A., Bader, S., Bowman, H., Domingos, P., Hitzler, P., ... & de Penning, L. (2017). Neural-symbolic learning and reasoning: A survey and interpretation. arXiv preprint arXiv:1711.03902.
16. Adams, S., Arel, I., Bach, J., Coop, R., Furlan, R., Goertzel, B., ... & Shapiro, S. C. (2012). Mapping the landscape of human-level artificial general intelligence. AI magazine, 33(1), 25-42.
17. Goertzel, B. (2014). Artificial general intelligence: concept, state of the art, and future prospects. Journal of Artificial General Intelligence, 5(1), 1-48.
18. Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. Behavioral and brain sciences, 40.
19. Mordatch, I. (2018). Concept learning with energy-based models. arXiv preprint arXiv:1811.02486.
20. Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. Cognitive science, 4(2), 195-208.
21. Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive psychology, 8(3), 382-439.
22. Behrens, T. E., Muller, T. H., Whittington, J. C., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. Neuron, 100(2), 490-509.
23. Hawkins, J., Lewis, M., Klukas, M., Purdy, S., & Ahmad, S. (2018). A framework for intelligence and cortical function based on grid cells in the neocortex. Frontiers in neural circuits, 12, 121.
24. Baddeley, A. (2000). The episodic buffer: a new component of working memory?. Trends in cognitive sciences, 4(11), 417–423.
25. Norman, D. A., & Shallice, T. (1986). Attention to action. In Davidson, R., Schwartz, G., & Shapiro, D. (Eds.), Consciousness and Self-Regulation, pp. 1–18. Springer.
26. Barnard, P. J. (1999). Interacting cognitive subsystems: Modeling working memory phenomena within a multiprocessor architecture. In Miyake, A., & Shah, P. (Eds.), Models of Working Memory, pp. 298–339. Cambridge University Press.
27. Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology. Ecology, 85(7), 1771-1789.
28. Somveille, M., Rodrigues, A. S., & Manica, A. (2018). Energy efficiency drives the global seasonal distribution of birds. Nature ecology & evolution, 2(6), 962.
29. England, J. L. (2013). Statistical physics of self-replication. The Journal of chemical physics, 139(12), 09B623_1.
30. Wigner, E. P. (1990). The unreasonable effectiveness of mathematics in the natural sciences. In Mathematics and Science (pp. 291-306).
31. Livio, M. (2010). Is God a mathematician?. Simon and Schuster.
32. Tegmark, M. (2014). Our mathematical universe: My quest for the ultimate nature of reality. Vintage.
33. Schmidt, M., & Lipson, H. (2009). Distilling free-form natural laws from experimental data. science, 324(5923), 81-85.

34. Udrescu, S. M., & Tegmark, M. (2019). AI Feynman: a Physics-Inspired Method for Symbolic Regression. arXiv preprint arXiv:1905.11481.
35. Zheng, D., Luo, V., Wu, J., & Tenenbaum, J. B. (2018). Unsupervised learning of latent physical properties using perception-prediction networks. arXiv preprint arXiv:1807.09244.
36. Iten, R., Metger, T., Wilming, H., Del Rio, L., & Renner, R. (2018). Discovering physical concepts with neural networks. arXiv preprint arXiv:1807.10300.
37. Schawinski, K., Turp, M. D., & Zhang, C. (2018). Exploring galaxy evolution with generative models. arXiv preprint arXiv:1812.01114.
38. Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. Cognitive psychology, 12(3), 306-355.
39. Polya, G. (1954). Mathematics and plausible reasoning volume 1, Induction and analogy in mathematics. Princeton University Press.
40. Holyoak, K. J., & Thagard, P. (1995). Mental leaps: Analogy in creative thought. Cambridge, MA: MIT Press⁄Bradford Books.
41. Nersessian, N. J. (2008). Creating Scientific Concepts. Cambridge, MA: MIT Press.
42. Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. Cognitive models of science, 15, 3-44.
43. Dorigo, M., & Birattari, M. (2010). Ant colony optimization (pp. 36-39). Springer US.
44. Colorni, A., Dorigo, M., & Maniezzo, V. (1992, December). Distributed optimization by ant colonies. In Proceedings of the first European conference on artificial life (Vol. 142, pp. 134-142).
45. Goss, S., Aron, S., Deneubourg, J. L., & Pasteels, J. M. (1989). Self-organized shortcuts in the Argentine ant. Naturwissenschaften, 76(12), 579-581.
46. Beckers, R., Deneubourg, J. L., & Gross, S. (1992). Trail and U-turns in the Selection of the Shortest Path by the Ants. J. of Theoretical Biology, 159, 397-415.
47. Binitha, S., & Sathya, S. S. (2012). A survey of bio inspired optimization algorithms. International journal of soft computing and engineering, 2(2), 137-151.
48. Yang, X. S. (2014). Nature-inspired optimization algorithms. Elsevier.
49. Olariu, S., & Zomaya, A. Y. (2005). Handbook of bioinspired algorithms and applications. Chapman and Hall/CRC.
50. Fauconnier, G. and Turner, M. (2003). The Way We Think.
51. Koestler, A. (1967). The act of creation. New York: Dell.
52. Boden, M. (2004). The creative mind: Myths and mechanisms (2nd ed.). London: Routledge.
53. Mednick, S. A. (1962). The associate basis of the creative process. Psychological Review, 69, 220–232.
54. Thagard, P. (1988). Computational philosophy of science. Cambridge, MA: MIT Press.
55. Thagard, P., & Stewart, T. C. (2011). The AHA! experience: Creativity through emergent binding in neural networks. Cognitive science, 35(1), 1-33.
56. Martinez, M., Krumnack, U., Smaill, A., Besold, T. R., Abdel-Fattah, A. M., Schmidt, M., Pease, A. (2014, December). Algorithmic aspects of theory blending. In International Conference on Artificial Intelligence and Symbolic Computation (pp. 180-192). Springer, Cham.
57. Confalonieri, R., Pease, A., Schorlemmer, M., Besold, T. R., Kutz, O., Maclean, E., & Kaliakatsos-Papakostas, M. (Eds.). (2018). Concept invention: Foundations, implementation, social aspects and applications. Springer.
58. Deift, P. (2006). Universality for mathematical and physical systems. arXiv preprint math-ph/0603038.
59. Newman, M. E. (2005). Power laws, Pareto distributions and Zipf's law. Contemporary physics, 46(5), 323-351.

60. Heppner, F., & Grenander, U. (1990). A stochastic nonlinear model for coordinated bird flocks. The ubiquity of chaos, 233, 238.
61. Conway, J. (1970). The game of life. Scientific American, 223(4), 4.
62. Wolfram, S. (2002). A new kind of science. Champaign, IL: Wolfram media.
63. Denning, P. J. (2017). Computational thinking in science.
64. Lloyd, S. (2006). Programming the universe: a quantum computer scientist takes on the cosmos. Vintage.
65. Zuse, K. (2013). Rechnender raum. Springer-Verlag.
66. Thagard, P. (2005). Why is beauty a road to the truth?. In Cognitive structures in scientific inquiry (pp. 365-374). Brill Rodopi.
67. Kuipers, T. (2002). Beauty, a Road to The Truth. Synthese131 (3), 291-328.
68. Girod, M. (2007). A conceptual overview of the role of beauty and aesthetics in science and science education.
69. Ivanova, M. (2017). Aesthetic values in science. Philosophy Compass, 12(10), e12433.
70. Dyson, F.J. (1987) Paul Dirac. Obituary Notice, American Philosophical Society Yearbook for 1986, 100-5.
71. McAllister (1996) Beauty and Revolution in Science. Ithaca, NY: Cornell University Press.
72. Zeki, S., Romaya J.P., Benincasa, D.M.T., Atiyah M.F. (2014) The Experience of Mathamtical beauty and its Neural Correlates. Frontiers in Human Neuroscience, 8: 1-12.
73. Zee, A. (1999) Fearful Symmetry: the Search for Beauty in Modern Physics. Princeton, Princeton University Press
74. Thuan, T. (2001) Chaos and Harmony: Perspectives on Scientific Revolutions of the Twentieth Century. New York, Oxford University Press.
75. Gross, D. J. (1996). The role of symmetry in fundamental physics. Proceedings of the National Academy of Sciences, 93(25), 14256-14259.
76. Xu, K., Li, J., Zhang, M., Du, S. S., Kawarabayashi, K. I., & Jegelka, S. (2019). What Can Neural Networks Reason About?. arXiv preprint arXiv:1905.13211.
77. Saxton, D., Grefenstette, E., Hill, F., & Kohli, P. (2019). Analysing mathematical reasoning abilities of neural models. arXiv preprint arXiv:1904.01557.
78. Zhang, C., Gao, F., Jia, B., Zhu, Y., & Zhu, S. C. (2019). Raven: A dataset for relational and analogical visual reasoning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 5317-5327).
79. Barrett, D. G., Hill, F., Santoro, A., Morcos, A. S., & Lillicrap, T. (2018). Measuring abstract reasoning in neural networks. arXiv preprint arXiv:1807.04225.
80. Davis, E., Marcus, G. (2019). Rebooting AI: building artificial intelligence we can trust. Knopf Doubleday Publishing Group.
81. Zenil, H., Kiani, N. A., Zea, A. A., & Tegnér, J. (2019). Causal deconvolution by algorithmic generative models. Nature Machine Intelligence, 1(1), 58.
82. Zenil, H. (2019). Compression is Comprehension, and the Unreasonable Effectiveness of Digital Computation in the Natural World. arXiv preprint arXiv:1904.10258.
83. Marcus, G. (2018). Deep learning: A critical appraisal. arXiv preprint arXiv:1801.00631.
84. Wu, T., & Tegmark, M. (2018). Toward an AI physicist for unsupervised learning. arXiv preprint arXiv:1810.10525.
85. Clune, J. (2019). AI-GAs: AI-generating algorithms, an alternate paradigm for producing general artificial intelligence. arXiv preprint arXiv:1905.10985.
86. Mitchell, M. (2019). Artificial Intelligence: A Guide for Thinking Humans. Farrar, Straus and Giroux.
87. Roscher, R., Bohn, B., Duarte, M. F., & Garcke, J. (2019). Explainable Machine Learning for Scientific Insights and Discoveries. arXiv preprint arXiv:1905.08883.

88. Colton, S., & Steel, G. (1999). Artificial intelligence and scientific creativity. Artificial Intelligence and the Study of Behaviour Quarterly, 102.
89. Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., & Pellegrino, D. (2009). Towards an explanatory and computational theory of scientific discovery. Journal of Informetrics, 3(3), 191-209.
90. Gobet, F., & Sala, G. (2019). How Artificial Intelligence Can Help Us Understand Human Creativity. Frontiers in Psychology, 10, 1401.
91. Sozou, P. D., Lane, P. C., Addis, M., & Gobet, F. (2017). Computational scientific discovery. In Springer Handbook of Model-Based Science (pp. 719-734). Springer, Cham.
92. Garcez, A. D. A., Gori, M., Lamb, L. C., Serafini, L., Spranger, M., & Tran, S. N. (2019). Neural-Symbolic Computing: An Effective Methodology for Principled Integration of Machine Learning and Reasoning. arXiv preprint arXiv:1905.06088.
93. Hill, F., Santoro, A., Barrett, D. G., Morcos, A. S., & Lillicrap, T. (2019). Learning to make analogies by contrasting abstract relational structure. arXiv preprint arXiv:1902.00120.
94. Chalmers, D. J., French, R. M., & Hofstadter, D. R. (1992). High-level perception, representation, and analogy: A critique of artificial intelligence methodology. Journal of Experimental & Theoretical Artificial Intelligence, 4(3), 185-211.
95. Hofstadter, D. R. (1995). Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought. Basic books.
96. Mitchell, M. (1993). Analogy-making as perception: A computer model. Mit Press.