Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 221 (2023) 227–231

Procedia Computer Science

www.elsevier.com/locate/procedia

10th International Conference on Information Technology and Quantitative Management

State space search revisited from perspective of deep learning

Zhengxin Chen

Department of Computer Science, College of IS&T, University of Nebraska at Omaha, USA

Abstract

Recent developments in deep learning have reshaped the landscape of artificial intelligence (AI). In this short research review note, we revisit the popular concept of state in state space search (as well as some related concepts) in AI from the perspective of deep learning. The major takeaway of this examination is that deep learning makes the philosophical concept of emergence operationalizable. In addition, along with other recent developments such as quantum computing, deep learning will have even more profound impact on the development of AI in near future.

Keywords: State space search, artificial intelligence (AI), deep learning

1. Introduction

Deep learning (see [10] for a recent introduction), as the landmark of contemporary AI (Artificial intelligence), shares with the traditional AI in the ultimate goal of intelligent problem solving. From a comprehensive and historical perspective, the idea of intelligent problem solving can be traced back to Polya's influential book on problem solving initially published almost 80 years ago [19], which heavily inspired research of traditional AI in its early days. In this short research review note, we revisit the popular concept of state in state space search (as well as some related concepts) in AI from the perspective of deep learning. The major takeaway of this examination is that *deep learning makes the philosophical concept of emergence operationalizable*. In addition, we point out that along with other recent developments such as quantum computing, deep learning will have even more profound impact on the development of AI in near future.

2. A quick review of state space search and related topics in traditional AI

2.1. States and state space search

State space search serves as the foundation for traditional AI problem solving and appears in the first chapter of every textbook of traditional AI (e.g., [21]). In the context of state space search, a state is a time snapshot representing certain aspects of the problem. A state contains all of the information necessary to predict the effects of an action and to determine if it is a goal state. A state is a representation of problem elements at a given moment. According to [6], problem solving systems have three components: a database,

10.1016/j.procs.2023.07.031

a set of operators to manipulate the database, and a control strategy for deciding what to do next. A State space is the set of all states reachable from the initial state by applying the appropriate control strategy.

Polya [19] presented four principles of solving a mathematical problem, but since these principles are quite general, their applications are far beyond mathematics. A particular interesting aspect of this book is that it contains a dictionary style heuristics, which are rule of thumbs that can lead to successful problem solving. Polya's book, along with AI pioneers' influential work, became the early inspiration for AI research, and heuristic search had its heydays in 1970s and 1980s.

2.2. A broader examination of encapsulated entities within AI

The concept of state is an encapsulated entity; so the examination of state space search as presented in the previous section leads to a broader examination of encapsulated entities within AI. Below are some examples.

(a) *Dynamic memory and case-based reasoning.* As a well-known research discipline, case-based reasoning, based on Schank's theory of dynamic memory [22], refers to using old experiences to understand and solve new problems. In case-based reasoning, a reasoner remembers a previous situation similar to the current one and uses that to solve the new problem. Case-based reasoning can be examined as a kind of classification algorithm [10]. Although the concept of case is different from the concept of state in state space search, just like a state, a case can be viewed as a high level semantic unit so it largely shares the idea of symbolic reasoning as state space search does, even it puts emphasis on episodic memory.

(b) *Analogy as structural mapping.* Another influential direction of work can be found in Gentner's structure-mapping theory [8] which views an analogy as a mapping of knowledge from one domain (the base) into another (the target); as a consequence, a system of relations which holds among the base objects also holds among the target objects. To visualize such a theory, we may imagine the base as a graph containing objects as vertices and their associations as edges; both the vertices and edges in the base are mapped to the target domain which is also made with a graph with vertices and edges. Both base and target analogs are high-level encapsulated symbols.

(c) *COGMIR: A cognitive model of information retrieval.* The idea of structural mapping can be generalized to go beyond analogical reasoning for knowledge integration, as shown in the work of COGMIR [5], which performs structural mapping in a graph with documents as the basic elements (or nodes). In COGMIR, a scheme, named query invoked memory reorganization, is used for knowledge integration. Unlike some other schemes which realize knowledge integration through subjective understanding by representing new knowledge in terms of existing knowledge, the proposed scheme suggests at storage time only recording the possible connection of knowledge acquired from different documents. The actual binding of the knowledge acquired from different documents is deferred to query time, depending on the actual needs of the query. Therefore, although there is only one way to store knowledge, there are potentially numerous ways to utilize the knowledge. Since the basic elements in COGMIR are documents, they are again encapsulated entities.

In summary, what are the common features of these techniques developed or matured in 1970s and 1980s? As we have already emphasized, they all manipulate highly encapsulated symbols using certain strategies or mapping rules. Another noticeable feature is that initial status of the involved systems is given as input. An apparent advantage of such kind of settings is that since the encapsulated symbols are high-level abstract entities, the reasoning process for problem solving can be significantly simplified. However, the shortcomings of this approach are also quite obvious. The concept of state as in state space search, the concept of case in case-based reasoning, as well as the concept of object in object-oriented paradigm, are all encapsulated entities. In state space search, the initial states are given, and from the initial states, new states can derived using appropriate control strategies. Where are these high-level abstractions coming from? It is somewhat mysterious to find the original source of origins. A simple, yet powerful answer eventually has arrived at deep learning era. The magic word for the behind-scene story can be summarized as one single word: emergence.

3. Deep learning makes philosophical concept of emergence operationalizable

3.1. The concept of emergence

Around the turn of this century, several interesting and broad examinations on the philosophical concept of emergence appeared [12,14]. Simply put, emergence occurs when an entity is observed to have properties its parts do not have on their own, properties or behaviors that emerge only when the parts interact in a wider whole. Emergence plays a central role in theories of integrative levels and of complex systems [27]. However, for a couple of decades, there has been an irony on the study of emergence, because although this philosophical concept is attractive, itself becomes a mystery, namely, how to make it happen?

The advent of deep learning comes to rescue. Deep learning, just like its close relative representation learning [2.3], is aimed to automatically learn effective representation (such as features or attributes) from the input data to facilitate data mining tasks such as classification or clustering [10]. By employing multi-level deep neural networks and appropriate algorithms, high level symbolic abstractions can come into being from low-level physical units through sophisticated inter-connections through many rounds of iterations (or "epochs" as sometimes it is referred to).

The main theme of this research note lies in the claim that deep learning has *operationalized* the concept of emergence. This is revealed in a number of recent studies. Below we will revisit state space search from a deep learning perspective.

3.2. State space search and deep learning

The relationship between state space search and deep neural networks is quite interesting. Many studies in deep learning involve the use of convolutional neural networks (CNNs) to deal with grid-type data. For example, [9] studied whether CNNs learn semantic parts in their internal representation by investigating the responses of convolutional filters and associating their stimuli with semantic parts. In addition, the authors explored several connections between discriminative power and semantics. They also did investigation to find out which semantic parts are the most discriminative and whether they correspond to those parts emerging in the network. This helps gain an even deeper understanding of the role of semantic parts in the network.

Soatto [24] presented an emergence theory of deep learning, started as a general theory for representations, and is comprised of three parts. Along with this research, reference [1] showed that that invariance to nuisance factors in a deep neural network is equivalent to information minimality of the learned representation.

An interesting experimental study was carried out by Lu et al [17] which noted that humans are able to learn abstract semantic relations (e.g., antonym, synonym, category membership) and use them to reason by analogy. A deep theoretical challenge is to show how such abstract relations can arise from nonrelational inputs, thereby providing key elements of a protosymbolic representation system. The authors developed a computational model and conducted experiments that exploit the potential synergy between deep learning from "big data" (to create semantic features for individual words) and supervised learning from "small data" (to create representations of semantic relations between words). In another study, by carefully designing data models for the visual scene, [13] showed that the emergence of this pattern is triggered by the non-Gaussian, higher-order local structure of the inputs, which has long been recognized as the hallmark of natural images.

In addition, deep learning can also provide similar role for case-based reasoning. For example, [17] used a deep neural network to explain its own reasoning for each prediction.

In summary, recent experimental studies have provided physical evidence of operationalizable emergence of symbolic states or cases, and this direction of research is showing increasing promises. It is also interesting to note that the notion of state and the method of state space search can benefit the study of deep learning as well, as shown in [11]; therefore, the studies of deep learning and state space search can benefit each other.

4. Issues on operationalizing emergence

There are various issues to be considered for operationalizing emergence in deep AI, such as methodologies needed to operationalize it, relationship with other research disciplines, etc. Some of these issues or scenarios are discussed below.

(a) *Let states and cases emerge: Integrated process of deep learning and higher-level tasks.* In this scenario, deep learning can serve as a preprocessing step so that states as in state space search and cases as in case-based reasoning can eventually emerge. In addition to examples already reviewed in previous section, we can also add another study [25] where the authors devised a deep auto-encoder (DAE) to discover hierarchical non-linear functional relations among regions, and then the regional features into an embedding space, whose bases are complex functional networks. In our view, more studies along this direction are still needed, so that a more general methodology can be developed.

(b) Quantum deep learning. Quantum computing refers to the manipulation of quantum systems in order to process information. The ability of quantum states to be in a superposition can thereby lead to a substantial speedup of a computation in terms of complexity, since operations can be executed on many states at the same time [23]. Unlike bits on classical computers, which at any point will represent 1 or 0, qubits are able to "hover" between these two states. Only when measured will the qubit "collapse" into one of its states.

This property is known as superposition and is critical for quantum computing tasks. References [7,26] provided two additional overviews on basics of quantum deep learning.

(c) Incorporating granular computing (GrC) into deep learning. A research field more established than quantum computing but has not shown its full potential for deep learning yet, granular computing (GrC) typically employs fuzzy set theory or rough set theory to deal with uncertainty. As noted in [28], there are three basic concepts that underlie human cognition: granulation, organization and causation and a granule being a clump of points (objects) drawn together by indistinguishability, similarity, proximity or functionality. As an example of this direction of development, [20] suggested hybridization of deep learning and rough set based granular computing.

5. Conclusion: Everything boils down to connections and emergence

Convolutional neural networks (CNNs) and other types of data structures in deep learning are latest examples of *connections*, which play important roles in science history, and the topic is made popular by James Burke's influential book and TV series [4]. In our view, connections in science take place at very different levels, and in many different forms. In this sense, this paper is devoted to connections. The relationship between connections and emergence of new concepts has been long recognized, but it is still lacking in in-depth studies.

In this paper we have reviewed part of history related to state space search and related development. Some of the most important findings can be summarized as follows. The concept of state as appeared in state space search has been an extremely influential in the development of AI, as evidenced in case-based reasoning and analogy as structural mapping. However, there is a potential of mystery for the states involved in state space search, namely, how they come from. In early research of AI, such kind of states are treated as a high-level summarization or encapsulation. This could be a good starting point to simplify the scenario for intelligent problem solving, but it is not really practical in many applications. The good news eventually comes from recent studies involving deep learning, which shows that the formation of a state could be the result of emergence of deep learning. Combined with related research disciplines such as quantum computing, the study of deep learning implies an enlightenment of an exciting new research era.

Recent developments of deep learning have also inspired efforts in other directions of re-examination of AI. For example, a philosophical examination of current status of AI has resulted a proposal of applying the concept of state space to AI itself [18], which is conceived as a three-dimensional space of self-learning (from rule-based to learning-based), generalization (from narrow to general AI), and grounding (the degree of semantic world anchoring). Other researchers have noted profound relationships between deep learning and explainable AI (XAI) [15]. By combining these and other developments motivated from deep learning, we hope that the reexamination of state space search can provide even more meaningful insights for future research.

References

[1] Achille A, Soatto S (2018) Emergence of Invariance and Disentanglement, in Deep Representations, Journal of Machine Learning Research, 19: 1-34

[2] Bengio Y, Courville , Vincent p (2012) Unsupervised Feature Learning and Deep Learning: A Review and New Perspectives, https://arxiv.org/pdf/1206.5538v1.pdf

[3] Bengio Y, Courville , Vincent p (2013) Representation Learning: A Review and New Perspectives, IEEE Transactions on Pattern Analysis and Machine Intelligence 35(8):1798-1828 (version 3, 2014, https://arxiv.org/pdf/1206.5538.pdf)

[4] Burke J, Connections: Alternative History of Technology (Time Warner International/Macmillan 1978) ISBN 978-0-333-24827-0 published in North America as Connections (Little, Brown and Company, 1978) ISBN 0-316-11681-5 and pbk: ISBN 0-316-11685-8.

[5] Chen Z (1988) COGMIR: A Computer Model for Knowledge Integration, Louisiana State University, Baton Rouge. https://digitalcommons.lsu.edu/gradschool_disstheses/4625/

[6] Cohen PR, Feigenbaum EA (1982). The Handbook of Artificial Intelligence (Vol. I) Kaufmann, William Inc

[7] Garg S, Ramakrishnan G (2020), Advances in Quantum Deep Learning: An Overview, https://arxiv.org/pdf/2005.04316.pdf

[8] Gentner D (1983) Structure mapping: a theoretical framework for analogy, Cognitive science, 7(2):155-170.

[9] Gonzalez-Garcia A, Modolo D, Ferrari V (2017). Do Semantic Parts Emerge in Convolutional Neural Networks? Int J Comput Vis (2018) 126:476–494

[10] Han J, Pei J, Tong H (2022) Data mining: concepts and techniques (4th ed.), Elsevier.

[11] Hauser M, Gunn S (2019) State-Space Representations of Deep Neural Networks (letter), Neural Computation 31, 538–554

[12] Holland JH (1999) Emergence: From Chaos To Order, Helix

[13] Ingrosso A, Goldt S (2022), Data-driven emergence of convolutional structure in neural networks. PNAS 119 (40) e2201854119, https://doi.org/10.1073/pnas.2201854119

[14] Johnson S (2002) Emergence: The connected lives of ants, brains, cities and software, Simon & Schuster

[15] Leventi-Peetz AM, Östreich T (2022) Deep Learning Reproducibility and Explainable AI (XAI), arXiv:2202.11452v3, https://doi.org/10.48550/arXiv.2202.11452

[16] Li O, Liu H, Chen C, Rudin C (2018) Deep Learning for Case-Based Reasoning through Prototypes: A Neural Network that Explains Its Predictions, AAAI.

[17] Lu H, Wu YN, Holyoak K (2019) Emergence of analogy from relation learning, PNAS, 116 (10): 4176-4181
[18] Lyre H (2020), The State Space of Artificial Intelligence, Minds and Machines (30)325–347, https://doi.org/10.1007/s11023-020-09538-3
[19] Polya G (1957) How to Solve It (2nd ed.). Doubleday Anchor, New York.
[20] Ropiak K, Artiemjew P (2020), On a hybridization of deep learning and rough set based granular computing, Algorithms, 13(3).
[21] Russel S, Norvig P (2020) Artificial intelligence: A modern approach (4th ed.), Pearson.
[22] Schank RC (1983) Dynamic memory: A theory of reminding and learning in computers and people, Cambridge University Press.
[23] Schuld M, Sinayskiy I, Petruccione F (2014), An introduction to quantum machine learning, Contemporary Physics, 56(2).
[24] Soatto S, Emergence theory of deep learning, http://www.ipam.ucla.edu/abstract/?tid=14550&pcode=DLT2018
[25] Suk A, Wee C-Y, Lee S-W, Shen D, State-space model with deep learning for functional dynamics estimation in resting-state fMRI, Neuroimage. 2016 April 01; 129: 292–307. doi:10.1016/j.neuroimage.2016.01.005
[26] Wiebe N, Kapoor A, Svore KM (2014) Quantum Deep Learning. *arXiv preprint arXiv:1412.3489*
[27] Wikipedia, Emergence. https://en.wikipedia.org/wiki/Emergence
[28] Zadeh LA (1997) Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems, Volume 90, Issue 2, 1 September 1997, Pages 111-127