Toward Compositional Generalization with Neuro-Symbolic AI: A Comprehensive Overview

Nassim Lateb, Farida Bouarab-Dahmani, Florence Sèdes

▶ **To cite this version:**

Nassim Lateb, Farida Bouarab-Dahmani, Florence Sèdes. Toward Compositional Generalization with Neuro-Symbolic AI: A Comprehensive Overview. 4th IEEE international Conference on Embedded & Distributed Systems (EDIS4 2024), University of Bechar-Tahri Mohamed, Bechar, Algeria, Nov 2024, Bechar, Algeria. à paraître. hal-04705213

HAL Id: hal-04705213
https://hal.science/hal-04705213v1

Submitted on 22 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Toward Compositional Generalization with Neuro-Symbolic AI: A Comprehensive Overview

Nassim Lateb, Farida Bouarab-Dahmani, Florence Sèdes

Mouloud Mammeri University of Tizi Ouzou, Tizi Ouzou, Algeria
IRIT, University of Toulouse 3 Paul Sabatier, Toulouse, France
lateb.nsm@gmail.com, bouarabfarida@ummto.dz, Florence.Sedes@irit.fr

Abstract—This review emphasizes the fascinating convergence of Neuro-Symbolic AI (NeSy) and Compositional Generalization (CoGe), examining how these models might potentially transform AI by enabling real human-like intelligence. This research contends that NeSy's capacity to combine the advantages of neural and symbolic techniques has enormous potential for addressing the CoGe dilemma. CoGe necessitates the ability to learn and use information in unexpected settings through the flexible assembly of existing building components. NeSy architectures, with their distinct combination of symbolic reasoning and flexible learning, provide a viable option for overcoming this critical hurdle. In this paper, we highlighted some of the most important concepts of both NeSy and CoGe, showcasing the cutting-edge research trends shaping these fields, we delve into their diverse techniques and methods. Drawing upon cognitive science studies and concrete AI-based works, we illustrate the multitude of possibilities for implementing CoGe within NeSy. Finally, we discuss the results performed by these studies, their commonalities, we then present our proposition and address the open challenges that lie ahead on the path towards true CoGe with NeSy.

Keywords—Inductive Learning - Compositional Generalization - Neuro-Symbolic AI - Knowledge Representation.

I. INTRODUCTION

The fundamental concept of generalization in AI is the ability to transfer gained knowledge and learning from previous experiences to deal with novel scenarios and guide actions. This necessary mechanism reflects the underlying character of human intellect, notably in terms of CoGe. Our knowledge representation is naturally compositional [1]. Human cognitive abilities to solve complex problems and represent abstract concepts rely significantly on it by applying common and simpler components learned in a particular context and attempting to generalize (transfer) them to novel and related settings via composition functions. Smolensky described this as “the Central Paradox of Cognition” [2]. The human brain's limited physical resources have been shown to develop powerful and versatile cognitive capacities, posing a well-known issue in neuroscience study. Adopting a CoGe architecture is critical for machine learning systems looking to improve their performance and dependability. Pure connectionist models, although effective, have numerous significant problems. These include an insatiable desire for enormous datasets to learn and extract meaningful patterns, the famed black-box effect, which impairs reasoning transparency, and the ongoing challenge of efficiently generalizing to unexpected scenarios. The human brain's limited physical resources have been shown to develop powerful and versatile cognitive capacities, posing a well-known issue in neuroscience study. Adopting a CoGe architecture is critical for machine learning systems looking to improve their performance and dependability. Pure connectionist models, although effective, have numerous significant issues. These include an insatiable desire for enormous datasets to learn and extract meaningful patterns, the famed black-box effect, which impairs reasoning transparency, and the ongoing challenge of efficiently generalizing to unexpected scenarios. Introducing NeSy AI systems, which are positioned as promising answers to the foregoing difficulties in practical domains such as healthcare, brain and behavior modeling, natural language processing [7], cybersecurity, and privacy [8], but their integration is not without obstacles. Challenges remain in seamlessly blending symbolic reasoning with neural networks, properly encoding information and embedding, and building models that balance learning and reasoning skills. Despite these obstacles, the increased curiosity in the field of NeSy systems is significant. The change from a single study published before 2016 to about 40 publications published between 2017 and 2021 illustrates the increasing enthusiasm and changing environment of NeSy research [3]. The aim of this paper is to reveal the active research still being conducted to determine how NeSy AI systems may fit and adopt successfully using the compositionality structure, focusing on particular strategies for greater generalization. It is divided as follows: **Section 2** introduces the NeSy AI, its categorization, knowledge representation and embedding, and active research directions. **Section 3** discusses the CoGe, starting with definitions, knowledge representation, highlighting the relation between CoGe, context-understanding and transfer learning, finishing with recent researches. **Section 4** emphasizes the combination of NeSy and CoGe by presenting our classification and relevant works, a cognitive science research and AI based studies, we also discuss theirs results and commonalities, we present the future research directions, our proposition and upcoming implementation

979-8-3315-1987-2/24/$31.00 ©2024 IEEE

model. Finally, **Section 5**, holds the paper's conclusion.

II. NEURO-SYMBOLIC AI

A. Categorization

NeSy AI architecture aims to unify the best of the both worlds, the neural and symbolic. There are mainly two principal propositions, the first one was a survey presented by *P.Hitzler and S.Bader* [6] in 2005 giving to NeSy integration according to eight dimensions, grouped into three aspects: *Interrelation*, *Language* and *Usage*, and recently, *Henry Kautz at AAAI 2022 Robert S. Engelmore Memorial Award Lecture* has presented his categorization [5] as six different modes of integrating symbolic and connectionist approaches:

- *Symbolic Neuro Symbolic:* the inputs and outputs are symbols, NLP systems including LLMs like GPT-3.
- *Symbolic [Neuro]:* neural modules are internally used in a symbolic problem solver.
- *Neuro | Symbolic:* neural and symbolic parts focus on different but complementary tasks.
- *Neuro:Symbolic ⇒ Neuro:* symbolic rules/knowledge are compiled into the architecture of neural networks.
- *NeuroSymbolic:* turning symbolic knowledge into additional soft-constraints in the loss functions of DNNs (Deep Neural Networks).
- *Neuro[Symbolic]:* fully integrated systems that embed a symbolic reasoning system inside a neural engine.

B. Knowledge Representation and Embedding

We have to emphasize how the knowledge is represented and embedded in pure connectionist, symbolic and NeSy AI systems, **Table 1** summarizes these two aspects of NeSy AI system.

TABLE 1.
Knowledge Representation and Embedding in NeSy AI

Models	Knowledge Representation	Knowledge Embedding
Connectionist Models	Sub-Symbolic Representations[1].	— Representing features. - continuous vector embeddings in a high-dimensional space. — capturing semantic relationships among entities.
Symbolic Models	— Explicit symbols and formal rules. — Represent knowledge structurally.	- Transforming discrete symbols and formal rules, into continuous vector representations.
NeSy AI Models	Combining symbolic and sub-symbolic knowledge representation techniques	Depending on the model's architecture and type.

[1] The internal formats used by AI models to encode information about the world, they're more abstract and hidden within the model's internal structure.

C. Open Challenges

The author in [1] has presented the upcoming challenges to overcome in NeSy realm, here are the main ones:
1) Scalability: NeSy systems are still struggling with large-scale symbolic/logic reasoning.
2) Compositional Generalization: a central aspect of human intelligence, among the most desirable characterizations that NeSy systems are expected to offer.
3) Automated Knowledge Acquisition: closely related to the concept of '*Learning to Reason*' [4]. NeSy systems are ignoring two crucial issues: i) how to acquire contextual knowledge ? ii) how to ensure that the acquired knowledge is the most suitable for the system's functionality ?

III. COMPOSITIONAL GENERALIZATION

A. Definitions

CoGe represents the ability to utilize a finite set of known components to understand and generate a vast array of novel combinations in different but related contexts [9]. These components can be represented as learned knowledge or concepts relying on methods like symbolic techniques.

N. Chomsky, in his book "*Aspects of the Theory of Syntax"*, defined it as "*infinite use of finite means*" [17] in the context of language, grammar, and linguistics. The combination process is performed systematically and most of the time hierarchically to keep a comprehensive meaning and to fit well with the novel contexts.

R. Montague [18] has worked on the compositionality principle in language, and he claimed that compositionality requires a relationship between expressions of a language and the meaning of those expressions. He introduced the concept of collective and global compositionality. The human cognitive system's ability to perform CoGe and *transfer learning*[2] is a complex interplay of neural mechanisms, inductive learning processes, and *cognitive biases*[3]. These elements work together to enable adaptive learning, problem-solving, and the application of knowledge across diverse contexts.

B. Knowledge Representation and Key Aspect

[2] ML technique that involves leveraging the knowledge gained from solving one problem and applying it to a different but related problem.
[3] Deviation from rationality in judgment, often influenced by subjective factors that can distort the decision-making.

According to [1], "*Our knowledge is naturally compositional*", in fact, it's a multifaceted topic that has been studied in cognitive science. The prevailing view is that the human brain employs a compositional and distributed representation of knowledge, depending on the performed task and the executed context. **Table 2** summarize it.

TABLE 2.
Compositional and Distributed Aspects

Aspects	Compositional	Distributed
Distributed Representation	✗	✓
Sparse Coding[4]	✗	✓
Hierarchy and Compositionality	✓	✗
Symbolic Representation	✓	✗
Combination of Features	✓	✗
Shared Representations	✓	✗

C. *Relation between CoGe, Transfer Learning and Context-Understanding*

The author in [29] unveils the comparison between a general model that was trained using all contexts against a system that is composed of a set of specialized models that was trained for each particular operating context. Authors proposed a local learning approach based on context-awareness by training two distinct tasks: classification and control task, this method uses contextual information to separate the input space, train and select the suitable ML models. This solution has showed significant results, it alleviates bias (domain-specific knowledge) to an approach that uses a unique general and versatile model.

W. Battaglia and his colleagues [10] highly supported the idea that "*The question of how to build artificial systems which exhibit combinatorial generalization has been at the heart of AI since its origins*". Their work focuses on how graph networks can support *relational reasoning*[5] and combinatorial generalization. They've also explored how using *relational inductive biases*[6] within DL architectures can facilitate learning about entities, relations, and rules for composing them. The relational inductive biases are a kind of transfer learning technique that aim to enhance a model's ability to understand and reason about relationships. Another useful example of recent researches that clearly demonstrates the impact of in-context examples on CoGe performance [11], they carried on with the raising question: "*what are the key factors that can make good in-context examples for CoGe?"*, they studied three potential factors: Similarity, diversity, and complexity. The experiments indicate that in-context examples should be structurally similar to the test case, diverse from each other, and individually simple.

In essence, this relationship can be summarized as the following: context-understanding shapes the compositionality of representations in a specific context, and leveraging this gained knowledge then acts as an inductive bias (transfer learning) guiding the model to generalize compositional rules to new contexts. **Figures 1** and **2** provide two distinguish examples about it:

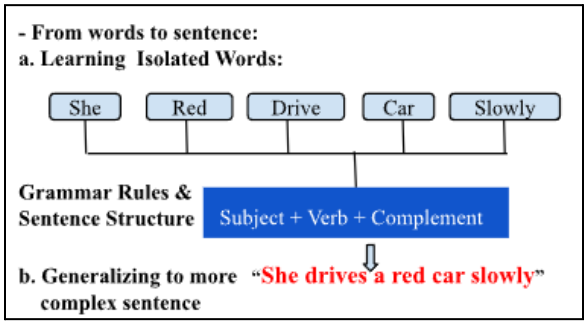

Figure 1. CoGe Ability in Human Cognitive System

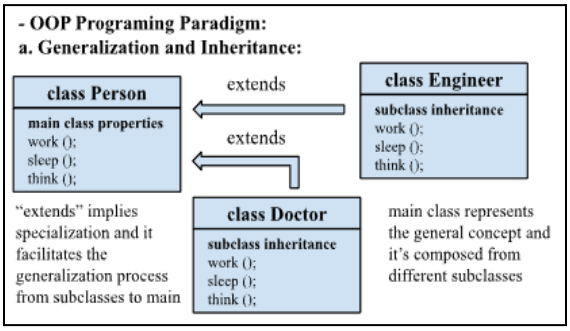

Figure 2. CoGe Ability in Software Engineering

D. *Upcoming Challenges*

The main question is: "*why connectionist and modern models lack of CoGe ability*" ? It is mainly due to:
1) Limited Symbolic Representations: struggle with incorporating symbolic representations and explicit rules, which are crucial for capturing the compositionality.
2) Lack of Hierarchical Abstraction: CoGe often requires hierarchical abstraction, where simple concepts combine to form more complex ones.
3) Implicit Representations: ANNs encode information in the weights and connections between neurons, leading to distributed and implicit representations for efficient learning, it also makes it difficult to interpret what a specific neuron or group of neurons represent.

[4] Involves representing information using a subset of active neurons for a given concept or representation.
[5] Models that can reason and draw conclusions based on the relationships between different elements in a given context.
[6] Refer to biases that encourage models to consider and exploit relationships between elements in the data.

IV. COGE THROUGH NESY AI

A. Classification

In this section, we have proposed a classification for CoGe approaches based on recent works:

1) Neural based CoGe: neural networks are used to learn compositional structures directly from data:

- *Attention Mechanisms:* mechanisms like *Transformers* focus on specific parts of the data that are most relevant for a particular task (identify and learn compositional structures).
- *Hierarchical Neural Networks:* models like *CNNs* for image or *RNNs* for text data process at different levels of granularity, allowing them to capture both individual elements and their relationships within a composition.

2) Symbolic based CoGe: encoding compositional knowledge using symbolic representations like rules or logic:

- *Knowledge Graph Embedding:* incorporating knowledge graphs (KGs) to represent domain knowledge symbolically (entities + relationships), techniques like *TransE* or *RotatE.*
- *Logical Reasoning Rules:* incorporating logic rules into NeSy models, derived from domain knowledge or expert insights, guiding it towards compositional structures.

3) NeSy Integration Techniques:

- *Neural Modules Augmentation:* integrating a symbolic reasoning module into ANN to guide the learning process, to ensure it adheres to the symbolic knowledge.
- *NeSy Co-Learning:* training the ANN and the symbolic module simultaneously. The symbolic module guides the ANN towards relevant compositional structures, and the ANN potentially refine the symbolic knowledge base.

B. Cognitive Science Research

Starting with some insights drawn from a psychological and cognitive sciences article released on August 2022 [13], the authors conducted an experience on 605 subjects (293 female and 312 male) to discover the determinants and limits of human transfer learning ability using curriculum learning for CoGe. They also explored how it might be modeled in a neural network. They did this by designing a task that required participants to learn and generalize a mapping function from symbolic cues to spatial locations. Furthermore, they also find that adult humans tend to learn *composable functions asynchronously*[7], and the human ability to handle novelty (generalization) seems to depend on transfer distance: near-transfer is when old problems and new ones share physical features, and it's the most successful method, far-transfer is when problems share common structure yet are superficially distinct, and it's less successful. In this paper [12], the authors studied the computational properties associated with CoGe in both humans and ANNs on a highly compositional tasks. Here are the steps:

- Characterizing the behavioral signatures of CoGe in a task that varied rule conditions across 64 contexts.
- Analyzing *fMRI data*[8] gained from the experience
- Designing pretraining paradigm aided by primitives to endow compositional task elements into ANNs.

The conclusions derived from this experience are:

- Humans generalize better to tasks with greater similarity structure to previous tasks (related contexts).
- Pretrained ANNs models exhibit more abstract representations, excellent generalization performance, and efficient learning.

C. Recent AI Works

In this section, we'll present four recent works, in the first paper. The authors [14] tackled the limitations issue of DL models when dealing with CoGe, the capability to learn compositional rules and apply them to unseen cases systematically, by introducing the Neural-Symbolic Stack Machine (NeSS)*).* NeSS combines the expressive power of neural sequence models with the recursion supported by the symbolic stack machine. The results have shown 100% generalization performance in four domains: the *SCAN*[9] benchmark, the task of few-shot learning of compositional instructions, the compositional machine translation benchmark, and context-free grammar parsing tasks. Curriculum Training technique scheme is used during the training process.

The second research [25] concerns the learning of compositional rules via neural program synthesis techniques. The authors propose a NeSy model that can learn entire rule systems from a small set of examples by combining neural networks and symbolic program representations. They demonstrate that their model achieves human-level performance on a few-shot artificial language learning task, improves over existing benchmarks on the SCAN challenge for CoGe, and can learn to interpret number words across natural languages.

The third paper has been carried out in May 2023 [15], the two authors addresses the CoGe challenge with respect to the two specific tasks of word problem-solving and visual relation recognition and propose a NeSy solution, using *DeepProbLog*[10] facing the issue of neural models and their inability to generalize compositionally, unless, by providing it with more data that has a sufficient number of instances of every combination of concepts, they are interested in. Incorporating task specific knowledge into purely neural models help address its inability to generalize compositionally.

[7] Discontinuities in learning that resemble those seen in child development.

[8] Functional Magnetic Resonance Imaging, it, provides insights into which regions of the brain are active during specific tasks.

[9] Benchmark to assess model's ability in compositional tasks.

[10] Neural probabilistic logic programming language that integrates deep learning through neural predicates.

The fourth paper research has been conducted in last Jan 2024 [16], the authors have proposed a novel approach based on NeSy architecture called the Compositional Program Generator (CPG). CPG has three key features: modularity, type abstraction, and recursive composition, that enable it to generalize systematically to new concepts in a few-shot manner. It uses context-free grammar rules to generate a hierarchical, abstract parse of the input, and generates rule-specific probability distribution parameters for probabilistic copy or substitution programs, using curriculum learning and evaluate on SCAN and COGS benchmarks.

D. Discussion

In the last section, we have listed some recent studies about the possibility of applying NeSy architecture to handle the CoGe issue, faced by the current neural network models. The first scientific research we've presented, shows that the human learns and generalize well compositionally. The second research confirms this, and follow by attesting that humans generalize better to tasks with greater similarity structure to previous tasks (related contexts). They have also compared how abstract representations are distributed across the entire cortex in a content-specific way during the execution of the compositional task with pretrained ANNs.

The recent AI works have showed many commonalities:

- Using NeSy as a foundation architecture to handle the different components from the input data.
- Curriculum Learning technique used during the training process for an incremental and structured learning.
- Leveraging pretrained NeSy models like DeepProbLog.
- Relying on specific knowledge as an input data to enhance the compositionality learning and context-understanding.
- The evaluation task is usually done on two popular compositionality benchmarks, SCAN and COGS.
- Common application domains: NLP (Natural Language Processing), semantic parsing, translation and language understanding, because of its natural compositionality (the meaning of the sentence equals to the sum of its words).

E. Proposition and Upcoming Implementation

Our contribution lies on a novel NeSy architecture applied in a new domain field, to tackle the CoGe challenge.

1) Generalization Through Contexts-Representation: Authors in [21] have defined context based on a corpus of 150 definitions: The context acts like a set of *constraints* that *influence* the *behavior* of a *system* (a *user* or a *computer*) embedded in a given *task*. In [22], authors have made a comparison between contextual and non-contextual features in ANNs for recommender system, they find out that adding dynamic environment features (day-time, location, season, etc…) outperforms the classical model based only on user's preferences.

In addition to ANNs , context can be handled relying on symbolic rules, knowledge graphs or First-Order Logic (FOL). Example: Rule-based representation: *IF (location = 'office' AND time = 'weekday' AND device = 'work computer') THEN (context = 'work')*. Using FOL: ∀x, y (location(x, 'office') ∧ time(y, 'weekday') ∧ device(z, 'work computer') → context(x, y, z, 'work')).

2) Compositionality through MTL Architecture: Our model is based on a Multi-Tasks Learning (MTL) architecture, It consists of training one main task model (usually a DL one) and a set of related tasks in the output layer. We try to improve the generalization on the main task by using the related tasks as an inductive signal [19][20]. **Figure 3** shows this architecture with details.

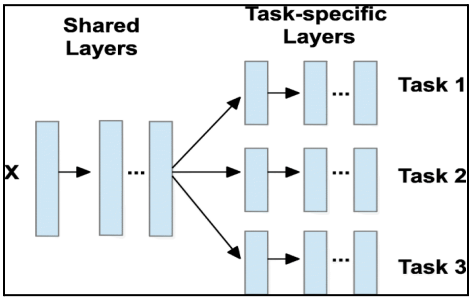

Figure 3. The MTL Architecture

3) Application: Computerizing of Habit-Forming Process: This 2023 research paper [23] emphasizes the introduction of ML method that discovers which of many context variables are associated with behavior and predicts how quickly habits form (Predicting Context Sensitivity), applied in two different large panels datasets.

Authors in this 2022 article [24] have proposed a method to enable intelligent systems to compute habit strength based on observable behavior, the advantage of using computed habit strength for behavior prediction was tested using data from two intervention studies on dental behavior change.

Comparing to these recent researches, our aim lies on classifying behaviors into: *Sporadic* or *Habitual* ones (Main Task), another way for analyzing consumer's behaviors. In neuroscience. "Habits" are defined as a set of hierarchical steps: context, cues, action, and reward. To succeed their modeling, we have summarized it as follows: ***Habits = Context [Behaviors (Repetition + Positive Reinforcement)].***

4) Hard Parameters Sharing Approach: it is the most commonly used approach to MTL in neural Networks. Following the figure 3 and the habit's equation, here are the implementation details:

- ***Shared Layers:*** Contextual-Embeddings: Context can

be modeled relying on symbolic rules, knowledge graphs or First-Order Logic (FOL). Example: Rule-based representation: *IF (location = 'office' AND time = 'weekday' AND device = 'work computer') THEN (context = 'work')*. Using FOL: ∀x, y (location(x, 'office') ∧ time(y, 'weekday') ∧ device(z, 'work computer') → context(x, y, z, 'work')).

- ***Task 01:*** Sequential Behavior Analysis: leveraging LSTM algorithm for sequential analysis, Relying on Symbolic Rules concerning the temporal patterns and recurring sequences of behaviors (daily or weekly repetition rate).
- ***Task 02:*** Positive Reinforcement: relying on Q-Learning algorithm to support the positive behaviors.
- ***Task 03:*** Habit Detection: binary classification algorithm, classifying behaviors into Habitual or Sporadic ones.
- ***Hierarchical Training:*** according to neuroscience, habits are a natural hierarchical process, curriculum learning is a promising method to mimic effectively the natural process.
- ***Loss Functions (LF):*** e.g., for M tasks, $LF = \sum Wi\ Li$, i=1 to M, Li represents each task's LF, Wi is an assigned weight.

V. CONCLUSION

CoGe is definitely a game-changer characteristic for novel AI models. To reach this challenge, the NeSy AI systems are presented as promising methods, and to foster its implementation, future researches must focus primarily on integrating reasoning ability into ANNs, and then designing well suited AI models architectures to handle CoGe. At the end, regarding its importance, CoGe must be applied in impactful domains like healthcare, NLP, and computer vision.

VI. REFERENCES

[1] Wenguan Wang, Yi Yang, Fei Wu, "Towards Data-and Knowledge-Driven Artificial Intelligence: A Survey on Neuro Symbolic Computing", IEEE transactions on pattern analysis and machine intelligence, 12 Oct 2023.

[2] Paul Smolensky, "Behavioral and brain sciences", Cambridge University Press, 1988.

[3] Md Kamruzzaman Sarker, Lu Zhou, Aaron Eberhart, Pascal Hitzler, "Neuro-Symbolic Artificial Intelligence: Current Trends", eprint arXiv:2105.05330, 14 Mai 2021.

[4] Roni Khardon, Dan Roth, "Learning to reason", Journal of the ACM Volume 44 Issue 5, pp 697–725, 01 September 1997.

[5] Henry A. Kautz, "The third AI summer: AAAI Robert S. Engelmore Memorial Lecture", 31 March 2022.

[6] Sebastian Bader, Pascal Hitzler, "Dimensions of Neural-symbolic Integration - A Structured Survey", Published in We Will Show Them, 10 November 2005.

[7] Djallel Bouneffouf, Charu C. Aggarwal, "Survey on Applications of Neurosymbolic Artificial Intelligence", 08 Sept 2022.

[8] Aritran Piplai, Anantaa Kotal, Seyedreza Mohseni, Manas Gaur, Sudip Mittal, Anupam Joshi, "Knowledge-enhanced Neuro-Symbolic AI for Cybersecurity and Privacy", July 2023.

[9] T.Wiedemer, P. Mayilvahanan, M. Bethge, W.Brendel, "Compositional Generalization from First Principles", arXiv.org, 10 July 2023.

[10] Peter W. Battaglia, Jessica B.Hamrick, Victor Bapst1, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, & 21 other authors, "Relational inductive biases, deep learning, and graph networks", Published 17 Oct 2018.

[11] Shengnan An, Zeqi Lin, Qiang Fu, Bei Chen, Nanning Zheng, Jian-Guang Lou, Dongmei Zhang, "How Do In-Context Examples Affect Compositional Generalization?", 09 Jun 2023.

[12] Takuya Ito, Tim Klinger, Douglas H. Schultz, John D. Murray ,Michael W. Cole, Mattia Rigotti, "Compositional generalization through abstract representations in human and artificial neural networks", NeurIPS 2022 36th Conference, 21 September 2022.

[13] Ronald B. Dekkera, Fabian Ottoa, and Christopher Summerfielda, "Curriculum learning for human compositional generalization", Pnas, Research Article, Psychological and Cognitive Sciences, 29 August 2022.

[14] Xinyun Chen, Chen Liang, Adams Wei Yu, Dawn Song, Denny Zhou, "Compositional Generalization via Neural-Symbolic Stack Machines", arXiv.org, 22 October 2020.

[15] Savitha Sam Abraham, Marjan Alirezaie, "Compositional Generalization and Neuro-Symbolic Architectures", CLeaR-Workshop Poster, 05 May 2023.

[16] Tim Klinger, Luke Liu, Soham Dan, Maxwell Crouse, Parikshit Ram, Alexander Gray, "Compositional Program Generation for Systematic Generalization", IJCAI 2023 Workshop, KBCG Oral , 28 September 2023.

[17] Noam Chomsky, "aspects of the theory of syntax", the m.i.t. press, published in 1965.

[18] Richard Montague, "Pragmatics and Intensional Logic", Vol. 22, No. 1/2, Semantics of Natural Language, II (Dec., 1970), pp. 68-94 (27 pages) Springer.

[19] Sebastian Ruder, "An Overview of Multi-Task Learning in Deep Neural Networks", ArXiv, Computer Science, Machine Learning, 15 June 2017.

[20] Shijie Chen, Yu Zhang, Qiang Yang, "Multi-Task Learning in Natural Language Processing: An Overview", arXiv, Computer Science, Machine Learning, 19 Sept 2021.

[21] Mary Bazire & Patrick Brézillon, "Understanding Context Before Using It", International and Interdisciplinary Conference on Modeling and Using Context, Springer, 2005: pp 29–40.

[22] G. Mustafa, I. Frommholz, "Comparing contextual and non-contextual features in ANNs for movie rating prediction", Conference paper, Lernen, Wissen, Daten, Analysen (LWDA 2016) At: Potsdam, Germany.

[23] A.Buyalskaya, H. Ho, K. L. Milkman, "What can machine learning teach us about habit formation? Evidence from exercise and hygiene", research article, Proceedings of the National Academy of Sciences 120(17), Pnas, April 17, 2023.

[24] C.Zhang, J.Vanschoren, A.van Wissen, D.Lakens, B.Ruyter & A. IJsselsteijn, "Theory-based habit modeling for enhancing behavior prediction in behavior change support system", Springer Umuai article, Vol 32, pages 389–415, 31 May 2022.

[25] Maxwell I. Nye, Armando Solar-Lezama, Joshua B. Tenenbaum, Brenden M. Lake, "Learning compositional rules via neural program synthesis", NIPS '20: Proceedings of the 34th International Conference on Neural Information Processing Systems December 2020 Article No.: 909 Pages 10832–10842.