Why Do We Need Neurosymbolic AI to Model Pragmatic Analogies?

Thilini Wijesiriwardene and Amit Sheth, *University of South Carolina, Columbia, SC, 29208, USA*

Valerie L. Shalin, *Wright State University, Dayton, OH, 45435, USA*

Amitava Das, *University of South Carolina, Columbia, SC, 29208, USA*

A hallmark of intelligence is the ability to use a familiar domain to make inferences about a less familiar domain, known as analogical reasoning. In this article, we delve into the performance of large language models (LLMs) in dealing with progressively complex analogies expressed in unstructured text. We discuss analogies at four distinct levels of complexity: lexical, syntactic, semantic, and pragmatic. As the analogies become more complex, they require increasingly extensive, diverse knowledge beyond the textual content, unlikely to be found in the lexical co-occurrence statistics that power LLMs. We discuss neurosymbolic AI techniques that combine statistical and symbolic AI, informing the representation of unstructured text to highlight and augment relevant content, provide abstraction, and guide the mapping process. This maintains the efficiency of LLMs while preserving the ability to explain analogies for pedagogical applications.

The ability to reason analogically, using a familiar domain to make inferences about a new one, is fundamental to human cognitive ability.[1,2] Analogies involve two domains: a familiar source domain and a less familiar target domain. For example, the familiar domain of cooking (source) provides insight into the less familiar domain of photosynthesis (target) by drawing comparisons and finding similarities between both.

The contemporary approach to simulating intelligent behavior with natural language inputs is the large language model (LLM). LLMs combine billions of parameters with self-supervised learning to capture statistical regularities in large data corpora. Given the substantial success of such neurally inspired approaches in satisfactorily addressing tasks such as classification, recommendation, and prediction, AI researchers have identified new ambitious task goals, including abstraction and analogies (https://youtu.be/aeMbLkONLUw). We suggest that LLMs can perform well on analogies at lower levels of an analogy taxonomy. However, what we call *pragmatic analogies* require relatively rare knowledge outside of the text for essential context. We advocate a neurosymbolic approach (https://bit.ly/3KjbE68) informed not only by data but also relevant knowledge (https://bit.ly/DKduality), typically represented using knowledge graphs (KGs) (https://bit.ly/KGOKN).

Here, we first present a taxonomy of analogies, highlighting the demands on knowledge outside of the analogy text. Then, we sketch our neurosymbolic approach, consistent with the broader cognitive science literature cited in regarding the need for external knowledge, an appreciation for the combinatorics of the mapping problem, and the problem of evaluating the necessarily imperfect analogy quality. Less concerned with fidelity to human processing, our own effort focuses on representing unstructured, rich analogies, where external knowledge guides initial and layered representations associated with deep learning mechanisms and supports explanation for pedagogical applications.

Digital Object Identifier 10.1109/MIS.2023.3305862
Date of current version 13 September 2023.

TYPES OF ANALOGIES

Past research on LLMs' abilities to model analogies primarily focused on the simplistic proportional analogies found on achievement and intelligence tests (e.g., "Fingers is to hand as toes is to what?"). Identifying the limitations of LLMs for analogical reasoning requires a principled taxonomy. We present a taxonomy, based on Wijesiriwardene et al.,[3] focused on the complexity and required external information and knowledge (Figure 1). Pragmatic analogies span several sentences (often a paragraph) that elaborate on both the source and target domains; contain multiple concepts or entities related by diverse relationships; contain abstractions (modeled as subgraphs); and require mapping concepts/entities, relationships, and subgraphs between the source and target contextualized by external knowledge and a purpose, often pedagogical.

Lexical Analogies (Proportional Analogies)

The first level in Figure 1 consists of lexical or proportional analogies. Proportional analogies follow the format of a : b :: c : d, where the relationship between a and b and the relationship between c and d remain consistent and can either be explicit or implicit. For instance, "Paris : France :: Rome : ?" requires retrieving the explicit relationship of "capital_of" from the first pair and applying it to the second pair to arrive at "Italy." However, the relationship between "caricature : drawing :: limerick : ?" is unlikely to be prestored but, rather, requires a more challenging construction. Explicit relationships such as "type_of," "capital_of," and "part_of" are still possible for LLMs to tackle because the data fed into these models at training time likely includes these explicit and specific relationships. However, the relationship between caricature and drawing does not fall into any preconstructed hyponymy or hypernym categories.

LLMs model lexical analogies effectively when they can depend on statistical regularities in data that inform the model. However, LLMs still struggle to capture rare and unusual relationships.

Syntactic Analogies

Syntactic analogies focus on the structural and grammatical similarities between phrases and examine word arrangement and grammatical relationships. These analogies are more complex than lexical analogies because they combine several words to form sentences, introducing relatively more complex relationships among the words, such as dependencies. For example, the sentences "A cat meowed at me" and "A dog barked at my brother" can be identified as analogues with similar syntactic and grammatical structures. Current LLMs are well suited for the task, with their capability to capture simple syntactic and grammatical structures likely present in the training data,[4] providing improved scalability compared to rule-based models.

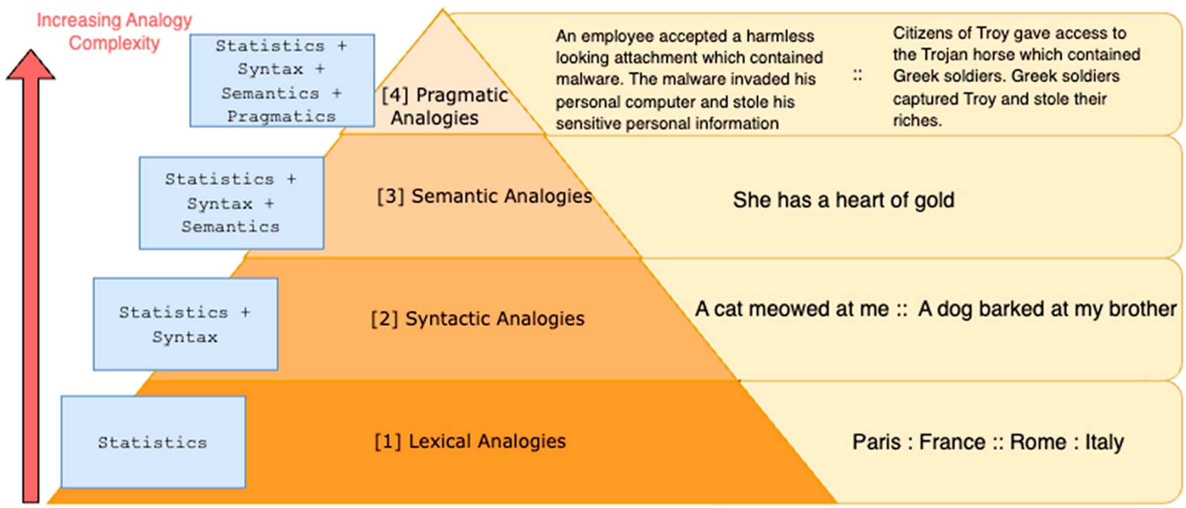

FIGURE 1. Taxonomy of analogies in order of increasing complexity of the associated computation as well as the types and sources of external knowledge necessary for modeling each type. Examples for each level of the taxonomy are on the right, and the information/knowledge needed to model analogies at that particular level are shown on the left.

Semantic Analogies (Metaphors)

Metaphor exemplifies the intermediate level of semantic analogies, largely expressed as a single proposition. Consider "She has a heart of gold," comparing two unrelated domains: a person's anatomical attribute (heart) and the precious metal known for its value and purity (gold). Metaphor elements are distant from each other from a literal perspective, and the domains are blended rather than mapped. Modeling such semantic analogies highlights the challenge of evaluating the understanding without the task-specific framings in levels 1 and 2, classically approached according to coherence, correspondence, and connection to external knowledge beyond the statistical regularities, syntax, simple lexicosemantics,[5] and grammatical structures captured by LLMs. The inability of LLMs to distinguish between meaningful and nonsense assertions casts doubt on the LLMs' ability to address metaphor.[6]

Pragmatic Analogies (Rich Analogies)

Pragmatic, or rich, analogies employ longer narratives spanning multiple sentences, often a paragraph. The initial interpretation requires access to context beyond what is explicit in the text itself, particularly general knowledge of how these domains are structured and function independently, and some process for recovering the abstraction supporting the identification of relevant underlying patterns, principles, or attributes that can be translated and applied from one domain to another. Generic datasets are unlikely to contain lexical co-occurrences across domains, which are the foundation of LLMs.[7]

LLMs encounter difficulties when dealing with complex analogies that require more abstract mappings.[4,8] Recent work using LLMs on a simplified forced-choice task is promising but still fails to reach human-level accuracy and, in any case, lacks explanatory capability and long-term retention.[9] Word order, critical to distinguishing roles, is only partially addressed with standard positional encoding methods.[10] Most critically, the structure necessary for mapping is distributed over lengthy textual description and requires the apprehension of large-window dependencies using very large models and extended training. This approach is impractical for repeated encounters with pragmatic analogies. Complementary multifaceted knowledge mitigates these challenges. In most domains, this knowledge is already curated by humans and made available via more generic KGs, such as DBpedia, ConceptNet, WordNet, and Freebase, as well as more domain-specific KGs, such as Bio2RDF and Greek Mythology KG.

A NEUROSYMBOLIC APPROACH TO MODELING PRAGMATIC ANALOGIES

Our technical approach fuses two complementary AI approaches: data-driven neural networks and knowledge-supported symbolic processing (https://bit.ly/3KjbE68). This fusion combines the efficiency and scope of deep learning using large datasets with the ability to reason using explicit knowledge, usually represented as KGs. We sketch our neurosymbolic approach in two primary sections: analogy representation and mapping. In each section, we identify specific functions that require knowledge, including the extraction and enrichment of concepts in the representation and distinct mapping tasks that are resolved with different computational methods.

Pragmatic Analogy Representation

Starting with unstructured text, we propose the recovery of a mediating graphical structure, which we refer to as an "analogy concept graph" (ACG). The ACG provides two critical functions. First, it captures the purpose of analysis, thereby directing subsequent processing. Second, it supports the recovery of an explanation for pedagogical applications.

Essential analogy concepts from the textual descriptions of both the source and target domains will be incorporated into the recovery of an ACG through a process we identify as "analogy concept extraction," where only the relevant concepts use KGs. A subsequent process, "ACG enrichment," consults KGs and incorporates comprehensive explanations for concepts and relations aiding the mapping process.

Pragmatic Analogy Mapping

We partition computationally intensive mapping into three different problems with three different computational solutions: entity-level mapping (ELM), relational-level mapping (RLM), and subgraph-level mapping (SLM).

ELM

We argue that ELM is relatively less important. "Citizens_of_troy" can be considered as a named entity and "employee" a generic entity, but mapping between these two entities does not give us any insight into the target domain, which is the cybersecurity malware attack. Similarly, comparing "trojan_horse" and "harmless_looking_attachment" at the surface level does not provide any insights into how a malware attack could be similar to the Trojan horse scenario in mythology.

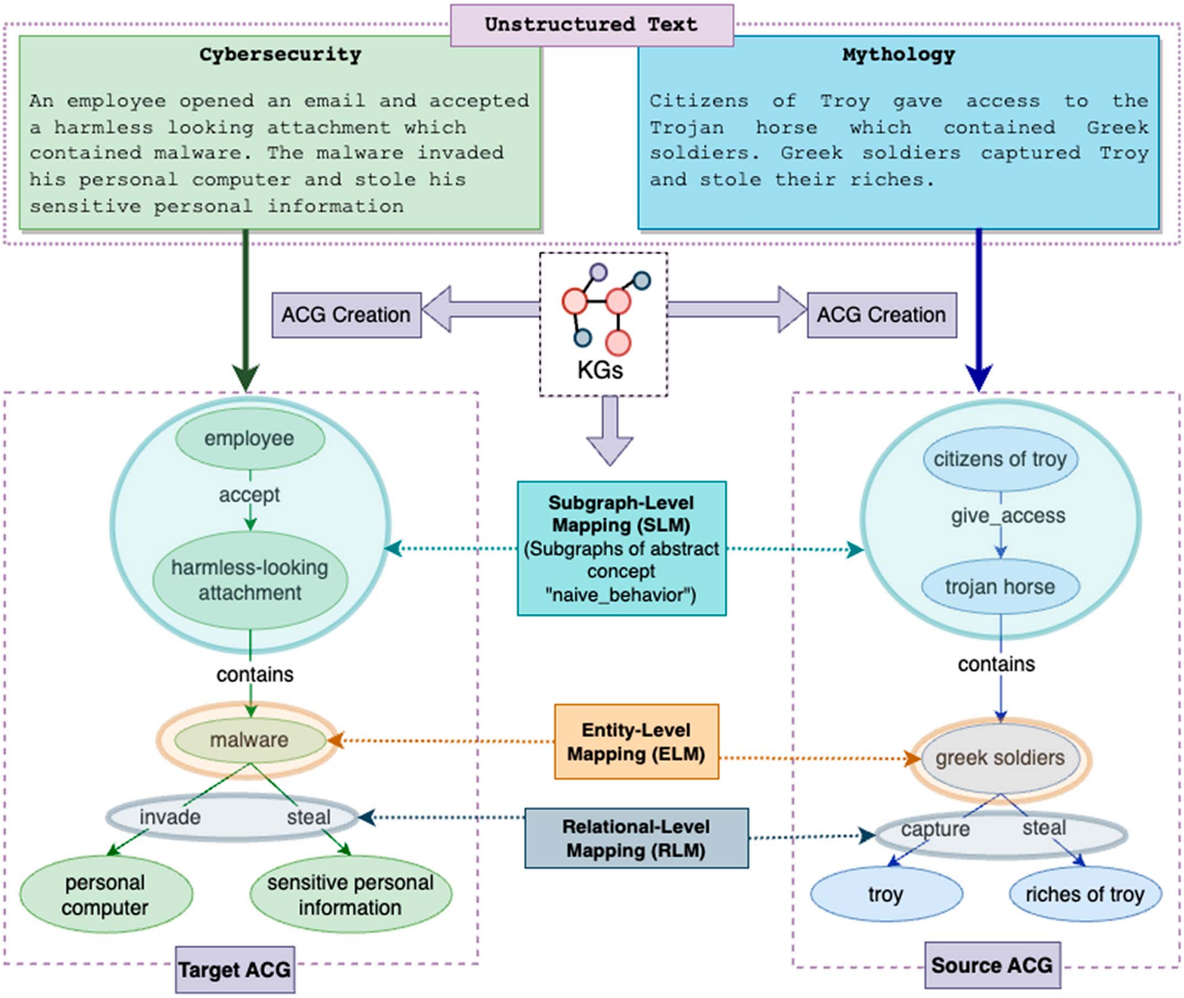

FIGURE 2. Process of modeling pragmatic analogies: the unstructured texts of the target domain (cybersecurity) and the source domain (mythology) are parsed to create source and target representations called analogy concept graphs (ACGs). Knowledge graphs (KGs) are utilized in this process. ELM, RLM, and SLM are performed after creating ACGs.

RLM

Analogical relations are ordered. LLMs do not reliably capture such intricate semantics and pragmatics of language. Relations are similar between the source and target domains (e.g., contain) in some cases and different in others (e.g., accepts vs. give_access). However, relations reflect domain semantics requiring abstraction in mapping. A model needs to identify what each relation means in each domain and how they are ordered. This is particularly challenging for stand-alone LLMs because the semantic and pragmatic distances between the relations and relational structures are higher; mapping the same relational word in the source and target domains lexically focusing on the surface level is easier, but when the relational words are far apart semantically and pragmatically, this strategy can be less effective and erroneous.

SLM

SLM is the highest level of abstraction in analogical mapping. In SLM, the mapping is done on the subgraph level. Subgraphs include a subset of entities and relationships and represent several abstract concepts that hold importance in the source and target domains. For example, in Figure 2, the abstract concept of "naive_behavior" is represented by both <employee, accepts, harmless-looking_software> in the target domain of cybersecurity and <citizens_of_troy, give_access, trojan_horse> in the source domain of Greek mythology and eventually mapped. (The hierarchical concept representation in KGs supports identifying these subgraphs.) To identify that the specific Greek mythology includes the abstract concept of "naive_behavior," the model must utilize already synthesized knowledge present in relevant domain-specific KGs. (In this specific case, to situate the meaning of trojan_horse and citizens_of_troy, the model needs the domain-specific KG on Greek mythologies.) Current LLMs do not capture these types of high-level abstractions well.

Subgraph mappings require approximate graph isomorphism (GI). Various techniques for computing NP-complete GI have emerged over time, from backtracking algorithms to graph-neural-network (GNN)-based methods. GNNs, as structure-driven models, dynamically adjust their network structure to capture intricate dependencies within an input graph. The iterative aggregation process in GNNs ultimately maps the subgraph structure into a vector space, where similarity calculations can be done quite efficiently.

PARTING THOUGHTS

This article highlights different types of analogies based on their complexity and requirements of information and knowledge, guided by a four-level taxonomy. We argued that LLMs generally model simple analogies, like lexical analogies and syntactic analogies, fairly well but that they fall short when modeling semantic and pragmatic analogies. We further explained how a neurosymbolic AI approach can be used to model pragmatic analogies by incorporating content beyond the text, capturing broad, rich, multifaceted knowledge in the form of KGs, and how the mapping is done between source and target ACGs at three distinct levels. Our neurosymbolic approach supports explanation, which is central to pedagogical applications.

Modeling pragmatic analogies presents a formidable challenge for LLMs, as it requires an understanding of context that transcends mere statistics, syntax, and semantics. LLMs must be combined with rich, nuanced, and multifaceted knowledge to acquire pragmatics.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation Award 2133842: "EAGER: Advancing Neuro-symbolic AI With Deep Knowledge-Infused Learning." A more detailed version of this article is available at https://arxiv.org/abs/2308.01936.

REFERENCES

1. K. Holyoak, D. Gentner, and B. Kokinov, "The place of analogy in cognition," in *The Analogical Mind: Perspectives from Cognitive Science*, vol. 119. Cambridge, MA, USA: MIT Press, 2001, pp. 10–13.
2. D. R. Hofstadter, "Analogy as the core of cognition," in *The Analogical Mind: Perspectives from Cognitive Science*. Cambridge, MA, USA: MIT Press, 2001, pp. 499–538.
3. T. Wijesiriwardene et al., "ANALOGICAL – A novel benchmark for long text analogy evaluation in large language models," in *Findings of the Association for Computational Linguistics*. Toronto, ON, Canada: Association for Computational Linguistics, 2023, pp. 3534–3549.
4. H. Li, "Language models: Past, present, and future," *Commun. ACM*, vol. 65, no. 7, pp. 56–63, 2022, doi: 10.1145/3490443.
5. I. Vulić, E. M. Ponti, R. Litschko, G. Glavaš, and A. Korhonen, "Probing pretrained language models for lexical semantics," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2020, pp. 7222–7240, doi: 10.18653/v1/2020.emnlp-main.586.
6. Y. Chang et al., "A survey on evaluation of large language models," 2023, *arXiv:2307.03109*.
7. E. Pavlick, "Symbols and grounding in large language models," *Philos. Trans. Roy. Soc. A*, vol. 381, no. 2251, Jul. 2023, Art. no. 20220041, doi: 10.1098/rsta.2022.0041.
8. A. Ushio, L. Espinosa Anke, S. Schockaert, and J. Camacho-Collados, "BERT is to NLP what AlexNet is to CV: Can pre-trained language models identify analogies?" in *Proc. 59th Annu. Meeting Assoc. Comput. Linguistics*, 2021, vol. 1, pp. 3609–3624.
9. T. Webb, K. J. Holyoak, and H. Lu, "Emergent analogical reasoning in large language models," *Nature Human Behav.*, early access, Jul. 2023, doi: 10.1038/s41562-023-01659-w.
10. M. Abdou, V. Ravishankar, A. Kulmizev, and A. Søgaard, "Word order does matter and shuffled language models know it," in *Proc. 60th Annu. Meeting Assoc. Comput. Linguistics*, 2022, vol. 1, pp. 6907–6919.

THILINI WIJESIRIWARDENE is a Ph.D. student at the AI Institute of South Carolina (AIISC), University of South Carolina, Columbia, SC, 29208, USA. Contact her at thilini@sc.edu.

AMIT SHETH is the founding director of the AIISC, an NCR chair, and a professor of computer science and engineering at the University of South Carolina, Columbia, SC, 29208, USA. Contact him at http://aiisc.ai/amit.

VALERIE L. SHALIN is a professor of psychology at Wright State University, Dayton, OH, 45435, USA, specializing in applications of technology to workplace activities. Contact her at valerie.shalin@wright.edu.

AMITAVA DAS is a research associate professor at the AIISC, University of South Carolina, Columbia, SC, 29208, USA. Contact him at amitava@mailbox.sc.edu.