AGen: Personalized Analogy Generation with Large Language Model

Shutong Wu(✉), Hecong Wang, and Zhen Bai

University of Rochester, Rochester, NY 14620, USA
{swu85,hwang99}@ur.rochester.edu, zhen.bai@rochester.edu

Abstract. Analogies facilitate understanding across domains, enabling individuals to navigate unfamiliar concepts using what they already know. Yet, crafting effective analogies requires extensive knowledge of both the source and target domains, making it a challenging task for educators—particularly when aiming to support novice learners. Recent advancements in Large Language Models (LLMs) have demonstrated their ability to generate analogies to explain scientific concepts. However, analogy in education must align with students' prior knowledge and cognitive resources, which shape how analogies are perceived. To address this challenge, we examine the theory and practices of using analogies in education and introduce a system that generates personalized analogies to facilitate learning. Preliminary evaluations show promising outcomes, and we anticipate deeper insights from future human-subject studies.

Keywords: Analogical Learning · Personalization · Large Language Model

1 Introduction

Analogies are central and ubiquitous in human cognition, as people often explore new situations by drawing parallels to familiar ones [12]. In the educational domain, analogies are powerful tools for facilitating students' understanding of unfamiliar concepts and increasing motivation by introducing a familiar source domain [15]. Incorporating analogies into science texts and teaching has enhanced student comprehension, particularly by elucidating the causal and relational structures of concepts [19].

However, manually creating effective analogies is a cognitively demanding task, especially for novices. Research indicates that generating high-quality analogies requires effort, expertise, and a nuanced understanding of the subject matters of both the source and target domains [14,25]. Moreover, the effectiveness of analogies in learning is in line with the "Expertise Reversal Effect" framework [21], which highlights that with growing expertise and cognitive capacity, learners require different types of analogies to support their understanding best. For example, novice learners often benefit from surface-similar analogies, while

A. I. Cristea et al. (Eds.): AIED 2025, CCIS 2591, pp. 242–249, 2025.
https://doi.org/10.1007/978-3-031-99264-3_30

experts prefer structurally similar analogies that highlight deeper causal relationships [12,14].

Recent advancements in Generative AI and Large Language Models (LLMs) [24] have enabled the automatic generation of natural language analogies in settings such as education and creative writing [3,7,22]. These systems produce novel and meaningful analogies, but there is scarce research that takes into account the different ways audiences with varying backgrounds perceive and process analogies. One exception is the recent work aimed at generating personalized analogies for AI concepts [4]. While this work demonstrates the potential of tailoring analogies to user profiles, it primarily focuses on high-level prompting strategies and lacks a structured pipeline for extracting and adapting conceptual knowledge.

To address these challenges, we propose a novel application for generating personalized analogies based on users' prior experience and information such as learner age and personal interests. Our initial evaluation shows promising results and demonstrates the potential to improve analogy generation and learning outcomes. Future work includes comprehensive user studies to assess generation quality and practical usability.

2 Related Work

2.1 Analogy and Analogical Learning

Analogy is a fundamental cognitive tool that allows humans to learn across domains by leveraging prior knowledge [18]. Its significance has been recognized since ancient times, tracing back to Aristotle's theoretical reflections on analogy [1], and continues to influence modern understanding. Analogies are effective in various contexts, including science education [6,13], children's learning [16,23], and creativity-focused activities [9,20]. For novices, analogies are critical in enhancing problem-solving capabilities by facilitating solution transfer [1] and reducing uncertainty during problem-solving tasks [5].

Despite its potential, creating high-quality analogies often demands expertise [26], making it difficult for novices to perform meaningful inferences in unfamiliar domains without expert assistance [15]. Moreover, the effectiveness of an analogy can vary significantly across different age groups and audiences. For instance, young children are more likely to rely on object-level matches and surface similarities, whereas adults tend to focus on relational similarities between domains [10]. This *relational shift* underscores the importance of tailoring analogies to the target audience's cognitive capabilities and prior knowledge. Similarly, individuals with varied cognitive abilities [17] and relational knowledge [8] are likely to interpret and accept analogies differently based on their prior knowledge and cognitive resources.

2.2 Analogy Generation with Large Language Model

LLMs, such as GPT [24], have opened up many possibilities in various domains, including analogy generation [3] and analogical reasoning [27]. Previous studies

		Prompt	Example Inputs	Example Analogy (target concept, analogy)
Knowledge Generation	Knowledge Sampling	You are an expert on the subject of [Concept]. Please create a comprehensive reading material for subject to be used as teaching materials for [Age] [Interest] [Other]. The knowledge and materials must be fit to the group's prior knowledge and cognitive level.	[Concept] Solar System [Age] [Interest] [Other] 9 years old at elementary school; drawing and comic book; good at art...	Imagine the **Solar System** as a vast comic book series, where each **planet** is a unique character with its own story and personality, and plays a specific part in the cosmic story. The **Sun**, as the central figure, is the superhero that holds everything together. The **rocky planets**, like Mercury and Mars, are the adventurous sidekicks... The **gas giants**, such as Jupiter and Saturn, are the powerful allies... The **ice giants**, Uranus and Neptune, add mystery and intrigue... Finally, the **dwarf planets** and the Kuiper Belt are the hidden gems
	Chain-of-Note Analysis	Analyze the knowledge clause by clause; For each clause, extract the explicit and implicit knowledge; For each extracted knowledge, phrase the knowledge as a simple subject-verb-object phrase. Here is the knowledge [Knowledge], and ...."	[Knowledge] • What is Solar System: ...... • The Members of Our Solar System: ...... • Fun Fact: Dwarf Planets...	
	Taxonomy Generation	Create a taxonomy of concepts and relations for a knowledge graph about [Concept] based on [Analyzed Knowledge]. A taxonomy consists of a set of concepts and a set of relations. For example ...	[Analyzed Knowledge] • The Solar System is like a big family in space... • At the center of this family is the Sun, a giant ball of hot, flowing gas...	
Analogy Generation	Analogy Generation	You will now teach about the concept [Concept] using analogies for the selected learner group [Age] [Interest] [Other]. The concept is accompanied by a combined knowledge summary [Analyzed Knowledge] and its taxonomy [Taxonomy]. You will follow principles below for teaching in an analogical approach: .... Given the principles, generate [n] detailed analogy and explanation tailored for such group. Let us think step-by-step the following aspects: ...	[Taxonomy] • Concepts: Solar System, Sun, Mercury, Venus, ... • Relations: is part of, orbits, is known for, contains, is equivalent to, ...	
	Analogy Filtering	Given the [n] analogies, determine if there is [m] analogies that are more suited to the audience [Age] [Interest] [Other]. Rank and filter [m] analogies based on these principles: ...		

Fig. 1. Prompt Design Overview.

have demonstrated that LLMs can be effectively prompted to generate meaningful analogies and provide corresponding explanations when given specific contexts [3,9,22]. Researchers have extended LLM-based analogy generation beyond proportional analogies(A::B to C::D) by incorporating frameworks like Structure Mapping Theory [11] to generate high-level analogies [29]. While recent work has begun exploring personalization in analogy generation [4], it focuses primarily on modifying prompt templates for specific audiences and lacks a structured mechanism for representing domain knowledge or filtering analogies for educational applicability.

3 System Implementation

To address the limitations of existing work and achieve a more comprehensive and structured analogy generation, we introduce AGen, an analogy-generation tool that, given a source concept and an audience profile, creates tailored, structured analogies. We investigated previously established frameworks and principles in analogical learning [12,15], and incorporated these into our system to generate more structured and comprehensive analogies. Our framework consists of two stages, which are *Knowledge Generation and Adaptation* and *Analogy Generation and Selection*. We currently use GPT-4o as the underlying LLM, given its high performance on various tasks. Our step-by-step prompt design can be found in Fig. 1.

3.1 Knowledge Generation and Adaptation

To achieve effective personalization, we begin by collecting the audience profile across three dimensions: *learner age*, *background knowledge*, and *learner interest*, which is similar to the prompting framework in [4]. However, our system addresses a key limitation of prior approaches—the lack of distinction between conceptual relevance and motivational appeal—by treating these dimensions differently to solve distinct personalization challenges and achieve structural generation. Specifically, we use *learner age* and *background knowledge* as conditioning inputs during the knowledge generation and adaptation stage. These factors shape a learner's cognitive readiness, domain familiarity, and ability to interpret relational structures—making them critical for generating analogies that are conceptually appropriate and structurally sound. In contrast, we use *learner interest* during the analogy generation and selection stage to filter and prioritize analogies based on motivational relevance. While interests can enhance engagement and retention, they should not interfere with the analogy's conceptual integrity.

Specifically, we implement three novel components—**Knowledge Sampling**, **Chain-of-Note Analysis**, and **Taxonomy Generation**—to shape analogy generation based on what the learner is likely to know and understand. We first sample the knowledge to develop a comprehensive knowledge of the source concept. We prompt GPT to generate background knowledge of a given science concept N times (N = 10) and sample them to form a combined background information text. We then analyze the background knowledge to extract explicit and implicit knowledge as simple subject-verb-object phrases using the Chain-of-Note technique [28]. The analyzed knowledge is then aggregated into a taxonomy of the given concept, which consists of a set of filtered concepts and relations; in our early experiments, we found that a taxonomy is helpful in preventing LLMs from generating redundant or irrelevant information that could result in noise in the subsequent generation stages. We then adapt the generated taxonomy to the audience profile to generate a conditioned knowledge base that reflects the audience's prior knowledge of the source concept.

3.2 Analogy Generation and Selection

To generate the personalized analogy, we first prompt LLM to generate K analogies (K = 10) based on the source concept, audience profile, and taxonomy. The LLM is prompted to create the analogy following the principles of analogical teaching [15], which includes principles such as using well-understood analogies and highlighting shared causal structure among analogies [15].

The above generation will still face limitations such as misinterpretations of certain sub-concepts or factually incorrect analogies. We further design the system to automatically filter the analogies by embedding the evaluation metrics in [2] and [12] into the prompts. Using the previously obtained taxonomy and additional user input, including *learner interest*, AGen filters the analogies only to output those easily understandable to the selected audience. An example of two tailored analogies towards different audiences is shown in Fig. 2.

4 Evaluation

We conducted a preliminary technical evaluation to test the personalization quality of the generated analogies in addressing intended audience groups. In the future, we will evaluate this tool's learning effectiveness and generation quality.

To test our tool, we selected a set of STEM science concepts from **Next Generation Science Standards (NGSS)**[1] by randomly sampling a group

[1] https://www.nextgenscience.org.

Example Analogy (target concept, analogy)		
6 years old, likes toys and cartoons like Peppa Pig	9 years old, likes drawing and comic book, good at art	16 years old, likes science and physics
Imagine the Solar System as a big, colorful carousel at the fair. The Sun is the tall pole in the middle, holding everything together, and each planet is like a fun horse that goes around and around. Mercury is the horse closet to the pole, and Neptune is way out on the edge. Each horse has a different color and style, just like each planet is special – Mars is red, Saturn has rings, and Uranus spins a little sideways! All the planets keep spinning around the Sun in a big circle, just like carousel horses moving together in a magical ride.	Imagine the Solar System as a vast comic book series, where each planet is a unique character with its own story and personality, and plays a specific part in the cosmic story. The Sun, as the central figure, is the superhero that holds everything together. The rocky planets, like Mercury and Mars, are the adventurous sidekicks... The gas giants, such as Jupiter and Saturn, are the powerful allies... The ice giants, Uranus and Neptune, add mystery and intrigue... Finally, the dwarf planets and the Kuiper Belt are the hidden gems	Imagine the Solar System as a giant atom. At its center, the Sun acts like the nucleus - a massive body that holds the system together.... Just as the nucleus is the dense core of an atom, the sun is the central powerhouse of the Solar System. Surrounding the Sun are the planets, akin to electrons orbiting the nucleus. These planets follow specific paths just like orbits, maintaining a delicate balance between the forces pulling them inward and those propelling them outward.

Fig. 2. AGen's Personalization Analogy Examples.

of concepts from different fields. This ensures that our experiments reflect the education curriculum. For further experiments, we will include more concepts.

4.1 Personalization Evaluation

We evaluate the system's ability to tailor analogies to audiences with different backgrounds and levels of expertise. We evaluated the system's personalization ability by classifying each analogy to the most likely audience profile. Specifically, we computed cosine similarities between the embeddings of each generated analogy and the embeddings of various audience profiles using the *text-embedding-3-small* model. The audience profile with the highest similarity score was selected as the predicted profile for each analogy. This approach allowed us to quantitatively assess how well the system aligns analogies with the intended audience characteristics.

The personalized evaluation examines the system's ability to tailor analogies based on audience profiles, specifically their ages and interests. Figure 3 provides confusion matrices summarizing system performance across these dimensions. Key findings are as follows: (1) **Strong alignment with user interests**: As shown in the right panel of Fig. 3, the system demonstrates high accuracy in predicting analogies aligned with users' declared interests. For instance, analogies tailored for individuals interested in comic books or sports are predominantly categorized correctly, with 14 accurate predictions in these categories. This strong alignment enhances the analogies' relevance, engagement, and potential impact on comprehension. However, minor misclassifications, such as sports analogies attributed to the "Animals" category, highlight areas for further refinement. (2) **Adherence to an audience age "upper bound"**: The left panel of Fig. 3 reveals that analogies respect an implicit "upper bound" for complexity based on the audience's age. Preschool-level analogies are correctly simplified and context-specific, with the system achieving 12 accurate classifications for this group. Conversely, analogies for College Undergraduates exhibit greater breadth and complexity. Interestingly, while analogies designed for older audiences could, in principle, be simplified and adapted for younger audiences, the reverse does not hold, reflecting the system's capacity to generate age-appropriate outputs.

The evaluation highlights the system's strengths in aligning analogies with users' interests and age-appropriate complexity, as well as areas where further adjustments to mitigate biases can enhance performance.

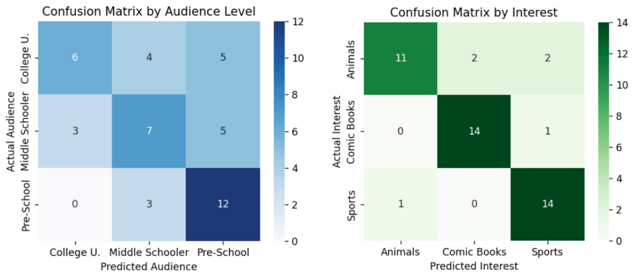

Fig. 3. Confusion Matrix based on two different settings.

5 Conclusion and Future Work

While AGen demonstrates strong potential, it also faces some limitations as an early-stage work. The reliance on LLMs for generating knowledge and taxonomy can result in incomplete or biased outputs due to inherent constraints in the model's pre-trained data, despite our efforts to condition the knowledge prospectively during the generation step. Moreover, evaluating the quality of analogies is inherently subjective, as it depends on users' prior knowledge and individual backgrounds. To address these challenges, future work will expand the scope of user studies and source concept domains to incorporate a more diverse group of participants, such as K-12 students and education experts.

This paper introduces AGen, an innovative analogy generation system that leverages advanced LLMs to create tailored, structured analogies for diverse educational contexts. By integrating knowledge adaptation and theory-driven generation methods, AGen represents a significant advancement in utilizing analogies to enhance learning.

Preliminary findings from our automatic evaluations show that our tool can produce personalized analogies with quality comparable to the existing baseline systems, indicating promising applications in improving concept comprehension and fostering user engagement. Future work will focus on conducting extensive user studies across varied demographics, exploring alternative and objective evaluation metrics, and integrating additional LLMs to enhance the diversity and robustness of analogy generation. These efforts aim further to establish AGen as a valuable tool for educational innovation.

Acknowledgments. Thanks to Xiaofei Zhou for significantly assisting the research.

References

1. Bartha, P.: Analogy and Analogical Reasoning. In: Zalta, E.N., Nodelman, U. (eds.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Fall 2024 edn. (2024)
2. Bhavya, B., Palaguachi, C., Zhou, Y., Bhat, S., Zhai, C.: Long-form analogy evaluation challenge. In: Mille, S., Clinciu, M.A. (eds.) Proceedings of the 17th International Natural Language Generation Conference: Generation Challenges, pp. 1–16. Association for Computational Linguistics, Tokyo (2024). https://aclanthology.org/2024.inlg-genchal.1/
3. Bhavya, B., Xiong, J., Zhai, C.: Analogy generation by prompting large language models: a case study of instructgpt (2022). https://arxiv.org/abs/2210.04186
4. CAO, C.C., Zoe, F., Lydia, Y.C., Jionghao, L., Ruizhe, L.: Llm-generated personalized analogies to foster ai literacy in adult novices. In: International Conference on Computers in Education (2024)
5. Chan, J., Paletz, S.B.F., Schunn, C.D.: Analogy as a strategy for supporting complex problem solving under uncertainty. Mem. Cogn. **40**(8), 1352–1365 (2012). https://doi.org/10.3758/s13421-012-0227-z
6. Coll, R.K., France, B., and, I.T.: The role of models/and analogies in science education: implications from research. Int. J. Sci. Educ. **27**(2), 183–198 (2005). https://doi.org/10.1080/0950069042000276712
7. Czinczoll, T., Yannakoudakis, H., Mishra, P., Shutova, E.: Scientific and creative analogies in pretrained language models (2022). https://arxiv.org/abs/2211.15268
8. D Gentner, M.R.: Language and the career of similarity. Technical Report 533, University of Illinois (1991)
9. Ding, Z., Srinivasan, A., Macneil, S., Chan, J.: Fluid transformers and creative analogies: exploring large language models' capacity for augmenting cross-domain analogical creativity. In: Proceedings of the 15th Conference on Creativity and Cognition, C&C '23, pp. 489–505. Association for Computing Machinery, New York (2023). https://doi.org/10.1145/3591196.3593516
10. Gentner, D.: Metaphor as structure mapping: the relational shift. Child Dev. **59**(1), 47–59 (1988). http://www.jstor.org/stable/1130388
11. Gentner, D., Markman, A.B.: Structure mapping in analogy and similarity. Am. Psychol. **52**(1), 45 (1997)
12. Gentner, D., Smith, L.A.: Analogical learning and reasoning. In: The Oxford Handbook of Cognitive Psychology, pp. 668–681. Oxford library of psychology, Oxford University Press, New York (2013). https://doi.org/10.1093/oxfordhb/9780195376746.001.0001
13. Gick, M., Holyoak, K.: Analogical problem solving. Cogn. Psychol. **12**, 306–355 (1980). https://doi.org/10.1016/0010-0285(80)90013-4
14. Goldwater, M.B., Gentner, D., LaDue, N.D., Libarkin, J.C.: Analogy generation in science experts and novices. Cogn. Sci. **45**(9), e13036 (2021)

15. Gray, M.E., Holyoak, K.J.: Teaching by analogy: from theory to practice. Mind Brain Educ. **15**(3), 250–263 (2021). https://doi.org/10.1111/mbe.12288. https://onlinelibrary.wiley.com/doi/abs/10.1111/mbe.12288
16. Haglund, J., Jeppsson, F., Andersson, J.: Young children's analogical reasoning in science domains. Sci. Educ. **96**(4), 725–756 (2012). https://doi.org/10.1002/sce.21009. https://onlinelibrary.wiley.com/doi/abs/10.1002/sce.21009
17. Halford, G.S.: Children's Understanding: The Development of Mental Models. Psychology Press (2014)
18. Hofstadter, D.R.: Epilogue: analogy as the core of cognition. In: The Analogical Mind: Perspectives from Cognitive Science. The MIT Press (2001). https://doi.org/10.7551/mitpress/1251.003.0020
19. Jaeger, A.J., Taylor, A.R., Wiley, J.: When, and for whom, analogies help: the role of spatial skills and interleaved presentation. J. Educ. Psychol. **108**(8), 1121 (2016)
20. Jiayang, C., et al.: Storyanalogy: deriving story-level analogies from large language models to unlock analogical understanding (2023). https://arxiv.org/abs/2310.12874
21. Kalyuga, S.: Expertise reversal effect and its implications for learner-tailored instruction. Educ. Psychol. Rev. **19**, 509–539 (2007)
22. Kim, J., Suh, S., Chilton, L.B., Xia, H.: Metaphorian: leveraging large language models to support extended metaphor creation for science writing. In: Proceedings of the 2023 ACM Designing Interactive Systems Conference, DIS '23, pp. 115–135. Association for Computing Machinery, New York (2023). https://doi.org/10.1145/3563657.3595996
23. Newton, D.P., Newton, L.D.: Using analogy to help young children understand. Educ. Stud. **21**(3), 379–393 (1995). https://doi.org/10.1080/0305569950210305
24. Ouyang, L., et al.: Training language models to follow instructions with human feedback (2022). https://arxiv.org/abs/2203.02155
25. Richland, L.E., Simms, N.: Analogy, higher order thinking, and education. Wiley Interdisc. Rev. Cogn. Sci. **6**(2), 177–192 (2015)
26. Turney, P.D.: The latent relation mapping engine: algorithm and experiments. J. Artif. Intell. Res. **33**, 615–655 (2008). https://doi.org/10.1613/jair.2693
27. Yasunaga, M., et al.: Large language models as analogical reasoners (2024). https://arxiv.org/abs/2310.01714
28. Yu, W., Zhang, H., Pan, X., Ma, K., Wang, H., Yu, D.: Chain-of-note: enhancing robustness in retrieval-augmented language models (2024). https://arxiv.org/abs/2311.09210
29. Yuan, S., Chen, J., Ge, X., Xiao, Y., Yang, D.: Beneath surface similarity: large language models make reasonable scientific analogies after structure abduction (2023). https://arxiv.org/abs/2305.12660