Proceedings of the International Conference on AI Transformation

2024

CORVINUS UNIVERSITY of BUDAPEST

香港中文大學深圳研究院
Shenzhen Research Institute
The Chinese University of Hong Kong

Partners

HUN-REN
Hungarian Research Network

KPMG

UNIVERSITY of DEBRECEN

MMA·MMKi

iFLYTEK

WeSpeakAI
SMART TOGETHER

SCITOP

Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal

AZ NKFI ALAPBÓL MEGVALÓSULÓ PROJEKT

Proceedings of the International Conference on AI Transformation

2024

香港中文大學深圳研究院
Shenzhen Research Institute
The Chinese University of Hong Kong

Budapest, 2024

Proceedings of the International Conference on AI Transformation

Editors: Prof. Yeung Yam & Prof. Péter Baranyi

ISBN 978-963-503-956-2

Publisher: Corvinus University of Budapest

Budapest, 2024

Table of Contents

AI-based simulation framework of supply chain resilience

Szabina Fodor, Andrea Gelei and Annamária Ónodi
Corvinus University of Budapest
E-mail: szabina.fodor@uni-corvinus.hu; andrea.gelei@uni-corvinus.hu;
annamaria.onodi@uni-corvinus.hu

Abstract: *Recent developments put supply chain resiliency into focus again. Its dynamics especially need a more profound understanding. The paper distinguishes two types of dynamics: (1) among the different operational performance measures and how they affect overall performance (e.g. EBIDTA). (2) in the counteraction to overcome the Ripple effect (e.g. short-term production planning). Based on dynamic control theory, we propose applying Deep Reinforcement Learning, a machine learning algorithm with excellent potential for tackling difficulties stemming from complex dynamic decision situations. The objective is to develop a conceptual framework for assessing the resilience of supply chains under unpredictable conditions. To achieve this, reinforcement-based simulations are conducted to analyse the performance of our framework in various scenarios. This interdisciplinary approach is a cutting-edge methodology for constructing a dependable and robust supply chain that can be theoretically implemented in other contexts.*

Keywords: *supply chain resilience; deep reinforcement learning; simulations*

1 Review of the literature and proposal of a model

Decades of relative political and economic stability have created global chains with a strong focus on efficiency. However, these chains have faced severe challenges in recent years. Resilient supply chain management is not new, but its importance has grown. Resiliency is widely interpreted as the capacity of a system to recover its critical functions following a disruptive event and restore its performance, which is captured by the time needed for this recovery [1]. However, how performance is understood is up for debate. However, the literature lacks consensus on appropriate resilience metrics; thus, it uses proxies instead of explicit metrics [2]. Managing supply chain disruptions has taught us that measuring performance by any single measure, especially operational ones, has limitations [3]. The lack of understanding of dynamism among different metrics limits existing research on supply chain resiliency. Thus, we need a more complex approach to modelling disruption- generated performance consequences and how it effects overall performance [2]. We propose a dynamic control theory-based supply chain simulation model that can handle two dynamics after a supply network disruption. Starting with performance outcomes. Second, modeling production scheduling dynamics.

2 Methodology

Such a simulation model incorporates two aligned optimisation issues: (1) the problem of production scheduling of different products and (2) the optimisation of the overall business performance (e.g. EBIDTA). Such a complexity is not easy to simulate, even in its simplified form. Reinforcement Learning (RL) is a machine learning algorithm that shows excellent

potential in tackling difficulties of complex dynamism. It determines how to interact with the environment to maximise the reward. As a result, it can be used as an adaptive controller for complex systems. A controller, often called an RL agent, aims to acquire the optimal control actions that maximise the system's long-term objectives across all of the dynamic system's potential states [4]. The RL paradigm demonstrates significant efficacy when implemented in the adaptive management of intricate and dynamic systems, such as supply chains [5]. The complexity of the supply chain environment can exceed the capabilities of the traditional RL technique. This is where Deep Reinforcement Learning (DRL) is applied. The term 'deep' indicates using an artificial neural network to predict potential sequences of actions and link them to long-term rewards, hence enhancing the controllability of the solution space. The agents utilising DRL can generalise the value of states that have not been previously seen or encountered during the training phase.

3 Closing remark

The paper focuses on dynamism in complex supply chains in the context of supply side disruption. To model the complex interplay of hard-to-predict performance implications generated by the Ripple effect and the counteraction made in response, it proposes a novel DRL-based simulation with a probabilistic framework where outcomes are partly random and partly under the decision-makers control by embedding artificial intelligence components into the simulations provides adaptive behaviour.

References

[1] Simchi-Levi, D. and E. Simchi-Levi, We need a stress test for critical supply chains. Harvard Business Review, 2020. 28.

[2] Golan, M.S., L.H. Jernegan, and I. Linkov, Trends and applications of resilience analytics in supply chain modeling: systematic literature review in the context of the COVID-19 pandemic. Environment Systems and Decisions, 2020. 40(2): p. 222-243.

[3] Olivares-Aguila, J. and W. ElMaraghy, System dynamics modelling for supply chain disruptions. International Journal of Production Research, 2021. 59(6): p. 1757-1775.

[4] Boute, R.N., et al., Deep reinforcement learning for inventory control: A roadmap. European Journal of Operational Research, 2022. 298(2): p. 401-412.

[5] Kegenbekov, Z. and I. Jackson, Adaptive supply chain: Demand–supply synchronization using deep reinforcement learning. Algorithms, 2021. 14(8): p. 240.

Ensuring Safety and Reliability in the Age of LLM Agents

Levente Szabados[1,2]
[1] Frankfurt School of Finance and Management
[2] University of Sopron
E-mail: l.szabados@int.fs.de

Abstract: *As Large Language Models (LLMs) increasingly permeate critical sectors like customer service, their reliability and safety have come under scrutiny. Despite their burgeoning deployment, notable failures have made headlines, underscoring the urgent need for rigorous evaluation and risk management frameworks. This paper delves into the unique challenges of ensuring the safety and reliability of LLMs, particularly in Retrieval Augmented Generation scenarios where traditional evaluation metrics falter. Through an empirical survey, we assess the risk consciousness among businesses regarding AI deployment and investigate methodologies for measuring LLM "goodness". We present practical use cases for safety evaluation and propose actionable, repeatable methodologies to enhance LLM reliability. Additionally, we offer insights into the future evaluation of LLM-based autonomous agents, contributing to the broader discourse on AI safety and effectiveness. Our work aims to equip stakeholders with the tools to navigate the complexities of integrating LLMs into business-critical operations, ensuring their beneficial and secure application.*

Keywords: *Large Language Models; Safety and Reliability; Retrieval Augmented Generation; AI Deployment*

1 Introduction

The adoption of Large Language Models (LLMs) in essential business operations is rapidly increasing. Their application ranges from enhancing customer service to streamlining finance operations, marking a transformative era in business technol- ogy. However, this transformation is not without its challenges, as demonstrated by the Air Canada chatbot incident[1]. This paper explores the critical need for safety and reliability in LLM deployments, with a special focus on Retrieval Augmented Generation (RAG) models, which add an additional layer of complexity to their evaluation.

2 Methodology

To understand the landscape of risk awareness and evaluation methodologies among businesses deploying AI, we conducted a comprehensive empirical survey. This survey aimed to gauge the level of risk consciousness in businesses regarding AI deployment and to identify the methodologies currently in use for assessing the efficacy and safety of LLM solutions. Additionally, we delve into the RAGAS[2] automated evaluation method and its successors, analyzing their practical applica- tion, strengths, and weaknesses in a business context.

3 Results and Discussion

The survey results reveal a broad spectrum of risk awareness among businesses, with significant discrepancies in the adoption of robust evaluation methodologies for LLM applications. Our examination of the RAGAS method and its successors uncovers their potential to offer detailed insights into LLM capabilities, espe- cially in complex scenarios involving RAG models. This analysis underscores the importance of adopting sophisticated evaluation tools to enhance the safety and reliability of AI technologies in business operations.

4 Conclusion and Future Work

The deployment of LLMs in critical business sectors necessitates a concerted effort to ensure their safety and reliability. Through our empirical survey and analysis of the RAGAS evaluation method, we highlight the need for sophisticated evaluation methodologies. Future research will focus on refining these methodologies and exploring the development of industry-specific standards to facilitate the safe and effective use of autonomous LLM agents.

References

[1] C. News, "Air Canada's chatbot gave a B.C. man the wrong information, now the airline has to pay for the mistake." https://bc.ctvnews.ca/ – 2 – 1st International Conference on Digital & Cognitive Corporate Reality Vol. XX, No. XX, 2024 air-canada-s-chatbot-gave-a-b-c-man-the-wrong-information-now-the-airline-has-to-pay-for-the-mistake-1. 6769454, 2023. Accessed: 2024-02-26.

[2] A. Name, "RAGAS: Automated Evaluation Method for Retrieval Augmented Generation Models." https://arxiv.org/abs/2309.15217, 2023. Ac- cessed: 2024-02-26.

Towards developing a framework for automated accessibility evaluation of web content from expert perspectives

Jinat Ara[1], Cecilia Sik-Lanyi[1,2]
[1]Department of Electrical Engineering and Information Systems, University of Pannonia
[2]Hungarian Research Network
E-mail: jinat.ara@virt.uni-pannon.hu

Abstract: *In this paper, we proposed a web content accessibility evaluation framework considering fifteen evaluation criteria from expert perspectives. The proposed framework performs through (a) identifying evaluation criteria from expert perspectives, (b) evaluating selected criteria by executing web accessibility evaluation algorithms incorporating different AI methods, and (c) validating the results through experimental analysis and user-centric study. The proposed method is dynamic in nature and can be applied to different platforms to evaluate multiple web pages.*

Keywords: *web accessibility, algorithmic evaluation, automated evaluation*

1 Introduction

WebAIM reported that in 2023, across the world, 96.3% of webpages do not ensure full accessibility as it is quite difficult to offer a completely accessible platform that requires great effort and careful observation [1]. With this in mind, many studies addressed the importance of accessibility needs that should be focused on in the development stage for providing complete support of accessibility criteria [2]. Regarding this issue, several techniques have been introduced to evaluate webpage accessibility considering several criteria including Web Content Accessibility Guideline (WCAG). However, studies reported that WCAG does not cover every aspect that may cause improper evaluation results [3]. Also, due to a lack of advanced techniques, the evaluation report might not be effective for the community. The main aim of this paper is to contribute to the accessibility perspective of digital platforms by proposing an automated web accessibility evaluation framework incorporating several AI techniques to determine webpage accessibility considering expert perspectives. From the expert study, we determined some additional criteria that might be valuable to incorporate into the evaluation process. Based on their feedback, we identified 15 criteria such as arbitrary information {sever status; webpage loading time; and webpage length}, and the content information {paragraph length; the ratio of Hyperlinks; webpage default language; user information; CAPTCHA; multiple language options; image ratio; text font family; text font size; text pattern; content type; audio/video content ratio}. The proposed system is performed by incorporating AI techniques, thus it is dynamic, and it can be integrated or implemented for any webpage evaluation.

2 Methodology

The proposed framework has four distinctive layers as shown in Figure 1. The first layer is responsible for data initialization which is performed by accessing the HTML code of the tested

webpage via the URL of the page. We used sublime text editor as a development framework and Python programming language to write the script. For HTML code access, we used an HTML parser which parses the HTML code and facilitates the data extraction process. The second layer is responsible for data extraction which performs the extracting process of all the necessary data or information from the HTML tree view or HTML source code of the tested website. The third layer is for extracted data evaluation through algorithmic observation which incorporates three different algorithms to evaluate every selected object included in this evaluation process. The proposed algorithms evaluate selected criteria considering several AI-based auxiliary methods. The final layer is responsible for output representation.

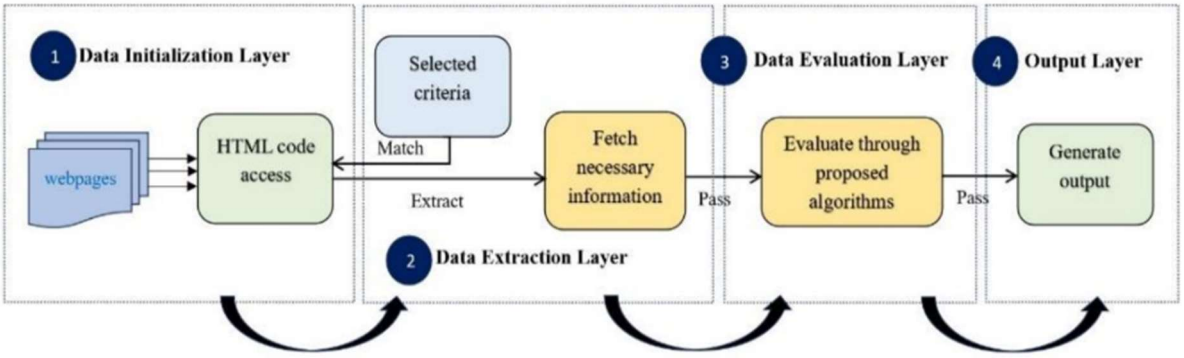


Figure 1
System architecture of the proposed framework

3 Conclusion

The proposed framework is experimented on 15 web pages to evaluate their accessibility status. Besides, the proposed method was also validated through a user study. The details of the user study and experiment results will be shown at the conference.

References

[1] 2023. WebAIM: The WebAIM Million - The 2023 Report on the Accessibility of the Top 1,000,000 Home Pages. https://webaim.org/projects/million/.

[2] Harper, S., & Chen, A. Q. (2012). Web accessibility guidelines: A lesson from the evolving Web. World Wide Web, 15, 61-88.

[3] Cooper, M., Sloan, D., Kelly, B., & Lewthwaite, S. (2012, April). A challenge to web accessibility metrics and guidelines: putting people and processes first. In Proceedings of the international cross-disciplinary conference on Web accessibility (pp. 1-4).

The always changing data problem of using AI in manufacturing - Using synthetic data from the digital twin to feed AI models

Zsolt Molnar
University of Miskolc, Institute of Logistics,
E-mail: zsolt.molnar.zsolt@outlook.com

Abstract: *Production is becoming increasingly flexible, which also requires the flexibility of the support system for the production. Artificial intelligence can be a way to provide control logic and control decisions for the flexible manufacturing lines .The biggest problem is supplying the neural network that controls the artificial intelligence with training data. In the article, we investigate how useful the synthetic data is during the life cycle of the production line.*

Keywords: *flexible manufacturing, layout, systematic layout planning, digital twin, artificial intelligence, decision table*

1 Lifecycle of a Production Line

Increasing the flexibility of production is the result of demand from two sides. On the one hand, the development of IT and automation in recent decades and the geopolitical changes of recent years appeared as an internal demand of manufacturing companies to be able to respond to market demands as flexibly and resiliently as possible. Flexible manufacturing systems consist of three main components: machines, material handling devices and control logic.

During the life cycle of the production line (design, installation and ramp-up, improvement, normal operation, end of life maintenance), different products are produced at different periods [1].

2 Lifecycle of Synthetic and Real Data

The traditional way the collection of data from the production line can start from Zsolt Molnar et al. The always changing data problem of using AI in manufacturing – 2 – the ramp-up period of the line. At that time, however, the quality of the data is generally not yet suitable for use in machine learning due to biases in it. We can only collect better quality data during the normal operation. However, this is often already too late if we are working on an AI model whose task is to control the production line. The situation is much better for synthetic data. Even in the design phase of the production line, the design data can be used to generate the initial batch of the synthetic data. Later, the quality of the synthetic data can be improved as the planning data becomes more accurate. Data mixing can also be used during data generation [2]. In data mixing the data package used to train the AI model contains both real measured data and synthetic data (Figure 1).

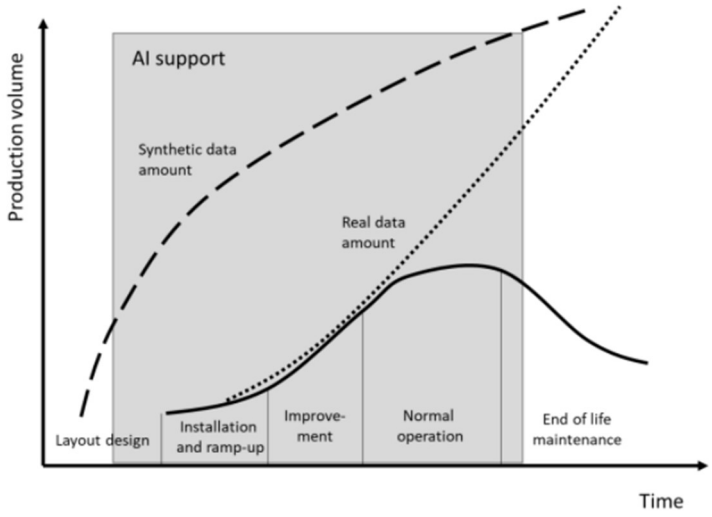


Figure 1
The availability of the real world collected and the synthetic data during the lifecycle.

The article examines the impact of changes in the data underlying artificial intelligence on the control of production processes. It's visible that due to frequent system changes in production, the collected data cannot be used or can only be to a limited extent. Therefore, by mapping the system, a digital twin must be built, from which synthetic data can be collected. The article presents the process through a real industrial example. The novelty of the method and the example is that it connects the life cycle of the production line with the life cycle of the digital twin. This is an exciting approach that will help to ensure that production lines can be used for longer in the future, thereby reducing investment and maintenance costs.

References

[1] Zsolt Molnár, Péter Tamás, Béla Illés, "The Life cycle of the layouts of flexible and reconfigurable manufacturing areas and lines," Advanced Logistic Systems – Theory and Practice, Vol. 15, No. 1 (2021), pp. 20-29.,
[2] F. H. L. S. M. B. G. C. C. M. S. N. C. A. W. James Jordon, "Synthetic Data - what, why and how?," The Alan Turing Institute, The Royal Society, London, 2022.

End-to-End Multi Level Encoding Methods of Visual Data Compression for Robust Monocular Visual ORB-SLAM

Omar M. Salih[1,2], József Vásárhelyi[1]
[1]Institute of Automation and Info-Communication University of Miskolc
[2]Department of Computer Engineering Techniques Northern Technical University
E-mail: omar.salih@uni-miskolc.hu

Abstract: *Recently, Visual SLAM is the most used state-of-the-art technique to implement SLAM tasks due to its reduced cost, lower sized, and affordability. Conversely, the intensive computational of VSLAM systems could not fit in a wide range of limited resources and energy mobile devices. Therefore, a possible solution is to split its functionalities between mobile devices and edge cloud. This point showed the necessity for efficient visual data compression methods to be integrated within ORB-SLAM systems. This work proposes a multi-level encoding method for visual data frame compression within VSLAM systems. The results demonstrated an outstanding performance compared to the most popular JPEG adopted on a ORb-SLAM system.*

Keywords: *ORB-SLAM; data compression; JPEG; image compression; visual perception; localisation and mapping*

1 Introduction

Cloud-based VSLAM solutions have been widely discovered recently. It enables limited computational power for mobile devices to implement SLAM operation efficintly. However, network latencies problem arises when working on low bandwirth networks [1,2]. The contribution of this article is to present a novel multi encoding visual data compression methods, used for compress input frames to high compression ratios while preserving the information quality. Showing the feasibility of integrating within the Compress Then Anayze (CTA) and Anayze Then Compress (ATC) frameworks of ORB-VSLAM. The proposed method is compared to the a corresponding VSLAM architecture utilized JPEG compression standard, which is most popular lossy compression standard.

2 Methodology

The compression algorithm is implemented on the mobile device, which receives the raw frame input data directly form the attached camera. It can be used in both CTA and ATC frameworks. in CTA, the mobile device compresses every input frame and transmit the encoded data to the cloud to decode to process the ORB- SLAM tasks. Regarding ATC framework, the mobile device deploys both the proposed image encoding algorithm along with the tracking module to maintain instant real-time operation. In this case, the tracking is accomplished on the raw sized data while the encoded frames are transmitted to the cloud for map reconstruction and loop closing operation. These frameworks options are selected depending on the application at hand and the computational power of the mobile device itself.

3 Results and Conclusion

The results demonstrated that the ORB-SLAM integrated with the proposed system overcomes the equivalent JPEG system in terms of compression ratio, PSNR, and SSIM. This clear outstanding superiority is attributed to the multi-encoding module added to the compression system. However, this addition increased the execution time, which can be seen clearly in low QF values. As the QF increased, more duplicated values and zeros are present in the data, yielding faster processing, low header and encoded data size, and higher compression ratios. Nevertheless, this higher execution time cannot affect real time operation if we consider the transmission latency, which is beyond the scope of this work. Future direction includes neglecting the compressed data header and harnessing the cloud computing power to directly decompress the results, which can significantly increase the compression ratio. Furthermore, research for efficient hardware implementation using cutting-edge devices, like Field Programmable Gate Arrays (FPGA), is needed since it is well suited to deploy the multiply accumulated operation in the proposed method.

References

[1] A. Tourani, H. Bavle, J. L. Sanchez-Lopez, and H. Voos, “Visual slam: What are the current trends and what to expect?,” Sensors, vol. 22, no. 23, p. 9297, 2022.

[2] J. Hofer, P. Sossalla, C. L. Vielhaus, J. Rischke, M. Reisslein, and F. H. Fitzek, “Comparison of analyze-then-compress methods in edge-assisted visual slam,” IEEE Access, 2023

Application of Process Discovery Methods for Identifying Successful Learning Procedures in Programming

Erika Baksáné Varga, Attila Baksa
University of Miskolc
E-mail: erika.b.varga@uni-miskolc.hu

Abstract: *Process mining is a set of techniques used to analyse event data to understand and improve operational processes. One way of achieving this is to discover the driving process of the activities that occurred in a system. Technically, process discovery algorithms are used to transform an event log into a process model which is representative of the activities registered in the given system. In this paper we report on an experiment with students' commit data concerning programming assignments. The commit events were registered by Github and we applied Alpha Miner, Heuristic Miner and Inductive Miner algorithms as implemented in the pm4py package to discover the learning process of first-year Computer Science students in an introductory programming course.*

Keywords: *process discovery; pm4py; learning process of programming*

1 Introduction

Within the domain of education, understanding how individuals learn and retain information is of great importance in creating effective instructional methods [1]. Traditional approaches to teaching often assume a one-size-fits-all approach, disregarding the fact that each learner has unique preferences and cognitive processes. As a result, this can lead to suboptimal learning outcomes and hinder the development of individuals' skills [2]. Therefore, the identification of successful learning processes can significantly enhance educational practices. To address this, the field of process discovery offers a promising avenue for uncovering and analysing the patterns and strategies that successful learners employ [3]. By applying process discovery methods to the study of learning procedures in programming, instructors can gain valuable insights into students' behaviour and can highlight problems that need more iterations to be solved.

2 Methods

In this research, we employed the pm4py Python program library developed by Fraunhofer Institute to implement algorithms and services connected to process mining. Since process discovery algorithms were in the focus of our investigations, we examined the Alpha Miner, Inductive Miner, and Heuristic Miner methods to reveal the process model from the event log.

3 Data and Results

Our experiment covered 59 practical problems propagated using Github Classroom as out-of-class assignments in an introductory programming course among 70 first- year students of a Computer Science BSc program. Students initiated all together 1678 commits, amongst which 869 were successful. By applying process discovery methods to this event log, we first created

a learning process model that is representative of the individual cases to identify bottleneck problems that hinder students' progress. The second experiment concentrated on producing the learning process model of successful students to be a showcase for other students.

References

[1] M. A. Ghazal, O. Ibrahim and M. A. Salama, "Educational Process Mining: A Systematic Literature Review.", 2017 European Conference on Electrical Engineering and Computer Science (EECS), Bern, pp. 198-203, 2017, doi: 10.1109/EECS.2017.45.

[2] I. Y. Kazu, "The Effect of Learning Styles on Education and the Teaching Process.", Journal of Social Sciences, 5(2), pp. 85-94, 2009, doi: 10.3844/jssp.2009.85.94

[3] W. van der Aalst, "Process Mining – Data Science in Action", Springer Berlin, Heidelberg, 2016, doi: 10.1007/978-3-662-49851-4

A Generator-Evaluator Framework for Automatic Question Generation from Program Codes

Jawad Alshboul, Erika Baksa-Varga
University of Miskolc, Faculty of Mechanical Engineering and Informatics
E-mail: alshboul.jawad@student.uni-miskolc.hu, erika.b.varga@uni-miskolc.hu

Abstract: *This research proposes a framework for automatic generation and evaluation of questions based on large language models. Large language models are causing a stir in understanding computer languages. They are used in coding and software development to generate questions about code automatically. For example, you could input a Python script, and these large language models could spit out a bunch of relevant questions. This makes evaluating code faster and super helpful for code reviews, online tech support, and teaching programming. Technology and software development must keep up with each other in a world that relies so much on technology. However, with so many large language models out there, it is crucial to compare how well they perform in terms of speed and quality of their output. That is where this research comes in. We are diving deep into several of these large language models to see how they compare against each other. We have defined a strict set of standards to measure them against, like how relevant, clear, and concise their responses are and how much ground they cover. By carefully evaluating each model based on these criteria, we get a better idea of what they are good at and where they could improve.*

Keywords: *Large Language Models; Question Generation from Code; Performance Comparison*

1 Introduction

Academic textbooks on computer programming contain texts, images, and codes. The latest studies are centered around crafting questions based on text, though there is also a branch of research delving into crafting questions from visuals or images [1]. The review paper [2] recommends Implementing question-generation approaches to generate questions on programming-related topics and develop evaluation metrics for these generated questions. Ontologies have been used to generate learning materials for educational purposes, as in articles [3] and [4]. Consequently, the authors in [1] proposed a hybrid approach for automatic question generation from program codes. Recently, the study of how computers understand human language, known as Natural Language Processing (NLP), has advanced greatly. This progress is mainly due to the availability of tons of texts for analysis, and the technology behind large language models has improved. These large language models are like the heart of NLP progress. They are incredibly good at understanding and creating human-like language, which has excited researchers.

2 Research Aim and Research Objectives

This research aims to unearth insights that can be pivotal in various applications. By recognizing and showcasing these top performers, we aim to provide valuable guidance to educators, developers, and researchers, offering them a basis for making

informed decisions about utilizing large language models in code-related tasks. The primary objectives of this research are as follows:

RO1. Develop an approach for evaluating and comparing the performance of LLMs in question generation from program codes.

RO2. Define a set of evaluation criteria, including relevance, clarity and coherence, conciseness, and coverage, to measure the quality of questions generated by LLMs.

RO3. Evaluate and rank the selected LLMs based on their performance generating questions from program codes.

References

[1] J. Alshboul and E. Baksa-Varga, "A Hybrid Approach for Automatic Question Generation from Program Codes," Int. J. Adv. Comput. Sci. Appl., vol. 15, no. 1, pp. 10–17, 2024, doi: 10.14569/IJACSA.2024.0150102.

[2] J. Alshboul and E. Baksa-Varga, "A Review of Automatic Question Generation in Teaching Programming," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 10, pp. 45–51, 2022, doi: 10.14569/IJACSA.2022.0131006.

[3] J. Alshboul, H. A. A. Ghanim, and E. Baksa-Varga, "Semantic Modeling for Learning Materials in E-Tutors Systems," J. Softw. Eng. Intell. Syst., vol. 6, no. 2, pp. 85–91, Aug. 2021.

[4] H. A. A. Ghanim, J. Alshboul, and L. Kovács, "Development of Ontology- based Domain Knowledge Model for IT Domain in E-Tutor Systems," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 5, pp. 28–34, 2022, doi: 10.14569/IJACSA.2022.0130505.

Creation of Digital Twins using Spatial Artificial Intelligence - A Pilot Study

Judit Szűcs, Tibor Guzsvinecz
Department of Information Technology and its Applications, University of Pannonia
E-mail: szucs.judit@zek.uni-pannon.hu, guzsvinecz.tibor@zek.uni-pannon.hu

Abstract: *This study explores the integration of spatial artificial intelligence (SAI) in the creation a digital twin of a university campus - namely the Zalaegerszeg University Centre of the University of Pannonia - for immersive virtual reality tours. Spatial AI techniques such as photogrammetry and LiDAR scanning were used to create detailed virtual replicas of campus buildings, landscapes, and infrastructure. With the help of these tours in virtual reality, users can explore the campus remotely, experiencing enhanced accessibility, interactivity, and engagement. Using the 4-factor version of the Presence Questionnaire 3.0 and the System Usability Scale, the virtual environment was evaluated by assessing the opinions of 18 university students. Our findings show the transformative role of spatial AI and virtual reality technology in redefining how environments are experienced and accessed. This can open new ways for spatial cognition research and other, immersive experiences.*

Keywords: *digital twin; digitization; human-computer interaction; spatial AI; virtual reality*

1 Introduction

Digital twins, virtual counterparts mirroring real-world entities or systems, have emerged as a transformative concept with large implications across multiple industries [1, 2]. This paper is an initial step in a larger research project, where our current goal is to test the level of realism of the 3D walkthrough we have produced.

2 Materials and methods

We used spatial AI techniques to digitize the Zalaegerszeg University Centre of the University of Pannonia and provide immersive virtual reality (VR) tours. During the study, two digital twins were created: a point-cloud and a version comprising of 360° photos. Using the 4-factor version of the Presence Questionnaire 3.0 and the System Usability Scale, feedback was gathered from 18 students.

3 Results and discussion

Feedback from participants in the VR tours was positive. They reported high level of presence in the virtual environment, while believing that it was easy to use. Using VR technology for exploration has significant implications: VR can serve as a valuable tool for orientation programs, campus tours, virtual classrooms, and human-computer interaction. By simulating real-world environments, VR can enhance spatial cognition and helps experiential learning opportunities.

4 Conclusion

In conclusion, our study demonstrates the potential of spatial AI and VR technology in digitizing university campuses and providing immersive campus tours. By offering such virtual experiences, VR can revolutionize how we explore and engage with certain environments. As technology continues to evolve, the integration of spatial AI and VR can transform experiences and advance research in spatial cognition and human-computer interaction.

Acknowledgements

This work has been implemented by the TKP2021-NVA-10 project with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the 2021 Thematic Excellence Programme funding scheme.
The authors would also like to thank Patrik Péter Kiss (http://vrgo.hu/) for his help in the digitization process.

References

[1] P. Baranyi, A. Csapo, T. Budai, and G. Wersényi, "Introducing the concept of internet of digital reality–part i," Acta Polytechnica Hungarica, vol. 18, no. 7, pp. 225–240, 2021.
[2] G. Wersényi, Á. Csapó, T. Budai, and P. Baranyi, "Internet of digital reality: Infrastructural background–part ii," Acta Polytechnica Hungarica, vol. 18, no. 8, pp. 91–104, 2021.

Implementation Guidelines for Ethologically Inspired Fuzzy Behaviour-Based Systems

Mohd Aaqib Lone, Szilveszter Kovács, Owais Mujtaba Khanday
Department of Informatics, University of Miskolc
E-mail: iitaaqib@uni-miskolc.hu

Abstract: *The adaptation of ethologically inspired behavior models for human- machine interaction has become a challenging research topic in recent years. Ethology is the study of animal behavior, including for example communication, predation, defense, aggression, and related behaviors and in laboratories. This paper presents a fuzzy behavior description language approach for analyzing animal aggression behavior. To analyze and classify the subjective impression of aggression behavior in a particular situation, fuzzy logic and fuzzy set theory approaches are used. This research aims to perform a meta-analysis of aggression behavior based on the fundamental values of animals and the possible ways of implementing the animal aggressive behavior in robots, ultimately aiming to enhance the adaptability and effectiveness of human-robot interaction and performance in various real-world scenarios, e.g. by expressing disagreement in the direction of the human operator in case of unclear, or unsafe cooperative situations. In both industrial and everyday settings, mobile robots and robotic vehicles are becoming increasingly prevalent. Integrating aggressive behavior into robotics is essential for boosting interactions between humans and robots, promoting safety in dynamic contexts, and getting a deeper understanding of animal behavior. It aids robots in asserting their presence, maneuvering around barriers, and efficiently adjusting to dynamic surroundings. This guarantees more seamless operations in industrial and daily environments, while also enhancing our comprehension of both robotics and ethology. This work presents graphical depictions of various animal behaviors, as well as trajectories, Gazebo simulations, and RViz visualizations of the animal robot, demonstrating the animals escape behavior.*

Keywords: *Ethologically Inspired Behavioural Models, Fuzzy Behaviour Modelling, Fuzzy Behaviour Description Language, Robot operating System, Gazebo, RViz*

References

[1] Taylor, Charles, and Alva Noë. The explanation of behaviour. Routledge, 2021.
[2] Michaud, François, and Monica Nicolescu. "Behavior-based systems." Springer handbook of robotics (2016): 307-328.
[3] Bateson, Patrick, and Kevin N. Laland. "Tinbergen's four questions: an appreciation and an update." Trends in ecology & evolution 28.12 (2013): 712- 718.
[4] Siciliano, Bruno, and Oussama Khatib, eds. Springer handbook of robotics. Springer, 2016.
[5] Vadakkepat, Prahlad, et al. "Fuzzy behaviour-based control of mobile robots." IEEE Transactions on Fuzzy Systems 12.4 (2004): 559-565.
[6] J.H. Connell: SSS: A hybrid architecture applied to robot navigation, Proc. IEEE Int. Conf. Robotics Autom. (1992) pp. 2719–2724
[7] Towle, Bradford A., and Monica Nicolescu. "Real-world implementation of an Auction Behaviour-Based Robotic Architecture (ABBRA)." 2012 IEEE International Conference on Technologies for Practical Robot Applications (TePRA). IEEE, 2012.

[8] Ronald C. Arkin. Behaviour-based robotics. MIT Press, 1998.
[9] Lehner, P.N., 1998. Handbook of ethological methods. Cambridge University Press.
[10] McFarland, D., Bösser, T. and Bosser, T., 1993. Intelligent behaviour in animals and robots. MIT Press.
[11] Hallam, B. and Hayes, G.M., 1992. Comparing robot and animal behaviour. The University of Edinburgh, Department of Artificial Intelligence.
[12] Kovács, Szilveszter. "Interpolative fuzzy reasoning in behaviour-based control." Computational Intelligence, Theory and Applications. Springer, Berlin, Heidelberg, 2005. 159-170.
[13] Adriansyah, Andi. "Design of Context-Dependent Blending (CDB) in Behaviour Based Robot Using Particle Swarm Fuzzy Controller (PSFC)." ICON-CSE 1.1 (2015): 45-50.
[14] Chang, Hyunjin, and Taeseok Jin. "Command fusion-based fuzzy controller design for moving obstacle avoidance of mobile robot." Future information communication technology and applications. Springer, Dordrecht, 2013. 905-913.
[15] Archer, J. “The organisation of aggression and fear in vertebrates; in Perspeaives in Ethology” - Vol. 2 (eds P.P.G. Bateson and P. Klopfer), Plenum Press, New York, 1976. pp. 231-298.
[16] Piller, I., Kovács, Sz.: Fuzzy Behaviour Description Language: A Declarative Language for Interpolative Behaviour Modeling. In: ACTA POLYTECHNICA HUNGARICA 16:9, 2019, pp.47-72.
[17] Piller, I.: EXPRAIL: https://github.com/piller-imre/exprail-python
[18] Piller, I.: FRIBE: https://github.com/piller-imre/fribe-python

Extending Virtual Force Field Navigation with Fuzzy Behaviour Fusion

Mohd Aaqib Lone, Szilveszter Kovács, Owais Mujtaba Khanday
Department of Informatics, University of Miskolc
E-mail: iitaaqib@uni-miskolc.hu

Abstract: *The increasing demand for human-robot cooperation requires the creation of solid object avoidance skills to guarantee safe and effective functioning. This paper introduces an innovative approach that combines fuzzy behavior fusion with virtual force field navigation to empower the robot to achieve its goal. The study encompasses animal behaviors, such as animal escape behavior, and how a robot (animal) will behave when an unfamiliar object or another unfamiliar robot approaches while the robot is trying to do its goal task. Utilizing virtual force field navigation and fuzzy behavior fusion has been demonstrated to be a successful approach for guiding robots toward their desired goals. The systems capacity to adjust to evolving conditions amplifies its overall efficacy and dependability. The proposed strategy has the potential to improve the efficiency of robot behavior in various sectors, such as manufacturing, logistics, and healthcare.*

Keywords: *Virtual Force Field Theory, Robot Operating System, Gazebo, Rviz, Fuzzy Behavior Fusion.*

References

[1] Gul, F., Rahiman, W. and Nazli Alhady, S.S., 2019. A comprehensive study for robot navigation techniques. Cogent Engineering, 6(1), p.1632046.
[2] Pronobis and B. Caputo. "COLD: cosy localization database," The International Journal of Robotics Research, vol. 28, no. 5, pp. 588-594, May
2009.
[3] Huntingford, F. ed., 2012. The study of animal behaviour. Springer Science & Business Media.
[4] Scovanner, P. and Tappen, M.F., 2009, September. Learning pedestrian dynamics from the real world. In 2009 IEEE 12th International Conference on Computer Vision (pp. 381-388). IEEE.
[5] Skarmeta, Antonio Gómez, Humberto Martínez Barberá, and Manuel Sánchez Alonso. "Learning behaviour fusion in autonomous mobile robots." IX ESTYLF 99 (1999). [6] J.H. Connell: SSS: A hybrid architecture applied to robot navigation, Proc.
IEEE Int. Conf. Robotics Autom. (1992) pp. 2719–2724
[6] Zadeh, Lotfi Asker. "Fuzzy sets as a basis for a theory of possibility." Fuzzy sets and systems 1.1 (1978): 3-28.
[7] Jom J. Kandathil1, Robins Mathew1, and Somashekhar S. Hiremath Modified bug-1 algorithm-based strategy for obstacle avoidance in the multi-robot system.
[8] K.-S. Tseng and C.-W. Tang, "Stream field based people searching and tracking conditioned on SLAM," IEEE International Conference on Robotics and Automation Workshop on People Detection and Tracking, May 2009.

[9] Sun, Q., Guo, Y., Fu, R., Wang, C. and Yuan, W., 2020. Human-Like Obstacle Avoidance Trajectory Planning and Tracking Model for Autonomous Vehicles That Considers the Driver's Operation Characteristics. Sensors, 20(17), p.4821.
[10] J. Borenstein, Member, IEEE, and Y. Koren,' Real-time Obstacle Avoidance for Fast Mobile Robots, 1989 IEEE. Reprinted, with permission, from IEEE Transactions on Systems, Man, and Cybernetics, Vol. 19, No.5, Sept./Oct. 1989, pp. 1179-1187.
[11] Piller, I., Kovács, Sz.: Fuzzy Behaviour Description Language: A Declarative Language for Interpolative Behaviour Modeling. In: ACTA POLYTECHNICA HUNGARICA 16:9, 2019, pp.47-72.
[12] ROS – Robot operating system, Available online: https://www.ros.org/

Comparative Analysis of NLP-Driven MCQ Generators from Text Sources

Richárd Németh, Asmae Azzi and Ferenc Erdős
Department of Information Technology, Faculty of Engineering Sciences, Széchenyi István University
E-mail: nemeth.richard@ga.sze.hu, asmae.azzi@sze.hu, erdosf@sze.hu

Recent studies on the impact of technology on student learning show that contemporary tools increase student engagement. One of the most effective technologies is natural language processing, a sub-field of AI that focuses on human languages and can be used in education by generating multiple-choice questions (MCQs) based on typed text. MCQs are a type of exam question in which the respondent is presented with different answer choices for a question, from which he or she must select the correct answer.

Our analysis involved testing different software with varied difficulty input texts. We evaluated text summarization, keyword extraction accuracy, question pertinence, and distractor quality (to assess comprehension and reasoning). However, it might be time-consuming to generate high-quality MCQs. As the AI field is rapidly evolving, multiple tools have been developed to help users in diverse areas, including education. This led to the creation of a multitude of software designed for the automatic generation of MCQs from an input text. The research began by examining relevant studies reviewed various automatic MCQ generation techniques, including open source programs, datasets and commercial solutions. We also examined a number of barriers to MCQ generation, such as the need to produce MCQs of high quality. This highlights the need for a comparative analysis to understand the advantages and disadvantages of each tool for instructors and quiz makers. We have two types of AI tools for MCQ generation: large language models (ChatGPT, Google Gemini) and MCQ-specific solutions (QuizGecko, Questgen). These programs typically turn declarative sentences into questions based on keywords and main idea. Recently, RNN, LSTM and GRU neural networks were used for word processing tasks, but with the so-called attention mechamism and positional encoding, transformers models are much more suitable for the task and capable of higher speeds, making it more efficient to use.

To ensure unbiased analysis, we built our own dataset for testing. This dataset included four texts of varying difficulty (CEFR levels A2 to C2), combining academic and public source materials. This variety reflects different writing styles and complexities, making the dataset well-rounded for assessing the adaptability of MCQ generation tools. In the examination, it was also important to define clear and objective evaluation metrics which are free to any biases to measure the efficiency of each application. The first part of the examination checks the clarity, relevance, and difficulty level of the questions generated by each tool. In order to have an efficient MCQ quiz, we should have cohesive and clear questions that match the complexity of the text. In that way, students can be efficiently assessed about their learnings depending on their English level and the topic covered. This method allows us to differentiate between the tools that can only generate basic questions and the ones that are more sophisticated. After defining the evaluation metrics, we proceeded with a collaborative evaluation:we invited ten participants to evaluate the different tools too based on the same metrics. Every participant was

asked to evaluate the tools based on the primary metrics that included: the quality of every question, the quality of each set of distractors, and the quality of the right answer.

Our assessment were carried out in three main categories. In the collaborative analyses, the quality of the questions, the generation of correct answers and the generation of distractors were analysed. For the secondary metrics, we compared non-functional characteristics of the tools, such as ease of use, speed of generating MCQs, and individual limitations. We then subjected each model to a performance evaluation, where we applied each of the classification categories of the confusion matrix we further developed to fit our purpse. After that, we grouped all of the 800 ratings from 10 participants into classes and we performed a distribution of the most common metrics used in ML model performance evaluation: Propotion, Accuracy, Precision, Recall and F1 Score, to compare the effectiveness of each tool.

Based on the results, the primary metrics yielded some noteworthy conclusions. ChatGPT and Gemini excelled in core metrics like question quality. Questgen and QuizGecko followed closely. When the tools were evaluated using secondary metrics, a new set of observations appeared Usability-wise, QuizGecko and Questgen shone with user-friendly interfaces for customizing quizzes. In terms of the speed of generation, most tools worked well, with the exception of QuizGecko, which had a noticeably longer processing time. As far as performance evaluation is concerned, the metrics show an even more mixed picture. In the competition within the category, Gemini beats ChatGPT and Questgen beat the underperforming QuizGecko, as we also find that in the results of our present test, there is no difference between general-purpose AI and purpose-built AI.

Our analysis helps instructors choose the best MCQ generator for their needs by revealing each tool's strengths and weaknesses. As technology improves and demand for customized assessments rises, MCQ generation is likely to become more effective and accurate. However, further research with broader applications and a wider sample size is needed for definitive conclusions.

Tailoring tourism decisions based on artificial intelligence to align with ideal preferences for accommodations

Judit Grotte, László Pittik
Kodolányi János University,
E-mail: grotte.judit@kodolanyi.hu, pitlik.laszlo@kodolanyi.hu

Abstract: *In today's fast-paced technological landscape, the process of booking accommodations has significantly improved in terms of convenience. Rapid technological advancements have not only revolutionized our lifestyles, but also spurred the emergence of new behavioral patterns and communication habits. The integration of AI-based decision-making has notably benefited various sectors, including the tourism industry, by providing optimized customer-centric solutions. Both traditional procurement processes and value-driven decision-making now heavily rely on Object-Attribute-Matrix versions for enhanced efficiency and effectiveness (c.f. https://miau.my-x.hu/miau2009/index.php3?x=e0&string=ideal-).*

Through the use of Artificial Intelligence, tourism businesses are able to collect, analyze, and manage extensive amounts of data. For instance, in the case of accommodations, this may encompass various aspects of the services such as smoking policies, parking availability, shower quality, provision of towels, and distances from key landmarks. This wealth of data enables hotels and online travel platforms to offer personalized tourism experiences to both business and leisure travelers, considering the location of the accommodation and the value-for-money factor. This approach differs from traditional methods which typically optimize based on a price-performance ratio. Customers now have the ability to specify their preferences, such as the proximity to the city center in relation to price. This direction-oriented preference system should be established for each attribute. Decision-making scenarios driven by human ideals involve an envisioned "dream accommodation," against which real accommodations are assessed based on their virtual multi-dimensional distances, rather than their actual performance characteristics. Notably, accommodations closer to the ideal standard may command higher prices. This represents a groundbreaking approach towards personalized customization through automation. In the methodology section of the presentation, the authors emphasize that accommodations have numerous attributes, but it is challenging to obtain genuine Online Accommodation Matching (OAM) from support portals. As a result, tourists must manually conduct analyses, which is a time-consuming process. However, the portals have structured data accessible through APIs and could be utilized if there is an economic incentive for the portals to integrate such a service. The consideration of prospects involves the potential automation and integration of this type of decision support within large portals as a specialized service.

Keywords: *customization, tourism, data-driven decision-making, AI, ideal-oriented price/performance*

References

[1] P. Baranyi, A. Csapo, and Gy. Sallai: „Cognitive Infocommunications (CogInfoCom)”, Springer International Publishing Switzerland, 2015, p. 191. 978-3-319-19607-7, 978-3-319-19608-4 eBook, doi: 10.1007/978-3-319-19608
http://www.springer.com/us/book/9783319196077#aboutBook

[2] P. Baranyi, Á. Csapó, T. Budai, and G. Wersényi, “Introducing the Concept of Internet of Digital Reality–Part I.”, Acta Polytechnica Hungarica, vol. 18, no. 7. 2021, pp. 225-240.
[3] I Horváth and B. Berki, "Unlocking the Potential of Immersive VR Systems: The Role of Gesture and Natural Controls," 2023 IEEE 2nd International Conference on Cognitive Aspects of Virtual Reality (CVR), Veszprém, Hungary, 2023, pp. 000093-000098, doi: 10.1109/CVR58941.2023.10394795
[4] Juhász, B.G. Kálmán, A. Tóth, and A. Horváth, “Digital Competence in the Visegrád Countries”, Digital Competence Development in a Few Countries of the European Union, Management and Marketing, vol. 17, no. 2, 2022, pp. 178-192. https://doi.org/10.2478/mmcks-2022-0010
[5] Ishwar Khatri: Information Technology in Tourism & Hospitality Industry: A Review of Ten Years’ Publications, Journal of Tourism & Hospitality Educaϴ on 9 (2019) 74-87

UN Peacekeeping in the Age of AI

Richárd Schneider[1,2]
[1]Africa Research Institute, Doctoral School on Safety and Security Sciences, Óbuda University
2Center for Geopolitics, Mathias Corvinus Collegium Foundation
E-mail: schneider.richard@bgk.uni-obuda.hu, schneider.richard@mcc.hu

---

Abstract: *The first peacekeeping operation was established in 1948 and since then, there have been more than 70 UN peace operations around the world. Over the years, UN peacekeeping has evolved to meet the demands of different conflicts and the ever- changing political landscape. UN peacekeeping, which was traditionally developed to deal with inter-state wars, has been increasingly adjusted to new (Kaldor, 2012) intra-state conflicts since the 1990s when armed oppositional militias and other non-state actors have become the key actors. Until the 1990s, super power rivalries during the Cold War paralyzed the functioning of the UN Security Council, therefore the priorities of international peacekeeping remained limited to maintaining ceasefires and stabilizing conflict situations with simple and narrow mandates (Oksamytna and Karlsrud, 2023). Then, since the beginning of the 1990s and especially from the 2000s, the UN started to shift and expand its field operations from the 'traditional' approach with military and law-enforcement priorities to more complex missions. This enthusiasm was outlined for example in the document titled "An Agenda for Peace" in 1992 (Boutros-Ghali, 1992) and the Brahimi Report in 2000 (UNGA, 2000). After the failures and tragedies between 1992 and 1995 (whether it comes to the tragedy in Somalia, the massacres in Srebrenica or the genocide in Rwanda, to mention a few) an exhaustive academic literature argued that a completely new approach is needed in the course of peace operations to better address the root causes of conflicts (e.g., Colletta and Muggah, 2009) and resolve humanitarian crises. Some scholars also argued that due to globalization and technological revolution new approaches were needed to address the legal and moral challenges induced by new innovations and technological revolution (Kaldor, 2019). During the last two decades, the UN has already issued several guidelines and strategies to better address the ever-changing nature of international politics and armed conflicts in the age of artificial intelligence (AI) (e.g., UN DPPA and UN Global Pulse, 2019). The intensifying technological revolution and related data risks are among the most urgent challenges that the organization should deal with. Even though the International Committee of the Red Cross found that AI-powered autonomous systems are not at all in compliance with and go beyond the traditional questions of existing rules of international humanitarian law (ICRC, 2018), Secretary-General António Guterres stated in 2018 in his Strategy on New Technologies that "engagement with new technologies is necessary for preserving the values of the UN Charter and the implementation of existing UN mandates" (UNSC, 2018: 4.). Now, the UN is to support the adoption of automated algorithmic systems for managing large-scale data and it has just started to enhance its AI-powered analytical and predictive capacities to better react in conflict situations (Pauwels, 2020; UN Global Pulse 2018). In Somalia, the UN used social media content analysis and data from public Facebook groups to detect how influencers and fake news thematize public discourse and identify trending topics related to peace and security (UN Global Pulse, 2018). In Yemen, itinitiated an AI-powered, large-scale dialogue with over 500 Yemenis to hold real- time conversations and make its peace mission more inclusive (OSESGY, 2020). Furthermore, based on a joint effort made by UN agencies and member states, a Smart Camp initiative has been launched in 2020*

to facilitate planning, management and decision making, enhance safety and security, minimize the mission's environmental footprint and improve day to day life in Dukra camp, along the border between Sudan and South Sudan. The presentation provides a brief summary on the evolution of UN peacekeeping and some of the latest innovation projects within the UN system. Among them, a special attention is devoted to an Ugandan case study where AI-powered radio content analysis methods were developed by the UN to monitor public discourse in host communities during a refugee crisis in 2015 and to explore how insights from public radio talks could provide real time information on different issues of concern. After conducting field research in Uganda and making expert interviews with data scientists (UN) in April, 2023, the purpose of this presentation is also to raise some awareness to the latest AI innovations and trends within the UN.

Keywords: *peace operations; United Nations; artificial intelligence; UN Pulse Lab*

References

[1] Boutros-Ghali, B., "An Agenda for Peace. Preventive Diplomacy, Peacemaking and Peace-keeping", New York: United Nations, 1992, 58. https://digitallibrary.un.org/record/145749?ln=en&v=pdf

[2] Colletta, Nat J., and Muggah, Robert, "Context matters: interim stabilisation and second generation approaches to security promotion" Journal of Conflict, Security and Development 9(4), 2009: 425−453. DOI: 10.1080/14678800903345762

[3] UNGA, "A/55/305-S/2000/809", United Nations General Assembly, 21 August, 2000, 58. https://peacekeeping.un.org/sites/default/files/a_55_305_e_brahimi_report.pdf

[4] ICRC, "The Humanitarian Metadata Problem: "Doing No Harm" in the Digital Era", Geneva: International Committee of the Red Cross, 2018, 130.

[5] Kaldor, M., „New and Old Wars: Organized Violence in a Global Era". Cambridge: Polity Press, 2012, 3rd Edition.

[6] Kaldor, M., "Peacemaking in an Era of New Wars", in: Thomas de Waal (ed.): Think Peace: Essays for an Age of Disorder. Carnegie Endowment for International Peace, 2019, 21-26. https://carnegieeurope.eu/2019/10/14/think- peace-essays-for-age-of-disorder-pub-79969

[7] Oksamytna, K. and Karlsrud, J., "United Nations Peace Operations and International Relations theory: An introduction", in: Kseniya Oksamytna and John Karlsrud (eds.): United Nations Peace Operations and International Relations Theory. Manchester: Manchester University Press, 2023, 1-30.

[8] OSESGY, "Cutting-Edge Tech in the Service of Inclusive Peace in Yemen", Office of the Special Envoy of the Secretary-General for Yemen. 3 August, 2020. https://osesgy.unmissions.org/cutting-edge-tech-service-inclusive-peace- yemen

[9] Pauwels, E., "Artificial Intelligence and Data Capture Technologies in Violence and Conflict Prevention. Opportunities and Challenges for the International Community". Global Center on Cooperative Security, Policy Brief, 2020, 20. https://www.collegesidekick.com/study-docs/713829

[10] UNDPPA and UN Global Pulse, "E-Analytics Guide: Using Data and New Technology for Peacemaking, Preventive Diplomacy and Peacebuilding". New York: UN Global Pulse, United Nations University Institute on Computing and Society, 2019, 26. https://www.unglobalpulse.org/document/e-analytics-guide- using-data-and-new-technology-for-peacemaking-preventive-diplomacy-and- peacebuilding/.

[11] UN Global Pulse, "Experimenting with Big Data and Artificial Intelligence to Support Peace and Security", 2018, 30. https://www.unglobalpulse.org/document/experimenting-with-big-data-and-ai- to-support-peace-and-security/

[12] UNSC, Strategy on New Technologies, 2018, 19. https://www.un.org/en/newtechnologies/

Folk Dance Lab – Measuring Dancer's Perception of Non-Isochronous Meter with Motion Capture

Anna Maria Bólya
HAA Research Institute of Art Theory and Methodology

Keywords: *nonisochronousmeter; folkdancemotioncapture; dancedidacticdevelopment; hungarianfolkdance*

1 State of the Art

The use of motion capture technology in the preservation and documentation of traditional dance forms has been demonstrated in a number of studies. [1][2][3]. Examples of three-dimensional digitisation of dance heritage already exist. [4] At the same time, the didactic development of the teaching of this important and advanced activity phenomenon of our kinetic artistic heritage has only scattered examples. [5]

2 Virtual Reality in Performing Arts - Dance Pilots

The present pilot is a dance curriculum development project, initiated in 2018. The benefits of virtual reality in the performing arts, which are based on three fundamental principles, are as follows:
1. A complete didactic reform to achieve kinetic and artistic goals.
2. A multidisciplinary dance research project to frame dance teaching material.
3. The implementation of pilots to preserve and publish dance cultural heritage.
Firstly, in collaboration with international universities, an innovative curriculum of dance teaching materials in a virtual space with a historical perspective was developed and tested. [6][7][8][9] Subsequently, footage of ballet and Hungarian folk dance was produced utilising motion capture. The material was subjected to a testing. The principal collaborators were the CYENS Centre of Excellence in Nicosia, the Csillagszemű Dance Ensemble, which employs the Timár-method, and Europabellett Sankt Pölten.

3 A specific study of folk dance motion capture data

One of the specific objectives of our testing of folk dance material is to assess the dancer's perception of non-isochronous dance music. The enculturation of music based on isochronous rhythms makes the perception and reproduction of non- isochronous music challenging. One of the primary challenges in folk dance education for numerous ethnic groups, including Hungarians, is the ability to transition between the two types of perception. [10]
A substantial body of research has demonstrated the significance of rhythm perception in body movement. [11][12][13] Consequently, we examined the relationship between music perception and the precision of dance execution in the context of non-isochronous music. The distinction in perception was illustrated by superimposing two frecorded versions of the same dance from the Hungarian male solo dance material entitled "Ritka legényes". [14] The visualisation facilitated the identification of the constant elements of the perception of non-isochronous music, thus enabling an assessment of the accuracy of music perception. This

allowed us to obtain data for measuring the perception of the efficacy of switching to non-isochronous music and musical adaptation in body movement.
The superimposition of motion-capture footage has proved an invaluable tool in the didactic development of the dancer's perception, facilitating the teaching of the switch between the two senses and following music. This has enabled the creation of more effective feedback throughout the process. The presentation will provide a concise overview of the findings from this pilot study on the measurement of perception.

beat 1	beat 2	beat 3	beat 4
I	I	I	I
2	2	2	2

Figure 1
"Sűrű legényes" dance music – isochronous meter.

beat 1		beat 2	beat 3
I.		I	I
3		2	2

Figure 2
Macedonian dance music – defined non-isohronous meter, a comparative South Slavic example.

beat 1	beat 2	beat 3	beat 4
I	I	I	I
2	between 1 and 2	2	2

Figure 3
"Ritka legényes" dance music – non-defined non-isochronous meter.
This is a metrical structure that can be situated between the Macedonian and the other Hungarian example.

Acknowledgements

The entire project was supported by the HAA Research Institute of Art Theory and Methodology, the International Visegrad Fund, the Arts and Research Ltd., the EFOP-3.6.3-VEKOP programme. We are grateful to the European Social Fund, the University of Debrecen, the MDPI publisher, and E-Rent Mobility Solutions for their support.

References

[1] Hegarini E., Dharmayanti and A. Syakur, "Indonesian traditional dance motion capture documentation," 2016 2nd International Conference on Science and Technology-Computer (ICST), Yogyakarta, Indonesia, 2016, pp. 108-111, doi: 10.1109/ICSTC.2016.7877357.

[2] Aristidou, Andreas, Alan Chalmers, Yiorgos Chrysanthou, Céline Loscos, Franck Multon, J. Parkins, Bhuvan Sarupuri, and Efstathios Stavrakis. "Safeguarding our dance cultural heritage." In Eurographics 2022-43nd Annual Conference of the European Association for Computer Graphics, pp. 1-6. 2022. https://doi.org/10.2312/egt.20221055

[3] Aristidou, Andreas, Anastasios Yiannakidis, Kfir Aberman, Daniel Cohen-Or, Ariel Shamir and Yiorgos Chrysanthou. "Rhythm is a Dancer: Music-Driven Motion Synthesis With Global Structure" in IEEE Transactions on Visualization & Computer Graphics, vol. 29, no. 08, pp. 3519-3534, 2023. https://doi.ieeecomputersociety.org/10.1109/TVCG.2022.3163676

[4] Aristidou, Andreas, Nefeli Andreou, Loukas Charalambous, Anastasios Yiannakidis and Yiorgos Chrysanthou. “Virtual Dance Museum: the Case of Greek/Cypriot Folk Dancing.” Eurographics Workshop on Graphics and Cultural Heritage, 2021. https://doi.org/10.2312/gch.20211405

[5] Bólya, Anna Maria, and László Koppány Csáji. "Benefits of Virtual Reality in Ballet Didactics." In 2023 IEEE 2nd International Conference on Cognitive Aspects of Virtual Reality (CVR), pp. 000103-000108. IEEE, 2023. https://doi.org/10.1109/CVR58941.2023.10394904

[6] Bólya, Anna Mária, Gilányi, Attila (Eds.) "Benefits of Virtual Reality in Art - Testing on a Dance History Teaching Case." Debrecen-Budapest HU: UD FH, UD FI, Arts and Research Ltd.

[7] Gilanyi, Attila, M. A. Bólya, and Katarzyna Chmielewska. "Early history of Hungarian ballet in virtual reality." In 2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 193-198. IEEE, 2019. https://doi.org/10.1109/CogInfoCom47531.2019.9089943

[8] Rácz, Anna, Attila Gilányi, Anna Mária Bólya, János Décsei, and Katarzyna Chmielewska. "On a model of the first National Theater of Hungary in MaxWhere." In 2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 000573-000574. IEEE, 2020. https://doi.org/10.1109/CogInfoCom50765.2020.9237848

[9] Gilányi, Attila, Anna Rácz, Anna Mária Bòlya, János Décsei, and Katarzyna Chmielewska. "A presentation room in the virtual building of the first National Theater of Hungary." In 2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 000517-000522. IEEE, 2020. https://doi.org/10.1109/CogInfoCom50765.2020.9237828

[10] Ullal-Gupta, S., Hannon, E., & Snyder, J., Tapping to a slow tempo in the presence of simple and complex meters reveals experience-specific biases for processing music. Plos One, 9, 7, e102962, 2014. https://doi.org/10.1371/journal.pone.0102962

[11] Haugen, Mari Romarheim. "An Embodied Perspective on Rhythm in Music– Dance Genres." Performing Time: Synchrony and Temporal Flow in Music and Dance 151-167, 2023. https://doi.org/10.1093/oso/9780192896254.003.0015

[12] Misgeld, O. and Holzapfel, A., Towards the study of embodied meter in Swedish folk dance. In Folk Music Analysis Workshop. 2018.
[13] Su, Yi-Huang. 2017. “Rhythm of Music Seen Through Dance: Probing Music– dance Coupling by Audiovisual Meter Perception.” PsyArXiv. November 8. https://doi.org/10.31234/osf.io/ujkq9
[14] Timár, Mihály, And Anna Mária Bólya. 2021. "Additions to the issues of European folklore processing in connection with the style of Sándor Timár and the Tanec ensemble in Skopje". Macedón Tudományos és Kulturális Közlemények 4, 1, 21-32. https://doi.org/10.56309/mtkk.v4i1.10024

Folk Dance Motion Capture LAB Evaluation Through The Visualization of BMN Benesh Movement Notation System - TRACK III

Barbara G. Haines
Dance Faculty, Musical Theatre Department, University of Mobile,

Abstract: *Dance is an ephemeral art form, created on the human body and presented to an audience for a moment. The typology of dance - technology is one of the last performing art forms to develop. Notation scores for dance, music, and theater have existed for some time, but the remaining notation systems2 that perpetuates dance literacy are relatively unknown except to a few trained notators. From early pioneers in the field [1a)] [1b)] [1c)] to late scholars, [2a)] [2b)] [2c)] it has been recognized that a translator (DanceWrite3 & Mac Life Form / Laban writer4) between existing notation systems, and a 3D avatar (based on human figure animation) could dramatically increase the ability to support dance education and its subfields in higher education institutions. This study is to analyze a post-production of an ongoing project [3] from a triangular methodological approach: instructor's perspective (dance and BMN), students' feedback (elementary5 and college6 level in dance and BMN) and choreographer's7 perspective (dance and LMA) to bring an experimental result to validate VMO [4]. It aims to find a solution for the missing semantic gap. It points us in the direction of why an application of the BMN system to OWL [5] to form a BeneshReader-Writer [6] notation editor for the technology (PhaseSpace Impulse X2) [PhaseSpace2014] used in the project would be a useful pedagogical tool to fill the missing semantic gap. And how this interface would allow for a complex, fluid approach between semiotics and phenomenology with an increased practice domain between art, science, and technology.*

Keywords*: Folk Dance Methodology, Applied Ethnochoreology, and Ethnomusicology, Supported Visual Effects, Collaborative Creativity, Presentational Choreographic Representation,*

1 Conceptualization

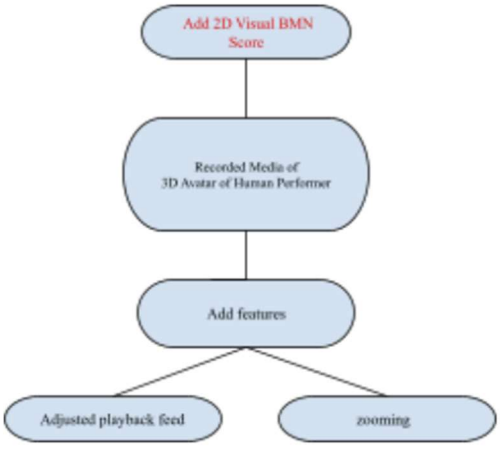
1) BMN's guide to building a Video Movement Ontology (VMO) for Folk Dance Mocap Lab with proposed goal: Establish a BWR-PS database, development of a PS-BWR automatic translation with added animation softwares to support classroom instruction and e-learning.

Mapping BMN for VMO Prototype:

- 3 D visual effect of body makes it more reader friendly
- Score Reading from Left to Right / Bottom to Top
- Movement and its timing (musical / non-musical)
- Presenting from the dancer's point of view - Behind
- Shape & Artistic Style of Movement Lines, Flexions, Extensions, direction faced, path of travel

Need: Check for correctness (Compare 2D Visual Notation Score with Computer generated Score, Ruling Key-Frames.

2) Studying with LMA during the Folk Dance Mocap LAB [3]:



LMA based 3D Avatar Recording Needs::

- No Visual Score - Add 2D Visual BMN Score
- Offers 3 D rotations to see (body kinesphere) extremities in space
- Add features: Adjusted playback feed / zooming
- Test on elementary level for in-classroom education with instructor's assistance
- Test on collegiate level for e-learning with the use of BMN
- Evaluate LMA in today's Folk Dance Education from choreographers' perspective

2 Conclusion

LMA is probably the most widely taught dance notation, but the Benesh system is the most widely used for its visual effects. To overcome the current semantic gap problem and to improve current teaching and learning methods: 1) formation of an integral system of methodological (BMN) and new technology development in the field of cultural studies, 2) creation of experimental sites for the conduct and implementation of scientific, theoretical and practical research, testing of innovative technologies - (Scientific & Methodological Laboratory), 3) correspond to the priority directions and problems of development of folk dance mastery, pedagogy of choreography, music theory, art management, social and humanitarian sciences. Increase the efficiency of scientific work through the implementation of fundamental and applied research in the field of choreographic art.

3 General Conditions

Literature, virtual performance, knowledge / cognitive skills, language design, origin choreographic intention / cultural heritage

A substantial body of research has demonstrated the significance of rhythm perception in body movement. [11][12][13] Consequently, we examined the relationship between music perception and the precision of dance execution in the context of non-isochronous music. The distinction in perception was illustrated by superimposing two frecorded versions of the same dance from the Hungarian male solo dance material entitled "Ritka legényes". [14] The visualisation facilitated the identification of the constant elements of the perception of non-isochronous music, thus enabling an assessment of the accuracy of music perception. This allowed us to obtain data for measuring the perception of the efficacy of switching to non-isochronous music and musical adaptation in body movement.

The superimposition of motion-capture footage has proved an invaluable tool in the didactic development of the dancer's perception, facilitating the teaching of the switch between the two senses and following music. This has enabled the creation of more effective feedback throughout the process. The presentation will provide a concise overview of the findings from this pilot study on the measurement of perception.

References

[1a] The Estate Project’s afterimages (Preserve, Inc. ‘s, newsletter of performing arts documentation and preservation)

[1b] Gray, J. “Dance in Computer Technology: a survey of applications and capabilities’. Interchange, 14(4):15-25, Winter 1984,

[1c] Calvert, T.W “Towards a Language for Human Movement.” Computers and the Humanities, 20:2, (1986), pp.35-43,

[2a] C, Michael. and T,Calvert. “Choreography and Animation: Tools for Visualization of Dance and Notation.” 2002

[2b] A, Ivarson, “Technology Adoption in the Performing Arts: The Case of ‘technology Adoption Capability’, 2005

[2c] S, Saad. and D, De Beul. and S, Mahmoudi. And P, Manneback. “An Ontology for Video Human Movement representation based on Benesh Notation

[3] A, Bólya,. MMK and A, Aristidou, CYENS and M,Tímár. and G, Tímár. Csillagszemű Folk Dance Ensemble “ Folk Dance Motion Capture Lab” CYENS Centre of Excellence, University of Cyprus, Nicosia, Cyprus, May, 2024

[4] Video Movement Ontology / The World Wide Web Consortium (W3C) Semantic Web technology https://www.w3.org/

[5] Ontology Web Language https://www.w3.org/TR/owl2-overview/

[6] MacBenesh Editor https://dancewrite.com/MacBeneshManual.pdf/

The Fairness of ChatGPT in Generating Scientific Literature Reviews

Judit Hermán[1], Kíra Diána Kovács[1], Orsolya Vásárhelyi[2*]
[1]Budapest University of Technology and Economics, Műegyetem
[2] Center for Collective Learning, Corvinus Institute of Advanced Studies & Institute of Data Analytics and Information Systems, Corvinus University of Budapest, Corvinus University of Budapest
E-mail: orsolya.vasarhelyi@uni-corvinus.edu

Abstract: *This study explores the fairness of ChatGPT in generating scientific literature reviews and citations. We used OpenAlex as our baseline dataset, and exported all 20 most popular scientific research topics within 26 dif erent main scientific research fields. We calculated the consistency of the reviews using average cosine similarity, and found that the answers are not so random for the same prompt, their consistency was between 0.39 and 0.48, the highest consistency was for the medical fields and lowest for arts, humanities and social sciences. We also found that in almost every field the average proportion of women was higher in the references provided by the ChatGPT than the papers in the OpenAlex database, which indicates that work produced by more gender-diverse teams is better represented than by the scientific community in general. This study highlights potential biases in LLM-generated academic content, underscoring the need for ongoing evaluation and refinement of these models to ensure equitable representation in scientific research.*

Keywords: *ChatGPT; gender bias; scientometrics*

1 Introduction

On 30 November 2022, OpenAI has published ChatGPT, an AI language model, which has abruptly changed the way we communicate, learn and conduct research.. The use of ChatGPT has also become widespread in knowledge production and scientific publication[1]. It holds promise in tackling long-standing inequalities in academia, such as helping non-native English speakers articulate their scientific discovery better, more efficiently, with a more coherent style[2]. Moreover, generative AI, trained on scientific datasets, could potentially generate reviews of the less biased and more thorough scientific literature reviews [3]. In scientific production, previous versions of ChatGPT have been shown to create factually incorrect answers and generate non-existing references [4].

2 Data and Methods

In this paper, we examine references generated by the latest ChatGPT model (GPT-4o) in 26 different research areas 1 . Specifically engineered a prompt that asked ChatGPT to write literature reviews in several subfields and provide references that contained authors' full names, article titles, journals, and publication year with DOI. We compared the resulting references by fields with all publications cited at least twice in OpenAlex by articles published within the

given subfield. Our analysis consists of the cited articles publication years, journals by impact category, and the gender setup of the co-author teams. We also investigated the consistency of the literature reviews the model wrote on the same topic.

3 Results

We investigated whether the references given by ChatGPT were biased compared to the references used by scientists within the research field in OpenAlex. We found that ChatGPT cites newer papers than what the distribution of citations would suggest in 65% of the main research fields. Our analysis also revealed that ChatGPT cites papers from different impact categories with similar probability, which results in a lower level of Q1 citations than the baseline would suggest. Finally, we computed the average ratio of women among cited papers and found that ChatGPT cites papers with a higher ratio of women than field-specific ratios would suggest. The Mann-Whitney test showed a significant difference for all fields except 4 (Business, Management and Accounting, Computer Science, Immunology and Microbiology, Mathematics). However, these results could be due to citing newer works, since the ratio of women has increased in scientific fields at the last 50 years.

References

[1] Xames, Md Doulotuzzaman, and Jannatul Shefa. "ChatGPT for research and publication: Opportunities and challenges." Xames, MD, & Shefa, J. ChatGPT for research and publication: Opportunities and challenges. Journal of Applied Learning and Teaching 6.1 (2023).

[2] Lund, Brady D., et al. "ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing." Journal of the Association for Information Science and Technology 74.5 (2023): 570-581.

[3] Wang, Hanchen, et al. "Scientific discovery in the age of artificial intelligence." Nature 620.7972 (2023): 47-60.

[4] Chelli, Mikaël, et al. "Hallucination Rates and Reference Accuracy of ChatGPT and Bard for Systematic Reviews: Comparative Analysis." Journal of Medical Internet Research 26 (2024): e53164.

Towards Incorporating AI to Generate Personalized and Adaptive Serious Game Settings

Anna Viktória Strbik, Dr. Bertalan Forstner
Department of Automation and Applied Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics
E-mail: strbika@edu.bme.hu, forstner.bertalan@aut.bme.hu

Abstract: *This paper explores a concept of personalized serious game configuration that adaptively adjusts game difficulty. Leveraging a neural network trained on players' performance data, the configuration of game parameters can be automated. The process involves three phases: data collection, data evaluation, and personalized recommendation generation. By iteratively applying these phases, changes in cognitive abilities can be detected and tailored game settings are provided continuously. Additionally, the method offers an adaptive game-based assessment to estimate cognitive abilities. This approach demonstrates a novel application of artificial intelligence in enhancing serious game design and cognitive assessment.*

Keywords: *serious game settings; neural network application; educational game configuration; cognitive assessment; adaptive game difficulty*

1 Introduction

The impact of digital games on cognitive development is supported by existing research suggesting positive correlation between usage of educational games and cognitive ability enhancement [1] [2]. Game-based learning further facilitates engaging and motivating environment, which augments learning experience. Notably, the willingness to persist in playing educational games is influenced by the perceived challenge [3]. Therefore, in designing serious games, it is essential to provide appropriate difficulty levels to maximize educational potential. Modifying game parameters likely provides insight into the parameters' relation to the game difficulty while enabling cognitive assessment [1]. Incorporating artificial intelligence to generate personalized game parameter settings, presents an innovative and effective approach [4].

2 Personalized Adaptive Game Configuration

An algorithm leveraging a specific neural network module was designed to automatically configure different game's parameters based on the participants' cognitive abilities and in-game achievements. The personalized configuration process comprises three phases: data collection, data evaluation and individualized game setting recommendation.

The objective of the first phase is to acquire sufficient data of game results of different players, and their relevant cognitive abilities. Normalized data is represented in three vectors capturing parameter values, individual ability values and the outcome of the game. This serves as training input to the neural network, resulting in each game's difficulty model to be represented by a specific trained neural network. The model facilitates personalized configuration based on the

user's cognitive abilities when the input of the AI-model is the expected outcome of the game and the player's abilities. In addition, it can estimate missing ability values. For this, game result and parameters are provided as input data.

The described phases can be applied iteratively which supports detecting changes in cognitive abilities. Remarkably, the assessment relies highly on training data. As a consequence, the scale on which ability values are measured is adaptive to the control group and flexibly changes with the actual achievements of players.

Acknowledgement

This work was supported by the European Union project RRF-2.3.1-21-2022-00004 within the framework of the Artificial Intelligence National Laboratory

References

[1] T. Tong, M. Chignell: "Developing a serious game for cognitive assessment: choosing settings and measuring performance", Proceedings of the Second International Symposium of Chinese CHI, Pages 70 - 79, April 2014

[2] F. C. Blumberg, K. Deater-Deckard, S. L. Calvert, R. M. Flynn, C. S. Green, D. Arnold, P. J. Brooks: "Digital Games as a Context for Children's Cognitive Development: Research Recommendations and Policy Considerations", Social Policy Report, Pages 1-33, April 2019, DOI: 10.1002/sop2.3

[3] L. Sun, M. Kangas, H. Ruokamo: "Game-based features in intelligent game- based learning environments: a systematic literature review", Interactive Learning Environments, February 2023, DOI: 10.1080/10494820.2023.2179638

[4] S. K. Banihashem, H. Dehghanzadeh, D. Clark, O. Noroozi & Harm J. A. Biemans: „Learning analytics for online game Based learning: a systematic literature review", Behaviour & Information Technology, September 2023, DOI: 10.1080/0144929X.2023.2255301

Different Type of Metrics to Evaluate Large Language Models

Tamás Pusztai, Péter Ekler
Budapest University of Technology and Economics, Departmento f Automation and Applied Informatics
E-mail: pusztai.tamas987@gmail.com, ekler.peter@vik.bme.hu

Abstract: *The Large Language Models (LLMs) are one of the most popular topic in nowadays. One of the first and biggest milestone in public expansion of Large Language Models was the introduction of ChatGPT toolset. Several information technology and industrial company started to research and apply the LLMs for the wide range of problems. But in that point it is worth to mention what are Large Language Models and what are they used for? One of the purpose of the article to make a quick introduction about Large Language Models and give some real life example about what they are used for in general. As their name indicates LLMs work on large datasets but it is worth to ask the question when a model is considered good performant and what criteria or so-called metrics can be used to evaluate the models. Since this task is not at all self-explanatory, it is worth examining several aspects that can be used to evaluate a model. In the recent years, extremely great progress has been made in terms of the development of the field of LLMs. Another purpose of this paper is to guide the reader through the various aspects that are most often used in the evaluation of large language models and give an insight into future development opportunities. The article introduces the basic concepts, some of the more important steps and elements of natural language processing and give a detailed description about the metrics used to evaluate the accuracy and performance of LLM models.*

Keywords: *llm; metrics; ai; evaluation*

References

[1] Murray Shanahan, "Talking about Large Language Models", Association for Computing Machinery, VOL. 67, 70-71, February 2024, 10.1145/3641526
[2] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, Wei Ye, Yue Zhang, Yi Chang, Philip S. Yu, Qiang Yang, Xing Xie, "A Survey on Evaluation of Large Language Models", Association for Computing Machinery, VOL. 15, 6-17, June 2024, 10.1145/3613609
[3] Cheng-Han Chiang, Hung-yi Lee, "Can Large Language Models Be an Alternative to Human Evaluations?", 3 May 2023, 2305.01937
[4] Yiheng Liua , Hao Hea , Tianle Hana , Xu Zhanga , Mengyuan Liua , Jiaming Tiana , Yutong Zhangb , Jiaqi Wangc , Xiaohui Gaod , Tianyang Zhongd , Yi Pane , Shaochen Xue , Zihao Wue , Zhengliang Liue , Xin Zhangb , Shu Zhangc , Xintao Hud , Tuo Zhangd , Ning Qianga , Tianming Liue and Bao Ge a, "Understanding LLMs: A Comprehensive Overview from Training to Inference", 2-22, 6 Jan 2024, 2401.02038
[5] Paul Thomas, Seth Spielman, Nick Craswell, Bhaskar Mitra, "Large Language Models can Accurately Predict Searcher Preferences", Association for Computing Machinery, 1-9, 11 July 2024, 10.1145/3626772

Estimating the difficulty level of puzzle games using artificial intelligence

Mercédesz Porkoláb, Bertalan Forstner
Department of Automation and Applied Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics
E-mail: 09dporkolabm@edu.bme.hu, forstner.bertalan@aut.bme.hu

Abstract: *This study presents a novel approach that utilizes a cognitive neural network to estimate the difficulty levels of cognitive games where NP-complete tasks make even an approximate estimation of the diffuculty cumbersome. The investigation emphasizes two primary aspects: the simulation of gameplay to produce extensive datasets for neural network training and the architecture of the cognitive neural network designed to achieve acceptable difficulty estimations. The results show that the combinated use of these methods provides a more accurate and consistent prediction of the difficulty level compared to the established methods.*

Keywords: *Convolutional neural network; Difficulty estimation; Cognitive games; Simulation; Machine learning*

1 Contributions

Research aimed at estimating the difficulty levels of cognitive development games has received considerable attention due to the fact that knowledge of the difficulty of the puzzles is essential from an educational and psychological point of view. Nilsson et al. (2017) used a convolutional neural network (CNN) to enhance difficulty assessment accuracy in Candy Crush Saga, achieving better results and reducing simulation time significantly [1]. Risi et al. (2021) highlighted the effectiveness of deep learning models for handling complex datasets in game difficulty estimation, emphasizing the need for robust datasets for consistent predictions [2]. Pusey et al. (2018) combined CNNs and reinforcement learning to capture gameplay patterns, demonstrating improved accuracy in difficulty estimation, though further optimization is needed for different game types [3]. Therefore, the open questions of the researches are mainly directed towards the production of training datasets and the appropriate neural network architecture. We developed a neural network model to estimate puzzle game difficulty by predicting the number of steps required to solve each puzzle. The model features input layers proportional to the 2D puzzle grid size, with convolutional and dense layers. This architecture is inspired by the hypothesis that the human brain processes board images at various abstraction levels.Key human strategies include placing the pieces in descending order of area, which has yielded the best results in our initial experiments. Other approaches involve the priorization of pieces based on their shape and the available spots. Our approach to estimating puzzle game difficulty was evaluated based on expert estimates, specifically those by László Mérő, together with other psychometricans and statisticians. We used a neural network model to predict the number of steps required to solve each puzzle with different typical strategies, comparing its predictions to expert-assigned difficulty levels. The quality of the results was validated against the difficulty classification established based on 4250 real gameplay solution statistics. The mean squared error (MSE) was used as the evaluation metric to measure the deviation of the

puzzles' difficulty order from the actual difficulty.The AI model achieved an MSE of 182, while the best expert's estimate had an MSE of 358. Additionally, AI naïve strategy had an MSE of 210, AI optimized strategy had an MSE of 300, and the average solution time had an MSE of 246. In contrast, the visual machine learning strategy had an MSE of 286. This significant reduction in MSE demonstrates the AI model's superior accuracy in estimating puzzle difficulty. The results indicate that our AI model can effectively estimate puzzle difficulty, outperforming traditional expert methods and suggesting potential for further refinement and application in various cognitive game scenarios.

Acknowledgement

Project no. TKP2021-NVA-02 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-NVA funding scheme.

References

[1] Nilsson, N., et al. "Estimating Cognitive Game Difficulty Levels Using AI." Linköping University. - 2017

[2] Risi, S., et al. "Deep Learning for Cognitive Game Difficulty Estimation." ACM Digital Library - 2021

[3] Pusey, S., et al. "AI Tools for Cognitive Game Difficulty Estimation." International Journal on Artificial Intelligence Tools. -2018

Fine Motor Skill Assessment Using Computer Vision

Panka Szalka, Dr. Bertalan Forstner
Department of Automation and Applied Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics
E-mail: szalka.panka@edu.bme.hu, forstner.bertalan@aut.bme.hu

Abstract: *This paper introduces a computer vision-based system to evaluate children's fine motor skills via the "Mondrian Blocks" educational game. By tracking movements during play, the system enables a data-driven assessment of these skills. Initial testing in a primary school classroom has provided promising insights and suggested potential for further development in child education and development.*

Keywords: *Computer Vision; Fine Motor Skills; OpenCV; MediaPipe; Child Development*

1 Introduction and Related works

Fine motor skills are crucial for a child's daily activities and are typically evaluated through observation and physical activities. Our study introduces a novel computer vision-based system using the "Mondrian Blocks" game for comprehensive data collection, which will aid further research in this area. Prior research such as Bebis, Egbert, and Shah [1] has provided a foundation for the application of computer vision in education. The application of computer vision to board games by Ren et al. [2], and Devyatkov and Alfimtsev's [3] techniques for enhancing human-computer interaction in games with computer vision, further emphasize the potential of such applications in various domains.

Figure 1
Recognised elements on the board

2 Contributions

We present a system for evaluating children's fine motor skills by combining computer vision with the "Mondrian Blocks" game. This user-friendly system autonomously initializes, recognizes the game board, divides it into fields, and assigns game colors based on proximity, eliminating the need for manual calibration or predefined setups, with the help of OpenCV and

MediaPipe technologies. The algorithm tracks the game board, archiving the current state for analysis if no changes or hand movements are detected for a certain period.

To validate our proposed system, we conducted an evaluation in a primary school environment. The students were engaged in the "Mondrian Blocks" game, and our system concurrently tracked and recorded their movements. In addition to the digital tracking, manual data collection was performed, offering a comprehensive assessment of fine motor skills. The data collected from this evaluation can be combined and analyzed, revealing deeper insights into child development and education.

Our evaluation demonstrated the potential of this system as an effective real-world tool for tracking, analyzing, and enhancing fine motor skills in children.

Acknowledgement

This work was supported by the European Union project RRF-2.3.1-21-2022-00004 within the framework of the Artificial Intelligence National Laboratory.

References

[1] G. Bebis, D. Egbert, and M. Shah, “Review of computer vision education,” IEEE Transactions on Education, vol. 46, no. 1, pp. 2–21, 2003.
[2] J. Ren, P. Astheimer, and I. Marshall, “A generic framework for vision based interactive board games,” in GAME-ON, 2003.
[3] V. Devyatkov and A. Alfimtsev, Human-Computer Interaction in Games Using Computer Vision Techniques, vol. 3, pp. 1210–1231. 05 2013.

Expressive Text-to-Speech Through Tales in Modern Standard Arabic

Rami Kammoun, Mohammed Salah Al-Radhi
Department of Automation and Applied Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics

Abstract: *Text-to-speech or speech synthesis has seen a wide evolution over the past decades, from the articulatory methods to the statistical means until today's deep learning-based solutions. Autoregressive solutions have pushed the limits with their ability to learn and keep track and alignments for longer phrases. An issue that was later on solved by Non-Autoregressive models where computation has run in parallel and helped in faster inference and longer accurate phrase generation thanks to the attention mechanism. Furthermore, other features that help predict more natural speech were extracted such as duration, pitch, and energy. All of the latter state-of-the-art models were built to more robust systems which proved highly effective for rich resource languages, especially English with its abundant high-quality audio recordings with their corresponding already clean transcription. For this work of research, we primarily focus on the Arabic language, notably, the Modern Standard Arabic (MSA). For a long time, the Arabic language has been considered a low-resource language, whereas lately, some advancements have been recorded either through newly published datasets or systems. We present in the following paper the first ever developed Deep Learning-based expressive TTS system in MSA. In previous research, expressive Arabic speech was synthesized using concatenative methods which proved unnatural and quality degrading over time. Despite recent developments, the Arabic language with its varieties still lags behind rich resources, notably English. Our goal is to retrain the already sourced Arabic datasets such as ASC (Arabic Speech Corpus), ClArTTS, and our newly natural speech dataset on more state-of-the-art systems like Arabic Text and Speech Transformer (ArTST), Expressive FastSpeech 2, and VITS, and test whether undiacritized speech could actually lead to better results considering the low nature of the latter dialect. To create our own expressive dataset, we automate a whole pipeline of trimming, splitting on silence, Arabic speech recognition, then, automatic diacritization. Another fact to consider is the effectiveness of self-supervised models such as VITS in bridging the gap between low and high-resource languages through MOS score evaluation, but we do not expect it to give better results seeing the small amount of data that we work with. Seeing that ArTST has already given much better results for MSA, fine-tuning it on the expressive few minutes of gathered speech would be the most effective in giving the highest values of subjective evaluation. The latter would represent the first step of generating the first natural expressive TTS models for Modern Standard Arabic.*

Keywords: *Text-to-speech; low-resource; autoregressive; self-supervised; expressiveness*

Question-answering on top of telecom related documents: successes and pitfalls

Máté Ákos Tündik[1], Márk Blaskó[2] and Ferenc Kovács[3]
Nokia
E-mail: { [1] akos.tundik, [2]mark.blasko}@nokia.com, [3] ferenc.2.kovacs@nokia-bell-labs.com

Abstract: *Nowadays, generative AI and especially LLM-based chatbots are becoming a commonly used toolset. Creating effective chatbots can be challenging for certain domains because of their own specialized rules, jargons or needs. In this paper we elaborate how to handle these specialties in case of telco related chatbots.*

Keywords: *Retrieval Augmented Generation, Domain specific Chatbot, LLM*

1 Building a telco chatbot

Industrial systems have complex product documentations, involving dozens of topics, hence it is difficult to find the right answer to a specific question. As leveraging generative AI is getting popular for many fields it raises the question whether it is possible to build a chatbot for handling domain-specific product documentations (in our case, for telecom industry) to relieve the daily job of product architects. The product for the case study is Nokia Telephony Application Server.

Telco product documentations have their own specific terms, containing plenty of abbreviations and images. Handling of these specialties raises some challenges such as: (1) handling and resolving acronyms and abbreviations, (2) multimodal input and output, (3) necessity of model fine tuning and (4) handling user feedback.

Figure 1 shows our RAG (retrieval augmented generation) based solution including acronym handling, image retrieval and user feedback. The GPT 3.5 model of OpenAI was chosen for text generation. It is hosted on MS Azure platform and accessible via REST API. Using 3rd party tools on top of internal documentation raises legal and compliance issues; EU AI Act must be followed [1]. Thus, some documents were excluded and questions that related to this part of documentation cannot be answered. The documents are chunked, then their text embeddings (BGE [2]) stored in a Milvus database, that are looked up during the retrieval. The fine- tuning experiments of the retrieval and the generation models are ongoing.

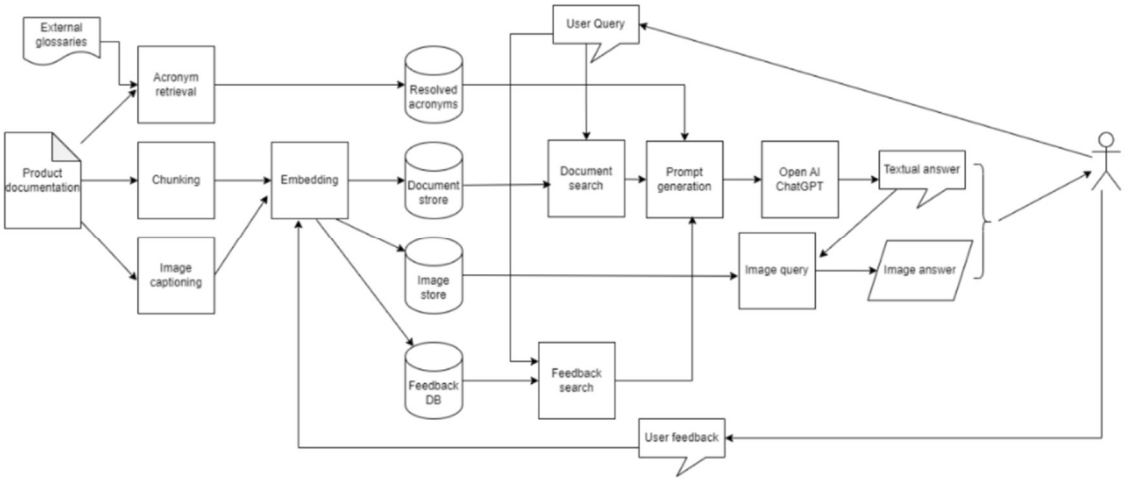


Figure 1
System architecture of our telco chatbot

A hybrid method is applied for resolving abbreviations as there are generic telco specific and product specific ones. External glossaries are available for the former ones, while the latter ones are extracted directly from product documentations. The acronym resolution is a recursive task as there are acronyms, which are based on acronyms. Finally, the most relevant abbreviations are added to the system prompt.

The product documentations contain several images, holding a lot of useful information for the users. However, these images have diverse functions, e.g., they describe the architectural design or the sequence diagrams support use cases. In our case, the images have descriptive captions; as a baseline, only the captions were used as a basis of text embeddings, applying the same BGE model as for the documentation chunks. The image retrieval is based on the generated answer. The retrieval engine tries to find the most fitting image to the answer, using cosine distance. Utilizing multimodal embeddings is an open direction for the future.

Collecting anonymous user feedback is a continuous task. Our solution keeps the possibility of collecting feedbacks in two ways: with simple thumbs up/down votes for the correctness of the generated answer and writing free-text comments. When a new question arrives, the system can find for the best matching previous questions, which has positive feedback, and it may be included into the generated prompt.

Acknowledgement

We thank to our colleagues participating in the subjective evaluation of the chatbot.

References

[1] Outeda, Celso Cancela. "The EU's AI Act: A Framework for Collaborative Governance.", Internet of Things, 101291., 2024
[2] Xiao, Shitao, et al. "C-pack: Packaged resources to advance general Chinese embedding.", arXiv preprint arXiv:2309.07597, 2023.

Enhancing Connectome Analysis through Multimodal Neuroimaging Fusion and Machine Learning Techniques

Artyom Puskin, Péter Ekler
Department of Automation and Applied Informatics, Budapest University of Technology and Economics
E-mail: puskin.artyom@edu.bme.hu, ekler.peter@vik.bme.hu

Abstract: *This paper looks at how combining different types of brain scans with machine learning can improve our understanding of the brain's connectome, which maps the connections between different brain regions. We review methods for merging data from structural MRI (sMRI), functional MRI (fMRI), diffusion MRI (dMRI), computed tomography (CT), positron emission tomography (PET), and electroencephalography (EEG). Our focus is on how these methods enhance the detailed mapping of brain networks. We discuss three main fusion approaches: feature - level, decision - level, and deep learning - based approaches and their roles in connectome analysis. We also explore practical applications and suggest ways to advance this research in the future.*

Keywords: *multimodal neuroimaging fusion; diagnostic accuracy enhancement; machine learning in medicine*

1 Introduction

The brain connectome is a detailed map showing how different brain regions are connected. Traditional brain imaging techniques provide useful but often incomplete views of these connections [1]. By combining different imaging methods, we can get a fuller picture of the connectome. Machine learning helps in merging this data to reveal more about how brain regions interact. This paper explores how integrating data from various imaging modalities improves our understanding of the brain's network and how these methods can be applied in real-world scenarios.

2 Background

The brain connectome includes both the physical connections between brain regions and their functional interactions. Structural MRI (sMRI) gives information about brain anatomy, while diffusion MRI (dMRI) shows the pathways of white matter. Functional MRI (fMRI) reveals brain activity, and electroencephalography (EEG) measures electrical signals in the brain [2]. Combining these imaging methods helps create a more complete map of brain connectivity. For instance, merging sMRI and dMRI data can show how different brain areas are physically connected [3], while combining fMRI and EEG data can reveal how these connections are used during brain activity [4].

3 Fusion Techniques

Different techniques are used to combine multimodal neuroimaging data for better connectome analysis. These can be grouped into three main types:

Feature-Level Fusion: This method combines features from different imaging types into one dataset. For connectome analysis, this might involve merging sMRI, fMRI, and dMRI data to create detailed maps of brain connections. Techniques like concatenation or merging images at the pixel level help integrate the information [5, 6].

Decision-Level Fusion: This approach combines the results from multiple models, each trained on different types of data, to make a final decision. In connectome analysis, decision-level fusion might involve combining predictions from models that analyze sMRI, fMRI, and EEG data separately. Methods like voting or averaging the results can improve accuracy [6]. Deep Learning-Based Fusion:

Deep learning methods, such as those using attention mechanisms or hierarchical models, integrate data at different levels. These techniques can focus on important features from various imaging types and combine them to build a more accurate connectome map. They are useful for capturing complex connections between brain regions [6].

4 Conclusions

This paper shows how combining different types of brain imaging data with machine learning can enhance our understanding of the brain connectome. By merging data from sMRI, fMRI, dMRI, CT, PET, and EEG, we can create more detailed maps of brain connections. Feature-level, decision-level, and deep learning-based fusion methods each provide unique benefits for connectome analysis. Advancing these techniques will improve our knowledge of brain connectivity and its implications for diagnosing and treating neurological and psychiatric disorders.

References

[1] R. E. Marc, B. W. Jones, C. B. Watt, J. R. Anderson, C. Sigulinsky, and S. Lauritzen, "Retinal connectomics: Towards complete, accurate networks," Progress in Retinal and Eye Research, vol. 37, pp. 141–162, 2013.
[2] S. Liu, W. Cai, S. Liu, F. Zhang, M. Fulham, D. Feng, S. Pujol, and R. Kikinis, "Multimodal neuroimaging computing: a review of the applications in neuropsychiatric disorders," Brain Informatics, vol. 2, pp. 167–180, Sep 2015. Epub 2015 Aug 29.
[3] L. Tan, Y. Chen, T. C. Maloney, M. M. Caré, S. K. Holland, and L. J. Lu, "Combined analysis of smri and fmri imaging data provides accurate disease markers for hearing impairment," NeuroImage: Clinical, vol. 3, pp. 416–428, 2013.
[4] P. A. Valdes-Sosa, J. M. Sanchez-Bornot, R. C. Sotero, Y. Iturria-Medina, Y. Aleman-Gomez, J. Bosch-Bayard, F. Carbonell, and T. Ozaki, "Model driven eeg/fmri fusion of brain oscillations," Human Brain Mapping, vol. 30, no. 9, pp. 2701–2721, 2009.
[5] J. Sui, T. Adali, Q. Yu, J. Chen, and V. D. Calhoun, "A review of multivariate methods for multimodal fusion of brain imaging data," Journal of Neuroscience Methods, vol. 204, no. 1, pp. 68–81, 2012.
[6] Y. Li, M. El Habib Daho, P.-H. Conze, R. Zeghlache, H. Le Boité, R. Tadayoni, B. Cochener, M. Lamard, and G. Quellec, "A review of deep learning-based information fusion techniques for multimodal medical image classification," Computers in Biology and Medicine, vol. 177, p. 108635, 2024.

Investigating the scaling properties of Self-Supervised Pretraining for Segmentation

András Kalapos, Bálint Gyires-Tóth
Department of Telecommunications and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics
E-mail: kalapos.andras@tmit.bme.hu

Abstract: *Self-supervised learning (SSL) is a powerful technique for pretraining neural net- works on unlabeled data, but its effectiveness on small-scale, specialized datasets is lim- ited. We investigate the properties of SSL methods on small-scale segmentation datasets and propose methods to evaluate and analyze their performance. Our results show that domain- specific SSL pretraining can be an efficient approach, and offer a promising compromise be- tween pretraining cost and downstream accuracy.*

Keywords: *Self-Supervised Learning; Semantic Segmentation; Pretraining; Scaling*

1 Introduction

Self-supervised learning (SSL) is a powerful technique for pretraining neural net- works on unlabeled data. It allows models to learn useful features, which are then fine-tuned on labeled data for specific tasks. This approach can improve accuracy and efficiency in situations where labeled data is scarce and expensive to obtain, especially in cases where legal constraints or special properties of the data prevent the use of public pre-trained models. Such scenarios are common in specialized domains where the data is private or the task is non-standard. We refer to this approach as domain-specific SSL pretraining, as opposed to generalist SSL pre- training on large-scale datasets that are independent of the downstream data. However, SSL methods are expected to learn general, task-agnostic features, which requires a large amount of data to be effective. With specialized, private datasets, this is not always guaranteed, limiting the performance achievable by domain- specific SSL. Moreover, most SSL methods are primarily developed and tested on the ImageNet [1] classification dataset and, by design, rely on the object- centric properties of such datasets. This often makes SSL pretraining challenging on non-object-centric tasks, such as scene understanding [2]. Additionally, com- mon SSL methods learn image-level features, which may not be optimal for tasks that require localized features, such as object segmentation. These trade-offs and challenges are important to address to enable and wider adoption of SSL.

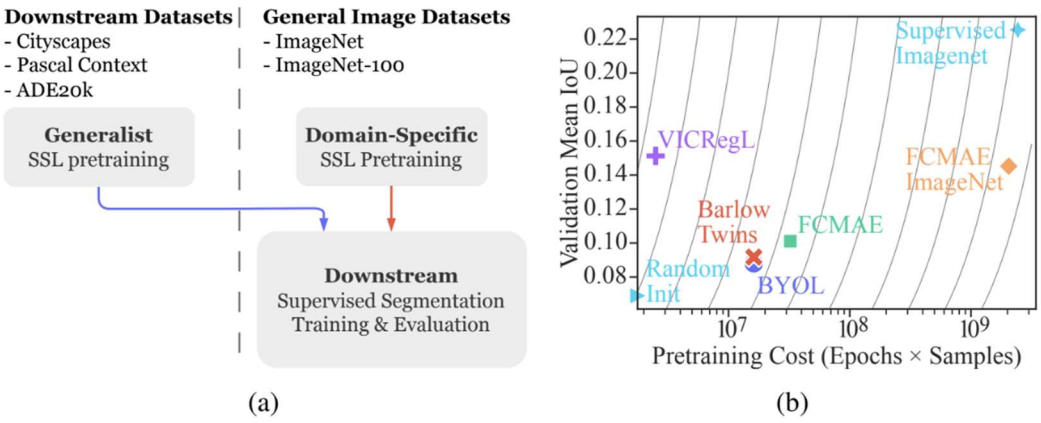


Figure 1
(a) Different pretraining strategies that we investigate. (b) Our results comparing SSL methods based on accuracy and computational cost on the ADE20k dataset. Lines represent constant Accuracy to Pretraining Cost ratio, with lower ratio values towards the left and up.

2 Methods

In this research, we propose methods to evaluate and analyze common SSL algo- rithms across several tradeoff questions: pretraining dataset scale, object centricity, and downstream tasks requiring localized features. Our examination includes pre- training computational cost, downstream accuracy, and convergence properties to assess the quality of learned features. We conduct experiments on Cityscapes, Pas- cal Context, ADE20k, ImageNet100, and ImageNet-1k [1]. We adapt a variety of SSL methods to these datasets and the segmentation task, including BYOL [3], BarlowTwins [4], FCMAE [5], and VICRegL [6]. We present our method in fig. 1(a) and report our results in fig. 1(b). The latter shows that, although SSL methods such as VICRegL provide lower final accuracy, their pretraining cost is orders of magnitude lower compared to pretraining on a large-scale dataset like ImageNet.

3 Conclusions

Our research investigates the properties of SSL methods on segmentation datasets, specific to the downstream task. We show that self-supervised pretraining on domain-specific data can be an effective, despite potential trade-offs in accuracy.

Acknowledgement

We acknowledge KIFÜ (Governmental Agency for IT Development, Hungary, https://ror.org/01s0v4q65) for awarding us access to the Komondor HPC facility based in Hungary.

References

[1] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255, 2009.

[2] S. Purushwalkam and A. Gupta, “Demystifying Contrastive Self-Supervised Learning: Invariances, Augmentations and Dataset Biases,” in Advances in Neu- ral Information Processing Systems, vol. 33, pp. 3407–3418, Curran Associates, Inc., 2020.

[3] J.-B. Grill, F. Strub, F. Altché, C. Tallec, P. Richemond, E. Buchatskaya, C. Do- ersch, B. Avila Pires, Z. Guo, M. Gheshlaghi Azar, B. Piot, k. kavukcuoglu, R. Munos, and M. Valko, "Bootstrap Your Own Latent - A New Approach to Self-Supervised Learning," in Advances in Neural Information Processing Sys- tems (H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, eds.), vol. 33, pp. 21271–21284, Curran Associates, Inc., 2020.
[4] J. Zbontar, L. Jing, I. Misra, Y. LeCun, and S. Deny, "Barlow Twins: Self- Supervised Learning via Redundancy Reduction," in Proceedings of the 38th International Conference on Machine Learning (M. Meila and T. Zhang, eds.), vol. 139 of Proceedings of Machine Learning Research, pp. 12310–12320, PMLR, July 2021.
[5] S. Woo, S. Debnath, R. Hu, X. Chen, Z. Liu, I. S. Kweon, and S. Xie, "Con- vNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders," Jan. 2023.
[6] A. Bardes, J. Ponce, and Y. LeCun, "VICRegL: Self-Supervised Learning of Local Visual Features," Oct. 2022.

Enhancing Auditory Component Detection in EEG/MEG Data Using Time-Frequency Deep Networks

Mohamed F. Issa[1*,3], Izhar Khan[1], Manuela Ruzzoli [1,2], Nicola Molinaro [1,2] & Mikel Lizarazu[1]
[1]BCBL, Basque Center on Cognition, Brain and Language, San Sebastian, Spain
[2] Ikerbasque, Basque Foundation for Science, Bilbao, Spain
[3] Faculty of Computers and Artificial Intelligence, Benha University, Benha, Egypt
E-mail: m.issa@bcbl.eu

Abstract: *This paper introduces Time-Frequency Deep Networks (TFDN)to detect auditory-related components in EEG/MEG data through Independent Component Analysis (ICA) for advancing neuroscientific research and clinical diagnostics.*

Keywords: *Time-Frequency Deep Network; Independent Component Analysis*

1 Introduction

ICA can capture signature information on how the brain processes and interprets spoken language, providing crucial insights into neural mechanisms involved in language interpretation and response. Traditionally, identifying ICA components associated with speech processing has been manual, time-consuming, and prone to error [1]. Nonetheless, time-frequency distributions have been extensively utilized in various applications, including deep learning-based speech assessments for patients with aphasia, demonstrating their effectiveness and broad utility [2].

2 Methods

1.1 Data Collection and Preprocessing

Ten native French-speaking adults with normal hearing participated in EEG and MEG recordings at CUB Hospital Erasme in Brussels under ethical approval, followed by preprocessing that included filtering, artifact removal, and ICA, setting the stage for applying the TFDN model.

1.2 Time-Frequency Deep Network (TFDN) Architecture

We implemented a Time-Frequency Convolutional Network (TFCN) for EEG classification using a scalable architecture that integrates a different scalogram layer from the continuous wavelet transform. This network includes convolution, max- pooling, batch normalization, and softmax layers optimized using the Adam optimizer.

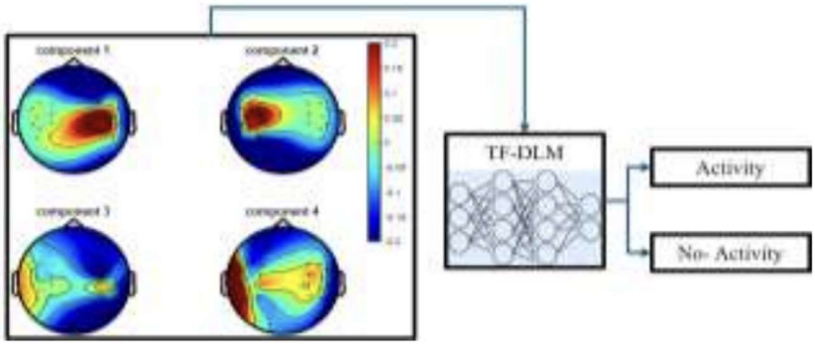


Figure 1
Components data passed to the Time-Frequency Deep Learning Model (TF-DLM).

3 Results

The TFDN model achieved an accuracy of 90 % for EEG data and 93% for MEG data in classifying components related to auditory activity. This high level of accuracy highlights the potential of TFCN as a robust tool for neuroimaging data analysis.

Acknowledgement

The study was backed by the Basque Government, the AEI (Severo Ochoa accreditation, and MICIU's project. M.L. and M.R. are supported by MICIU's Ramón y Cajal program and Ikerbasque.

References

[1] M. F. Issa, G. Tuboly, G. Kozmann, and Z. Juhasz, "Automatic ECG Artefact Removal from EEG Signals," Measurement Science Review, vol. 19, no. 3, 2019, doi: 10.2478/msr-2019-0016.

[2] A. Kumar, S. S. Mahmoud, Y. Wang, S. Faisal, and Q. Fang, "A Comparison of Time-Frequency Distributions for Deep Learning-Based Speech Assessment of Aphasic Patients," International Conference on Human System Interaction, HSI, vol. 2022-July, 2022, doi: 10.1109/HSI55341.2022.9869452.

Modeling neural networks with hypergraphs via tensor-based views

Csaba Hajdu[1], Ádám Csapó[2]
[1]Department of Informatics, Széchenyi István Egyetem
[2]Institute of Data Analytics and Information Systems, Corvinus University of Budapest
E-mail: hajdu.csaba@ga.sze.hu, adambalazs.csapo@uni-corvinus.hu

Abstract: *Virtual systems and physical simulations are becoming increasingly widespread in creating digital twins for industrial and consumer applications. However, these systems often require complex layouts and scene descriptions involving special tools. Additionally, simulations usually inhibit complex relationships between elements, which are complicated to describe using traditional methods. This paper introduces a new approach to physical modeling by leveraging hypergraphs to describe high-order relationships between physical elements. This method offers a more intuitive and efficient way to represent complex physical structures, streamlining the modeling process.*

Keywords: *virtual systems; digital twins; robotic systems; simulations*

1 Introduction

Virtual systems are widely used to validate, verify, and, more recently, interact with (through digital twins) real physical systems. A persistent challenge in deploying such systems is the complexity of the scene-creation process. Furthermore, these descriptions include complex relationships that are difficult to describe with currently available widespread description formats, even including many-to-many relationships (i.e., describing references and relationships).

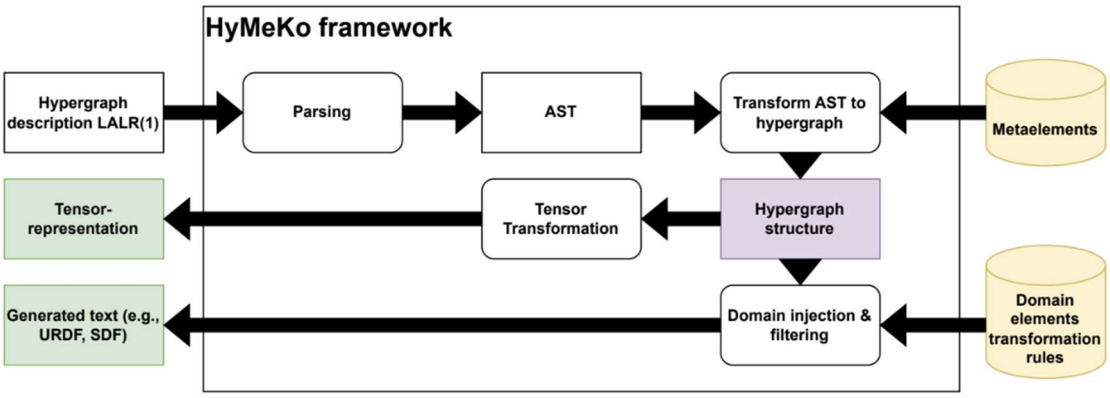


Figure 1
Tensor and text generation using the framework.

This paper introduces a novel approach to describing scene descriptions of environments in simulations, offering a description and modeling methodology that remains independent of the final output format. This flexibility allows the same scene description to be reused across

multiple simulation and virtual reality platforms (e.g., Gazebo, Maxwhere[1]), significantly streamlining the process and reducing redundancy. Current formats, like URDF, use XML, leading to a verbose and rigid format [2], even with the use of XML-macros (XACRO) [3]. The framework and the methodology is based on the theory of hypergraphs [4].

2 Methodology

In the proposed framework, the scene model is described with the help of the HyMeKo description (the authors' LALR-based framework1), that describes kinematic structures using complex relationships, like kinematic loops. The language describing graph relationships are more concise than XML or YAML based languages. An instance is then parsed into an abstract syntax tree (AST) and transformed into the hypergraph structure, with meta elements injected of a specific domain. The hypergraph structure can transform into a text-based representation (i.e., the application input format). The possible tensor-based representation is not as emphasized in the current scenario, but it can give insight into the scene structure using algebraic methods. The process overview is depicted in Figure 1.

The text generation in the hypergraph works by traversing the transformed hypergraph and filtering on the elements required by the target domain. For example, in the case of kinematic descriptions, the kinematic link and joint elements are required (all defined as part of the meta elements, also a hypergraph-based model). To enhance the traversal and filtering process, a hypertree compositional hierarchy is maintained in the organization of the elements: that is, a hypernode and the hyperedge are composable objects, capable of containing other elements (a property that is exhibited in the generalizations of hypergraphs). In the case of kinematic elements, the following mappings are defined to describe scene trees and kinematic structures: **Hypernodes** represents scene objects, kinematic links, geometries, and inertial parameters, while **hyperedges** represents joints and references (e.g., to materials). Element types are given as stereotypes so a node can exhibit kinematic link properties to allow filtering during traversal. The references are unfolded during text generation, as the description allows the reuse of elements (i.e., in the case of shared visual & collision geometries). Some results is shown in Figure 2.

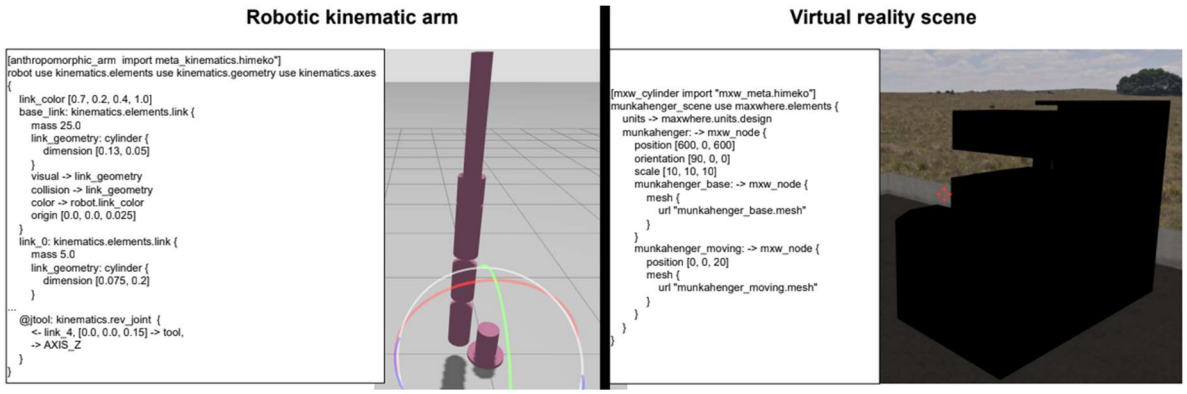


Figure 2
Scene generation result (robotic arm and virtual reality scene respectively)

3 Conclusions

This paper introduced a methodology to describe scenes for simulators (e.g., Gazebo) and virtual reality systems (e.g., MaxWhere). The methodology allows the description of complex

relationships, such as kinematic loops, as there are precise mappings of physical and geometrical elements to hypergraph elements, while providing more concise description than XML or YAML. The text generation is resolved using a simple graph traversal method by fixing the hierarchical composition in a tree-based fashion. The generated descriptions are runnable in target simulators while maintaining the model separately from the application description. Future work includes creating more complex descriptions, like humanoid robots or industrial scenes, and investigating the possibilities of validation analysis of physical structures based on the algebraic properties of the tensor view.

References

[1] A. Csapo, A. Sudár, B. Berki, B. Gergely, B. Berenyi, and C. Czaban, "Measuring Virtual Rotation Skills in MaxWhere," Sept. 2020.

[2] D. Tola and P. Corke, "Understanding URDF: A Survey Based on User Experience," 2023. _eprint: 2302.13442.

[3] N. Albergo, V. Rathi, and J.-P. Ore, "Understanding Xacro Misunderstandings," ArXiv, vol. abs/2109.09694, 2021.

[4] A. Bretto, Hypergraph Theory: An Introduction. Springer Publishing Company, Incorporated, 2013.

Driving Digital Transformation: Data Analytics, AI, and Business Development Solutions for SMEs and the Public Sector

Andras Hajdu, Attila Tiba
Faculty of Informatics – DATA-EDIH Consortium, University of Debrecen

Abstract: *Our competence provides a comprehensive approach to enhancing the digital maturity of SMEs and the public sector. By integrating advanced technologies, conducting tailored consultations, and fostering collaborations through matchmaking, we aim to drive sustainable economic growth and improve operational efficiency for businesses.*

Keywords: *Digital transformationm, SMEs, AI-based solutions*

The integration of digital technologies is essential to maintaining and enhancing competitiveness. The goal of the Data-EDIH project is to develop the digital maturity of small and medium-sized enterprises (SMEs) and the public sector, with a particular focus on the healthcare and food industries. Our services include data asset assessment, data analysis, data visualization, digital security, and the implementation of AI-based solutions. Our competencies also include conducting consultations to support digital transformation, organizing digitalization training and workshops, and supporting the practical application of advanced data analytics methods. Additionally, we provide business consultancy for SMEs, including capital raising preparation and digital business development. Our aim is to increase the innovation capacity and efficiency of businesses, thus promoting sustainable economic growth.

Our approach to digital transformation emphasizes not only technological implementation but also the alignment of these solutions with the strategic objectives of the organizations. We conduct thorough initial assessments to understand the specific needs and challenges faced by SMEs and public sector entities, and through tailored consultations, we facilitate a seamless transition into the digital phase. Through matchmaking services, we help SMEs connect with each other or even with larger corporations, fostering partnerships that lead to mutual growth. By focusing on sustainable, scalable solutions, we help businesses improve their internal workflows and market competitiveness, enabling them to adapt to evolving market conditions and regulatory demands.

As part of our comprehensive approach, we employ living labs and testbeds to allow businesses to experiment with new digital technologies in real-world settings before full-scale implementation. This environment accelerates innovation by providing a controlled space to validate solutions, ensuring that digital transformation efforts are not just a temporary boost but a foundational change that drives continued success and innovation. By leveraging data-driven insights, AI tools, and robust business strategies, we enable organizations to not only improve operational efficiency but also create long-term value, leading to sustainable digital growth.

Acknowledgement

This project was supported by the Government of Hungary and the Parliamentary State Secretariat of the Ministry of Public Administration and Regional Development. The Data EDIH consortium was granted a budget support of HUF 720,000,000 under the Government Decision 1548/2023 (XII. 13.) on the provision of the own contribution costs of certain direct EU tenders.

Smart Learning Platforms: Combining AI and Moodle for Enhanced Learning Experiences

Zoltán Király, Tamás Molnár
Multimedia and E-learning Technical Center University of Debrecen

Abstract: *The use of artificial intelligence has become a part of our everyday lives and its application in education is also spreading, especially the ChatGPT. This work aims to develop a chatbot prototype and connect it to Moodle via REST API services, allowing the chatbot to access course materials and provide assistance to students.*

1 The artificial intelligence can be utilized in various educational contexts in several ways.

The chatbot assists in preparing outlines for presentations by generating key points. It is also capable of summarizing essential parts of lengthy texts. The ChatGPT can categorize textual data based on mood, theme, intent and emotion as well as detect unwanted content, such as plagiarism or hate speech. The chatbot can extract relevant keywords from texts, helping students identify the main themes and concepts. ChatGPT can also generate thought-provoking discussion questions tailored to specific topics, encouraging student engagement and critical thinking. This AI tool can be used in pair assessments (like a Moodle Workshop). The teacher gives the topic of assignment. The student sends the prompt to the ChatGPT and then evaluates the ChatGPT's completion and explains why the answer is right or wrong.

2 Chatbot for education is an important use of ChatGPT, the bots provide new ways for students to interact with learning platforms

The general chatbot serves guest users who do not need to authenticate. It provides quick answers to frequently asked questions about courses and the portal. These bots can guide users through login procedures, password resets and other common inquires (FAQs). This service provides automatic responses in 24/7 hours and making teachers' work easier.
One unique feature of the chatbots is their ability to interact with course materials. It can provide activities, self- assessment quizzes as well as explains the glossary terms.
Through the "chat with your documents" service, students can query the bot about specific topics, receiving relevant explanations directly from the course content.
The tutorbot, on the other hand is designed for authenticated users and acts as a personal assistant tailored to individual needs and competency levels. This chatbot acts as a personal learning "mate", guiding students along personalized learning paths. It sends notification if the latest quiz result was insufficient or alerts students if they haven't logged in for a while. The tutorbot is connected to the course calendar, allowing it to access event information and send alerts or reminders. It can notify students about upcoming assignments, quizzes and deadlines.

3 Various possibilities of linking Moodle and ChatGPT

ChatGPT chatbot is a large language model optimized for chat. It uses artificial intelligence to analyze and reproduce natural language, imitating human communication at a high level and providing realistic and intelligent answers to user questions (prompts).
Moodle is one of the most popular Learning Management Systems (LMS) in the world, widely used by various universities and millions of users worldwide. At our university, we also use the Moodle system, so we pay special attention to linking Moodle and ChatGPT.

4 Using the Moodle plugin is one way to create the Moodle – ChatGTP connection

Moodle Plugin is an application that extends the functionality of Moodle. The “OpenAI Chat Block” plugin provides an easy way to integrate ChatGPT into Moodle. Once installed, this plugin allows to add the new “OpenAI Chat” block to the courses. This block includes the “Source of truth” text area, where teachers can insert the special context of the course. This helps the bot provide more accurate responses. However, the language model's token limit restricts the amount of context text that can be added, this limit can be avoided by the Retrieval-Augmented Generation (RAG) method, which will be discussed later.

5 Moodle REST API web services and Python Client (webhooks)

By utilizing the Moodle REST API web services alongside a Python client with webhooks, an efficient connection between Moodle and ChatGPT can be established. The method acts as a bridge between Moodle and ChatGPT, making AI capabilities available within the platform.
A user's request, initiated through the Moodle course interface is transferred by the Python Client to ChatGPT. The response (or completion) from ChatGPT, along with custom content retrieved from a vector database, is then displayed in the Moodle course.
The Moodle web service offers a wide range of functionalities, enabling users to interact with various aspects of the platform. For example, the get_site_info function retrieves essential information about the Moodle instance, such as its name or version.
To manage courses, several functions are available, including **course_get_categories, course_get_courses**, etc. which enable users to retrieve course-related information. Additionally, the web service allows users to create new users on the Moodle site and enroll them in courses.
The web service also includes functions related to quizzes, such as **mod_quiz_get_quizzes_by_courses** and **mod_quiz_get_user_attempts**, which allow instructors to utilize quizzes effectively within their courses. These functions enable users to retrieve information about quizzes and access user attempts on these quizzes. To keep track of important dates and activities, the function **core_calendar_get_calendar_events** enables users to check and list calendar events, ensuring that they stay informed about upcoming deadlines and events within the Moodle environment. The mod_glossary_get_entries_by_search function enables users to search for glossary terms and retrieve their corresponding definitions.

6 Chat with your own documents using ChatGPT

This is a very useful AI service for education. Large language models usually answer correctly questions that they have encountered many times during their teaching. The limitation of the service available on the ChatGPT - OpenAI website or in the "OpenAI Chat Block" is that only

a limited amount of specific content (context) can be provided, so it cannot answer specific questions.
This limit can be removed by using the embedding and Retrieval-Augmented Generation (RAG) method. Embedding converts text into numbers. The language model will then be able to measure the similarity of texts and find text patterns. For each word in the context, it will take into account nearby words. The probabilities are stored in the form of vectors and the distance between two vectors measures the similarity of the texts.

6 RAG = Retrieval-augmented generation

RAG is a technique for augmenting LLM knowledge with additional data to get more accurate answers. RAG integrates GPT with databases. The custom context is extracted from various sources, such as the user's individual documents, Moodle pages, Moodle books and web pages then converted into embeddings (numerical representations). The embeddings are stored in a vector database.
When a user submits a prompt, RAG creates an embedding of that prompt and compares it with the stored embeddings to retrieve the most relevant contexts from the vector database. Based on the extracted content, the GPT generates a new prompt that combines the user prompt, the selected contexts and the conversational memory. This enhanced "super" prompt is then sent to ChatGPT, which generates the final response.

7 Summary

By leveraging ChatGPT's capabilities, educators can enhance teaching and learning experiences, making education more effective and engaging. The aim of this work was to develop a prototype chatbot and connect it to Moodle via REST API services, enabling the chatbot to access course learning materials and provide assistance to students. The chatbots have the potential to significantly improve e-learning experiences by offering personalized support, efficient information access, and a more engaging learning environment. By leveraging the capabilities of both general and tutorbots, educational institutions can create a more effective and enjoyable learning journey for their students.

Cognitive Schema Fitting (CSF): An Autonomous Multi-Agent AI Framework for Dynamic Prompt Engineering in Large Language Models

Gábor Orbán, Nimród Mike
Corvinus University of Budapest, Institute of Data Analytics and Information Systems
E-mail: gabor.orban1@stud.uni-corvinus.hu, nimrod.mike@uni-corvinus.hu

Abstract: *We introduce Cognitive Schema Fitting (CSF), a novel multi-agent AI framework that enhances large language models by decomposing tasks into subproblems and applying adaptive cognitive schemas for their solution. Agents iteratively refine outputs, dynamically switching schemas to improve accuracy. Future work will focus on optimizing schema selection and validating performance gains.*

Keywords: *cognitive schema fitting; multi-agent systems; large language models; prompt engineering.*

1 Introduction

Despite significant advancements in large language models (LLMs) and prompt engineering inspired by human cognition, AI systems struggle to emulate the adaptive problem-solving capabilities inherent in human thinking, particularly for complex tasks. This limitation hinders AI efficiency and precision across diverse contexts [1–3].

2 Cognitive Schema Fitting (CSF)

We propose Cognitive Schema Fitting (CSF), a multi-agent AI framework that emulates human cognitive adaptability [4]. CSF decomposes user inputs into subproblems and assigns each to an AI agent representing a specific cognitive schema [5]. Agents iteratively refine their outputs, dynamically switching schemas to enhance solution accuracy, mirroring human strategy shifts when facing challenges.

2.1 Methodology

CSF comprises user input processing, problem decomposition, schema selection, iterative self-evaluation, and dynamic schema switching. Inputs are parsed to identify subproblems, which are then broken down into manageable components [6]. The most suitable cognitive schema is assigned to each subproblem based on its characteristics. Agents continuously assess and refine their outputs [7], adopting alternative schemas when more efficient approaches are identified. Parallel computing is employed to process subproblems simultaneously, optimizing inference time and scalability.

2.2 Expected Contributions

This research introduces cognitive schemas as a foundational framework for prompt engineering, bridging cognitive science and AI. It develops an adaptive multi-agent architecture

that autonomously applies different schemas to meet specific problem requirements [5]. The efficacy of the architecture will be validated through performance benchmarking on diverse datasets.

2 Conclusion and Future Work

CSF enhances LLM problem-solving by integrating dynamic schema selection and a multi-agent approach, improving responsiveness and accuracy in complex tasks. Future work will refine schema selection algorithms, expand the cognitive schema repertoire, and conduct empirical evaluations to quantify performance gains, reinforcing CSF's potential as a transformative AI framework.

References

[1] Zhou, Y., Schramowski, P., & Kersting, K. (2022). "Teaching Algorithmic Reasoning via In-context Learning.", arXiv preprint arXiv:2211.01910. Retrieved from https://arxiv.org/pdf/2211.01910

[2] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., ... & Zhou, D. (2022). "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models.", arXiv preprint arXiv:2201.11903. Retrieved from https://arxiv.org/abs/2201.11903

[3] Weston, J., & Sukhbaatar, S. (2023). "System 2 Attention (S2A): Leveraging Cognitive Attention in LLMs.", arXiv preprint arXiv:2311.11829. Retrieved from https://doi.org/10.48550/arXiv.2311.11829

[4] Shapiro, D., et al. (2023). "Conceptual Framework for Autonomous Cognitive Entities (ACE).", arXiv preprint arXiv:2310.06775. Retrieved from https://doi.org/10.48550/arXiv.2310.06775

[5] Talebirad, Y., & Nadiri, A. (2023). "Multi-Agent Collaboration: Harnessing the Power of Intelligent LLM Agents.", arXiv preprint arXiv:2306.03314. Retrieved from https://doi.org/10.48550/arXiv.2306.03314

[6] Wang, X., et al. (2022). "Self-Consistency Improves Chain of Thought Reasoning in Language Models.", arXiv preprint arXiv:2203.11171. Retrieved from https://doi.org/10.48550/arXiv.2203.11171v

[7] Liu, H., Teng, Z., Cui, L., Zhang, C., Zhou, Q., & Zhang, Y. (2023). "LogiCoT: Logical Chain-of-Thought Instruction-Tuning.", arXiv preprint arXiv:2305.12147. Retrieved from https://doi.org/10.48550/arXiv.2305.12147

Paperless Office Provided by Legaltech Tools

Miklós Zorkóczy LL.M.
Pázmány Péter Catholic University, Hungary & Legalaise Zorkoczy Law Office
E-mail: miklos@legalise.hu

Abstract: *Legal terminology empowers words and expressions with specific meanings. Legal language used by lawyers is just like another language, with special wording, jargon and there are special legal instruments behind of the terminology. Using technology in the legal domain is not coding the law, tech engineers and lawyers should cooperate to let LegalTech tools work in full capacity with their effectiveness. In this way lawyers can be backed by technology, and in return for this their clients will receive enhanced customer experience. Lawyers can't be substituted, but the legal tasks can be speeded up, and costs can be decreased. Large Language Models and Blockchain technology will be the next generation of LegalTech tools, just lawyers need to be careful until the tech achievements will be accurate and accountable enough for the quality legal work. The study gives a detailed presentation of how current LegalTech tools can support legal tasks, and what are the risks of next generation of technology.*

Keywords: *LegalTech, BigData, large language models (LLM), AI, machine learning, smart contract (blockchain)*

Find the Study in English:
https://ojs.mtak.hu/index.php/pazmany_law_review/article/view/16113

SmartBook: AI-Powered Personalized Learning Experience

Ildikó Horváth[1,2], Péter Baranyi[1,2], Ádám Csapó[1,2], Anna Sudár[1,2], Borbála Berki[1,2], Péter Földesi[2]
[1]Corvinus University of Budapest
[2]Hungarian Research Network
Email: ildiko.horvath@uni-corvinus.hu

Abstract: *The SmartBook project integrates Artificial Intelligence (AI) and Virtual Reality (VR) to develop personalized digital learning environments. The framework applies adaptive content layout methods based on cognitive profiling to align with individual learning styles and preferences. It adjusts content and layout in real-time to meet each user's needs. The primary goal is to revolutionize personalized education by providing an intelligent, automated approach to optimizing learning experiences. Initial user experiments demonstrate the framework's effectiveness, showing measurable improvements in user engagement, memorization, and time efficiency. This project paves the way for the next generation of adaptive educational technologies.*

Keywords: *Virtual Reality; Artificial Intelligence; Personalized Learning; Cognitive Profiling; Adaptive Content Layouts*

1 Introduction

In traditional educational systems, the lack of adaptation to individual learners' preferences and cognitive styles reduces the effectiveness of teaching materials. Modern technologies such as Artificial Intelligence (AI) and Virtual Reality (VR) provide tools to address these challenges. The SmartBook project applies these technologies to create personalized, interactive, and adaptive digital learning and training environments. By focusing on cognitive profiling and real-time data analysis, the SmartBook framework offers a new approach to education and training, adjusting content to align with individual users' needs and improving their learning experience.

2 Methodology

The goal of the SmartBook project is to enhance human performance through three interconnected pillars: evaluating ICT tools, personalizing content, and implementing adaptive IT solutions. The first pillar addresses the evaluation of ICT tools by developing a methodology to measure operational complexity and assess digital interfaces. This is followed by validation through empirical studies, which informs cognitive load optimization and supports 3D VR planning. The second pillar focuses on whether personalized content can improve performance. AI-driven user profiling is used to create individualized digital content, validated through experiments that demonstrate better memorization and time efficiency. The third pillar involves the implementation of adaptive solutions. A template language is developed to define personalized schematic layouts, complemented by dynamic content arrangement. Validation and refinement result in automated, personalized VR environments that optimize engagement and learning outcomes.

3 Results

Experimental validation of the SmartBook framework revealed that users consistently preferred spatial layouts matching their individual learning styles. The chart highlights that users favored learning-style-matched layouts with 2D content over blank monitors, particularly in assimilating, converging, and diverging learning styles. This confirms a strong correlation between layout preferences and cognitive styles. Furthermore, the presence of specific 2D content influenced users' choices, demonstrating the importance of aligning layout design with information-gathering preferences.

4 Conclusion

The SmartBook project confirmed the effectiveness of personalized spatial layouts in addressing users' learning style preferences. By incorporating specific 2D content, the system improves engagement and aligns with individual cognitive needs. Future work will build on these findings to further develop adaptability and scalability across diverse user groups.

References

[1] Horváth, I.; Csapó, Á.B. Motivations and Tools Relevant to Personalized Workspaces in VR Environments. Electronics 2023, 12, 2059. https://doi.org/10.3390/electronics12092059

[2] Horváth, I. Personalized learning opportunity in 3D VR. 2020 11th IEEE International Conference on Cognitive Infocommunications CogInfoCom), Mariehamn, Finland, 2020, pp. 000425 000430, doi: 10.1109/CogInfoCom50765.2020.9237895.

AI Solutions for Addressing Cognitive Bias in Tabletop RPG Experiences

Márton Vránics
Hungarian Research Network
Email: vranics.marton@gmail.com

Abstract: *This project introduces a novel AI-driven system designed to enhance gameplay in Tabletop Role-Playing Games (TTRPGs) by identifying and addressing cognitive biases. The framework uses advanced algorithms to analyze player behaviors, decisions, and dialogue, detecting biases such as confirmation bias, anchoring, and groupthink. It then generates real-time, adaptive scenarios and character interactions tailored to mitigate these biases. This dual-purpose tool not only enriches gameplay with immersive, bias-aware storytelling but also serves as a training resource for corporate HR departments, fostering creativity and balanced decision-making. Experiments demonstrate measurable improvements in player engagement, problem-solving, and narrative satisfaction.*

Keywords: *Tabletop Role-Playing Games; Cognitive Bias Mitigation; AI Tools; Corporate Training*

1 Introduction

Cognitive biases often influence decision-making in TTRPGs, leading to unbalanced gameplay and limited creative outcomes. This project addresses these challenges using Artificial Intelligence to detect and counteract biases in real-time. The dual-purpose application targets both individual players seeking enhanced gameplay and corporate entities requiring gamified tools for team-building and bias awareness training.

2 Methodology

The system's development is based on three pillars:

- Bias Detection: AI algorithms analyze gameplay data to identify patterns of cognitive biases, such as confirmation bias and groupthink.
- Adaptive Interventions: Once biases are identified, the AI generates personalized scenarios and character actions to challenge biased thinking.
- Real-Time Adjustments: Dynamic narrative elements are tailored during gameplay to maintain engagement and improve decision-making.

Validation is conducted through TTRPG sessions, testing the effectiveness of bias mitigation and assessing user satisfaction. VR environments that optimize engagement and learning outcomes.

3 Results

Experimental data show that AI interventions lead to a significant reduction in cognitive biases during gameplay. Players exposed to bias-aware narratives demonstrated higher levels of creativity, problem-solving, and narrative immersion. Corporate trials can be promising because of the tool's potential for fostering team dynamics and reducing workplace biases through gamified scenarios.

4 Conclusion

This project establishes the potential of AI in creating bias-aware environments within TTRPGs. By seamlessly integrating bias detection and mitigation, the tool enhances gameplay while providing educational value. Future research will focus on expanding adaptability for diverse gaming systems and increasing scalability for broader corporate applications.

References

[1] I. Horváth, "Behaviors and Capabilities of Generation CE Students in 3D VR," 2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Naples, Italy, 2019, pp. 491-494, doi: 10.1109/CogInfoCom47531.2019.9089988.

[2] Horváth, I., Berki, B., Sudár, A., Csapó, Á., Baranyi, P. (2024). Cognitive Aspects of Virtual Reality. In: Cognitive Aspects of Virtual Reality. Studies in Big Data, vol 156. Springer, Cham. https://doi.org/10.1007/978-3-031-68130-1_4.

The Human Face of AI Transformation

Enikő Zsófia Kovács, István Kerek, Anita Rozgonyi
VOSZ Budapest & Pest County,
Üzlet és Pszichológia, Budapest, Hungary
E-mail: kovacs.zsofia@uzlet-pszichologia.hu

Abstract: *This study examines the human aspects of AI transformation in Hungarian businesses, focusing on how business psychology can enhance AI adoption. In partnership with VOSZ Budapest and Pest County, a survey of 31 items revealed both the potential benefits of AI and significant challenges, such as skill gaps, emotional resistance, and concerns about employee displacement. To address these, the study developed specialized services, including tailored workshops, change management toolkits, and employee training programs. By combining psychological insights with practical strategies, this research provides a roadmap for businesses to effectively integrate AI while prioritizing workforce development and adaptability.*

Keywords: *Artificial Intelligence; Business Psychology; Change Management; Learning and Development*

1 Introduction

In today's fast-evolving business environment, shaped by AI and automation, organizations face both challenges and opportunities. By applying business psychology, we help companies maximize their AI investments through targeted upskilling and reskilling of employees. This approach equips workforces with the skills and psychological tools needed to embrace technological innovations, accelerating ROI and maintaining a competitive edge. Combining psychological insights with practical business strategies, we aim to build a culture of adaptability, continuous learning, and resilience, enabling businesses to thrive in an AI-driven market.

2 Methodolog

As a first step, we launched an online survey in collaboration with VOSZ (National Association of Entrepreneurs and Employers Budapest and Pest County), the largest private business organization in Hungary, representing 46% of the country's GDP. The survey was designed to assess the attitudes, practices, investments, and training needs of Hungarian businesses regarding AI transformation. It targeted 4,000 companies from various industries, with a specific focus on engaging business owners and CEOs, our primary audience. The survey included 31 carefully designed questions aimed at identifying key challenges and priorities in AI adoption. Data collection was completed in September 2024, providing critical insights into the state of AI readiness across the Hungarian business landscape.

3 Results

The results revealed that while 79% saw AI's positive impact on performance, 75% rated their AI knowledge as moderate or lower, and 77% used AI tools infrequently. Key challenges included a lack of expertise (80%), insufficient financial resources (50%), and data protection concerns (70%). When it comes to training, 88% of respondents recognized its necessity, with the greatest focus on improving hard (81%) and IT (65%) skills. These findings highlight the need for tailored strategies to support businesses in overcoming AI transformation challenges.

4 Conclusion

The survey findings clearly show that businesses need substantial support in AI transformation, particularly in training and change management. In response, we have developed three specialized services—AI Ignition, AI Gear Change, and AI Full Speed—to help organizations effectively integrate AI and maximize returns on their investments. These services are designed to equip teams with the necessary skills and tools to navigate AI adoption smoothly. Looking ahead, we plan to conduct more in-depth research in 2025, targeting larger Hungarian companies to refine strategies for successful AI transformation and workforce development

References

[1] Huang, M.-H., & Rust, R. T. (2023). Artificial intelligence in services. ResearchGate. https://www.researchgate.net/publication/368624051_Artificial_Intelligence_in_Services

[2] Huang, M.-H., & Rust, R. T. (2017). Artificial intelligence in service. ResearchGate. https://www.researchgate.net/publication/322958062_Artificial_Intelligence_in_Service

Learning-Driven Segmentation of NF- kB p65 Translocation in Fluorescence Microscopy Using Clustering Algorithms

Attila Tiba[1] , Ferenc Fenyvesi[2], Ildiko Bacskay[2], Katalin Reti- Nagy[2], Agnes Rusznyak[2], Tamas Giraszi[1], Balazs Harangi[1*]
[1]University of Debrecen, Faculty of Informatics
[2]University of Debrecen, Faculty of Pharmacy
E-mail: {tiba.attila, giraszi.tamas, harangi.balazs}@inf.unideb.hu, {fenyvesi.ferenc, bacskay.ildiko, retinagykatalin,rusznyak.agnes}@pharm.unideb.hu

Abstract: *Fluorescence microscopy is widely used in pharmacological studies to investigate protein dynamics, such as the translocation of the p65 subunit of NF- κB, which is essential in inflammatory responses. Traditional methods of analyzing p65 translocation, which involve manual cell-by-cell fluorescence intensity comparison between the cytoplasm and nucleus, are labor-intensive and prone to error. To streamline this process, we developed a machine learning-based segmentation tool that automates fluorescence image analysis. Using K-Means and DBSCAN clustering algorithms, the tool accurately distinguishes between nucleus, cytoplasm, and intercellular regions, allowing for rapid and precise calculation of nucleus-to-cytoplasm intensity ratios. The automated system reduces manual workload by up to 80% and improves reproducibility and accuracy. User-friendly features, such as custom brush tools and quick-fill capabilities, enhance the flexibility of the workflow. This tool represents a significant advancement in fluorescence microscopy-based pharmacological research, enabling more efficient analysis of inflammatory processes.*

Keywords: *Fluorescence microscopy, image segmentation, automated analysis*

1 Introduction and proposed method

Fluorescence microscopy and image analysis are widely used in pharmacological studies. In our experiments the p65 subunit of the NF-kB protein was investigated in cultured cells by antibody labelling. The translocation of p65 from the cytoplasm into the cell nucleus is informative in inflammatory processes.

The fluorescence intensity from the region of cytoplasm and cell nucleus should be distinguished and their ratio is determined for the evaluation of the experiments. The evaluation is conventionally carried out manually cell by cell, with the software of the microscope, however it is a time-consuming process.

To address this challenge, we implemented a machine learning-based segmentation approach that significantly reduces the manual workload. By employing clustering algorithms like K-Means and DBSCAN, we were able to automate the segmentation of fluorescence microscopy images. This method allows for the clear distinction between the nucleus, cytoplasm, and intercellular regions through color-based clustering. The automated system not only improves accuracy but also cuts down the time required for analysis.

Our tool incorporates various user-friendly features such as custom brush tools and quick-fill capabilities to further streamline the workflow. The program will open several windows, each presenting the cells, the nucleus, the segmented image, and the original image separately. Segmentation results are displayed in an intuitive table format, where each cell is assigned a unique identifier, along with the calculated nucleus-to-cytoplasm ratios. The tool also overlays these ratios on the original images, providing a clear visual reference for users (as shown in Fig. 1). By reducing manual input and offering enhanced precision, our method offers a robust solution for fluorescence microscopy-based analyses of NF-kB p65 translocation.

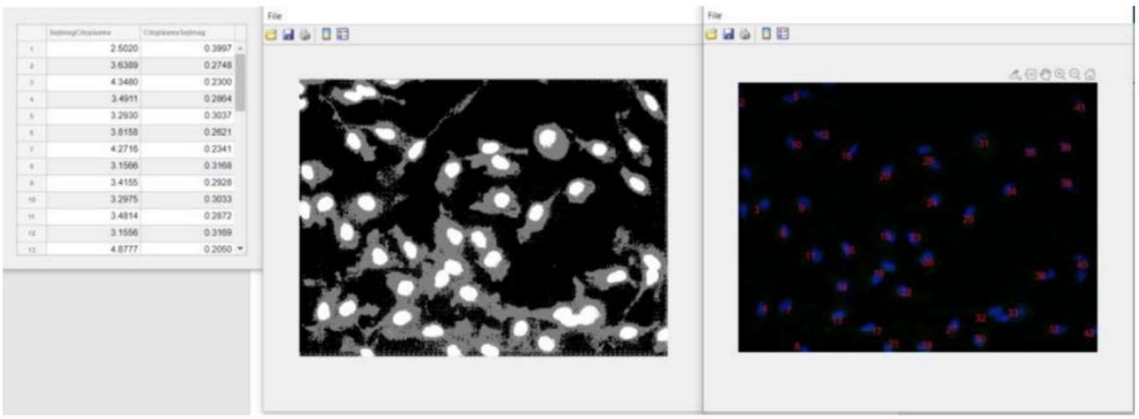


Figure 1
The output windows show the results of automated analysis.

Acknowledgement

This work was supported by Research was supported by the project GINOP-2.3.4- 15-2020-00008 supported by the European Union, co-finances by the European Social Fund. Special thanks to Microsoft Hungary Ltd. for supporting and funding this research work.

Explaining Graph Neural Networks through Subgraph Selection Policy

Dániel Unyi[1], Bálint Gyires-Tóth[1]

[1] Department of Telecommunications and Artificial Intelligence,
Faculty of Eletrical Engineering and Informatics,
Budapest University of Technology and Economics
Műegyetem rkp. 3., H-1111 Budapest, Hungary
unyi.daniel, toth.b@tmit.bme.hu

Abstract: Graph Neural Networks (GNNs) have shown strong performance in various graph-based tasks, but understanding the rationale behind their decisions remains challenging. In this paper, we propose an explainability algorithm that uses reinforcement learning to identify key subgraphs influencing GNN predictions. A mask generator neural network is trained using a custom reward function to select the optimal explanatory subgraph. To evaluate our approach, we created a dataset of random graphs, each connected to one of six possible motifs. The results demonstrate the mask generator's ability to correctly identify these motifs.

Keywords: graph neural networks; reinforcement learning; explainability

1 Introduction

Graph Neural Networks (GNNs) are a powerful tool in deep learning, well-suited for tasks involving graph-structured data. Despite their success in applications like social networks, recommender systems, and the natural sciences, GNNs operate as black boxes - in a way that obscures the processes leading to their predictions. GNN explainability is still in its early stages due to the challenges posed by graph data. Recent methods have started addressing these challenges, but a significant gap remains in uncovering how GNNs arrive at their decisions.

For further details on GNN explainability, see [1, 2].

2 Methodology

We consider a graph dataset $\{G_i\}_{i=1}^N$ with labels $\{y_i\}_{i=1}^N$. Given a GNN trained on this dataset, parameterized by ϕ, which produces a probability distribution $p_\phi(f_i|G_i)$, our objective is to identify a subgraph g_i from each graph that explains the GNN's prediction for G_i. More specifically, the GNN should predict the same label for the subgraph g_i (indicating that the subgraph is sufficient) and should predict randomly for its complementary $\bar{g}_i$ (indicating that the subgraph is necessary).

To achieve this, we introduce a selector GNN, parameterized by θ, designed to generate node selection probabilities $p_\theta(s_i|G_i)$. From these probabilities, we sample the node mask $m_i \sim \text{Bernoulli}(z_i)$, where a mask value of one means that the corresponding node is included in the explanatory subgraph, and a value of zero means that the node and its connections are excluded. This way, we get the explanatory subgraph $g_i = m_i \circ G_i$ and its complementary $\bar{g}_i = \bar{m}_i \circ G_i$, which are then evaluated using the original GNN to obtain $p_\phi(f_i|g_i)$ and $p_\phi(f_i|\bar{g}_i)$.

Since the sampling operation is non-differentiable, we can use the REINFORCE algorithm [3] to perform unbiased gradient estimation for the selector. We propose a custom reward function tailored to this task: minimizing the cross-entropy between the explanatory subgraph and the whole graph, maximizing the cross-entropy between the complementary subgraph and the whole graph, and a budget B for the mask size. This formulation ensures that the selected subgraph g_i is both necessary and sufficient to explain the GNN's prediction:

$$\min \text{CE}(p(f_i|G_i), p(f_i|g_i)) + \max \text{CE}(p(f_i|G_i), p(f_i|\bar{g}_i)) + \text{ReLU}(||p(s_i|G_i)|| - B)$$

3 Results

To evaluate the proposed method, we generated a synthetic dataset by first creating 3000 random Barabási-Albert (BA) graphs as base graphs. Then, one of six possible motifs (house, house-x, diamond, 5-cycle, 6-wheel, or 5-star) was randomly attached to each base graph. The attachment involved connecting a random node from the motif to the base graph and connecting other motif nodes to random base nodes with a 0.1 probability.

Our algorithm identifies the motif nodes with up to 98.95% accuracy when selecting 20% of the nodes. In comparison, PGExplainer [4], a similar parameterized explainer method, achieves only 61.85% accuracy under the same selection conditions. The main difference between our method and PGExplainer is that PGExplainer uses the Gumbel-Sigmoid gradient estimator [5], which results in biased gradients and introduces more hyperparameters.

References

[1] H. Yuan, H. Yu, S. Gui, and S. Ji, "Explainability in graph neural networks: A taxonomic survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 5, pp. 5782–57996, 2022.

[2] C. Agarwal, O. Queen, H. Lakkaraju, and M. Zitnik, "Evaluating explainability for graph neural networks," *Scientific Data*, vol. 10, no. 1, p. 144, 2023.

[3] R. J. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning," *Machine Learning*, vol. 8, pp. 229–256, 1992.

[4] D. Luo, W. Cheng, D. Xu, W. Yu, B. Zong, H. Chen, and X. Zhang, "Parameterized explainer for graph neural network," *Advances in Neural Information Processing Systems*, vol. 33, pp. 19620–19631, 2020.

[5] E. Jang, S. S. Gu, and B. Poole, "Categorical reparameterization with gumbel-softmax," *In International Conference on Learning Representations*, 2017.

Modeling neural networks with hypergraphs via tensor-based views

Csaba Hajdu[1], Ádám Csapó[2]

[1] Department of Informatics
Széchenyi István Egyetem
Győr, Hungary
hajdu.csaba@ga.sze.hu

[2] Institute of Data Analytics and Information Systems
Corvinus University of Budapest
Budapest, Hungary
Email: adambalazs.csapo@uni-corvinus.hu

Abstract: Neural networks have seen widespread adoption due to their robust performance across various applications. However, these networks' formal description and structural analysis remain an area of active research. This paper introduces a novel framework that leverages the algebraic structure of neural networks, transforming their architecture into tensor-based representations. By integrating algebraic properties into the learning process as part of the cost function, the framework aims to generate well-structured, deterministic networks. This methodology offers a descriptive and transformative tool for enhancing neural network architecture and improving learning outcomes as part of an iterative process. Future work will focus on exhaustive validation using prominent architectures.

Keywords: neural networks; hypergraphs; transformation; tensor algebra; tensor metric

1 Introduction

Neural networks have gained prominence in recent applications due to their versatility and robust performance in pattern-matching tasks (e.g., real-time object detection, playing real-time strategy games, ChatGPT). However, the architectures are becoming increasingly complex in size and parameters, making learning and development

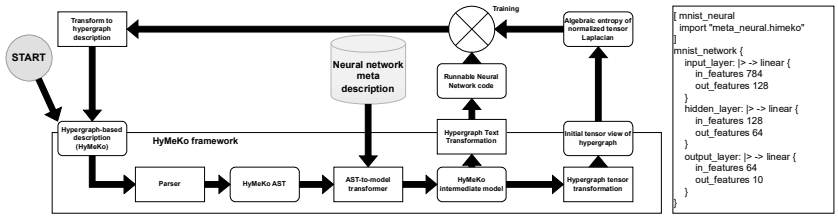


Figure 1
Overview of the process of using algebraic entropy in the learning process, with an example of neural network description in HyMeKo

challenging. Furthermore, formal verification remains an active research area due to model complexity and the difficulty to ensure reliability and correctness [1].

This paper introduces a novel hypergraph-based framework to describe neural networks in a modular and extendable way, using an intermediate language to represent neural networks. Hypergraphs [2] have recently become a prominent tool for modeling neural networks due to their ability to capture complex, high-order relationships between elements [3]. The framework goes beyond traditional modeling by providing a direct tensor-based transformation of neural network architecture, grounded in the hypergraph representation, enabling a more efficient and structured approach to neural network design. The main purpose is to avoid repetitive coding and reliance on verbose ONNX (XML-based) formats to model neural networks.

2 Methodology

The framework describes the neural network architecture, which is based on the authors' framework[1]. This description is parsed into an abstract syntax tree (AST), then transformed into an intermediate model and refined with the specific meta elements of the neural network domain. Then, a text is generated (runnable PyTorch or Tensorflow code). Alongside with the code a tensor-view representation is generated (based on previous work [4]), which is then used for the algebraic metric analysis of networks, which is added to the cost function in the backpropagation process. The recalculation of the metric is iterative during the training, and after a new network is created, it can be reused for tensor-based analysis as well. The process is depicted in Figure 1.

Two approaches are used for the modeling of neural networks with hypergraphs. *Factor view* results with the layers and the result nodes (activation results) are depicted on the graph, focusing on connection patterns. The dimensionality of data is indefinite, enabling further adjustment. *Dataflow view* depicts dimensionality

[1] https://github.com/kyberszittya/himeko_lang

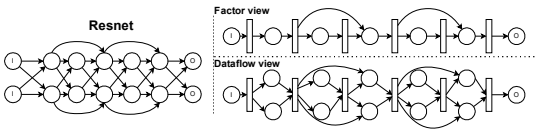


Figure 2
Example of views (factor & dataflow view) of a neural network architecture (ResNet)

and activation, as the nodes are connected to an activation layer in a linear hypergraph (example shown in Figure 2). Utilizing the described tensor view of any of the hypergraph view H, the algebraic entropy can be calculated using the hypergraph's aggregated normalized laplacian $\hat{L}(H) = \frac{L(H)}{\sum_{x \in V} D(x)} = \frac{D(H) - A(H)}{\sum_{x \in V} D(x)}$. Using the $(\hat{\lambda}_i)_{i \in \{1,2,\ldots,n\}}$ eigenvalues of $\hat{L}(H)$, the algebraic entropy is $I(H) = -\sum_{i=1}^{n} \hat{\lambda}_i \log_2 \hat{\lambda}_i$. With cross-entropy of H_1, H_2 hypergraphs as $I_H(H_1, H_2) = -\sum_i \hat{L}(H_1) \circ \log \hat{L}(H_2)$ providing a similarity metric ($\circ$ denotes element-wise matrix multiplication). Kullback-Leibler divergence can be utilized in each subsequent step to calculate the overall progress and error minimization on consecutive runs, as $\mathcal{D}_{KL}(H_1, H_2) = -\sum_i H_1 \circ \log\left(\hat{L}(H_1) \circ \frac{1}{\hat{L}(H_2)}\right)$, by adding to the $J(\theta)$ cost function $J(\theta)_{t+1} = J_t(\theta) + otherparts + \mathcal{D}_{KL}(H_t(\theta), H_{t+1}(\theta))$.

Conclusions

This paper introduces a novel method for modeling neural networks using a hypergraph-based framework. The framework, focusing on dataflow and tensor-based representations of various architectures, generates both tensor-based views and neural network programs (e.g., PyTorch, TensorFlow). By incorporating algebraic entropy into the learning process, it aims to improve the structure and determinism of networks. The future work includes validating the approach on standard models, exploring new similarity kernels for tensor representations, and investigating the integration of spiking neural networks (SNNs).

References

[1] M. König, A. W. Bosman, H. H. Hoos, and J. N. v. Rijn, "Critically Assessing the State of the Art in Neural Network Verification," *Journal of Machine Learning Research*, vol. 25, no. 12, pp. 1–53, 2024.

[2] A. Bretto, *Hypergraph Theory: An Introduction*. Springer Publishing Company, Incorporated, 2013.

[3] F. Wang, K. Pena-Pena, W. Qian, and G. Arce, "T-HyperGNNs: Hypergraph Neural Networks Via Tensor Representations," *TechRxiv*, 2023.

[4] C. Hajdu and A. Csapó, "Tensor-based Format for Exchanging Hypergraphs between Cognitive Entities," in *2022 IEEE 1st International Conference on Internet of Digital Reality (IoD)*, pp. 000033–000038, June 2022.