Can LLMs Effectively Summarize Explainable AI Output on Long-Horizon Tasks?

Kevin Fu
Georgia Institute of Technology
kfu38@gatech.edu

Abstract

Explainable AI work tends to focus on producing explanations on single-shot tasks for domain experts. However, this is insufficient to explain the behavior of robotic control policies, as robots operate over continuous time windows, producing a large amount of data, and are often operated by or coexist with non-experts. In this work, I propose using an LLM to tackle both issues, based on their ability to summarize and present information in simple language. I prompt an LLM to faithfully interpret the output data from an existing explainable AI technique, GraphMask-Explainer, which is run on the learned control policy of a robotic navigation task. Results from a small user study suggest that our method is more interpretable than baseline approaches while still maintaining faithfulness to the original data. However, these effects do not hold in all scenarios, and I leave improving the robustness of this method to future work.

1 Introduction

The field of explainable AI (xAI) aims to improve the trustworthiness and reliability of AI systems by providing insight into how they work. Modern AI relies on deep neural networks, which are notoriously difficult to analyze. Much of the xAI literature has focused on making AI explainable to domain experts, which is logical as this can help AI researchers improve their own systems. However, a recent study (Feldhus et al., 2022) has wondered if large language models (LLMs) can help make AI explainable for non-expert users. They study the effectiveness of running standard xAI methods, which often give numerical outputs, then using an LLM to interpret said outputs in natural language. For instance, one of their experiments uses Integrated Gradients (Sundararajan et al., 2017) to extract a saliency map for a sentence classification task, then appends the saliency scores to each word of the sentence before asking an LLM to explain the model's classification. This was more effective for non-experts than the standard saliency heatmap.

A limitation of this method is that interleaving the input data and xAI system output before feeding it to an LLM makes the size of the input data a bottleneck. For explaining the behavior of an AI system over longer-horizon tasks, this simple strategy may exceed the context window limits of an LLM. This is a practical problem for applying xAI techniques to robotics control polices, which can operate over long periods of time. More broadly, analyzing a large volume of data with standard xAI methods can quickly become unwieldy.

To address this limitation, I propose reformatting the xAI outputs as text values in a table, rather than combined together at a one-to-one ratio with the original input data, then prompting an LLM to interpret the table and provide a brief summary. This greatly condenses the information gathered from a continuous control policy. I argue that this will provide faithful, yet easily digestible verbal explanations of lengthier xAI outputs, which no other work has done yet to the best of my knowledge. Since it is well-established that LLMs may hallucinate (Huang et al., 2023), I conduct a small user study to assess the effectiveness of this technique. Results show that this technique has promise in producing more easily interpretable, faithful explanations than existing approaches.

2 Related Work

2.1 Explainable AI for GNNs

Many modern day AI applications are discriminative models trained with machine learning, meaning they are black boxes which simply output yes or no without any reasoning behind their choice. The study of explainable AI (xAI) aims to create AI models which can be probed for explanations as to why they make certain classifications (Gohel et al., 2021). This desire for explainability has naturally followed the trends of deep learning research, from computer vision (Gohel et al., 2021) to natural language processing (Danilevsky et al., 2020; Zhao et al., 2024). Similarly, xAI techniques for graph neural networks (GNNs) have recently become more common as GNNs are applied more broadly (Ying et al., 2019; Schlichtkrull et al., 2020; Kakkad et al., 2023). In this work, we utilize GraphMask-Explainer (Schlichtkrull et al., 2020) to analyze the most relevant edges of a GNN-based robot control policy.

2.2 LLMs as Interpreters for Non-Experts

This work aligns with the recent trend of using LLMs to improve the quality of AI robotic systems using natural language guidance. This strategy has been applied to automating assessment of failure modes (Liu et al., 2023), defining short-term motion plans (Tziafas et al., 2023; Xie et al., 2023; Majumdar et al., 2023), and even both at once (Brohan et al., 2023). The common thread is showing how LLMs can connect simple instructions to complex robotic control policies.

In this work, we tackle the reverse problem: how do we connect complex robotic control policies back to simple natural language? While there is some existing work wondering whether LLMs can interpret data in this manner, to our knowledge, no prior work has focused on using an LLM for faithful compression of large volumes of xAI output, instead directly using an LLM for reasoning over raw data (Lu et al., 2022) or interpreting much smaller xAI outputs (Feldhus et al., 2022).

3 Method

In this section, I first explain the robotics task and policy we analyze, then explain how we use the LLM to improve the quality of traditional xAI methods, and finally detail the user study used to evaluate the proposed approach.

3.1 Learned Multi-Robot Navigation

My collaborators and I (see Section 6) first train a policy to solve a multi-robot navigation task, where robots aim to travel to pre-assigned goal zones without colliding with one another. Notably, robots must communicate with one another to succeed at this task, as each robot can only observe its own position, velocity, and goal zone. We train a decentralized policy for each robot using MAPPO (Yu et al., 2022), a multi-agent reinforcement learning algorithm, and incorporate a GNN (Scarselli et al., 2009) to facilitate learned communication between them. The GNN represents robots as nodes and communication links between them as edges. The policy uses this GNN to decide which information is relevant, and outputs movement actions for each robot at each timestep. Thus, we can interpret the learned communication scheme by identifying the important edges in this GNN at each timestep. To achieve this, we first fully train the policy, then apply GraphMask-Explainer (Schlichtkrull et al., 2020) to the GNN. This allows us to analyze which robots are paying attention to which other robots, at specific points in time. The final output takes the form of a timestamped list of directed edges, where an edge being present means that specific robot-to-robot communication link is in use.

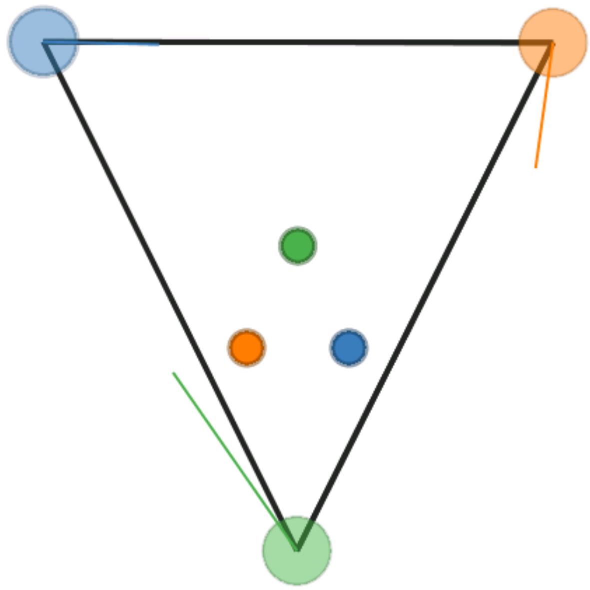

Figure 1: An example frame from the baseline video explanation. Edges here (drawn as black lines) between the robots (larger circles) represent active communication links as the robots navigate to their goals (smaller circles, matching color). Note: this figure is produced by my external collaborators (see Section 6).

A key issue with applying GraphMask-Explainer in this way is that because it produces a new explanation for each timestep, the resulting interpretations can be unwieldy to analyze for longer-horizon tasks, making this task an ideal candidate for our method. Our baseline approach will be to produce videos of the robots moving in real space, and draw edges between them on certain frames based on the output of GraphMask-Explainer. My collaborators and I concluded that this video format is a more interpretable form of explanation than the raw output itself. A single frame of this video approach is shown in Figure 1.

3.2 Interpreting xAI Data with LLMs

My two main hypotheses are that an LLM, given the same xAI model output described above to produce a video, can (1) produce an explanation which is more easily interpretable than a video (or the raw data) would be, and (2) do so while staying faithful to the xAI output.

To test these hypotheses, I build a semi-structured table from the timestamped output of GraphMask-Explainer, following the table format in Lu et al. (2022) because they showed this format improves LLM performance on mathematical reasoning tasks. I then produce concise interpretations with a two-stage approach. First, I prompt the LLM to interpret the data table, given some instructions on what each column heading means. I include specific requests to reference the original input data in the response, in an attempt to keep the interpretation faithful to the data, as well as asking the model to highlight any collisions, as this is the most relevant failure case of a navigation task from a robotics standpoint. Second, inspired by recent findings which indicate LLMs are effective zero-shot summarizers (Zhang et al., 2024), I ask the LLMs to summarize their own findings with a follow-up prompt: "Can you summarize your findings in one paragraph, while still maintaining specific references to the original data?" Without this second stage, I found that the LLM responses were overly verbose and tended to contain spurious information, such as highlighting the starting locations of the robots.

I followed this procedure for the trained policies of three separate scenarios of a navigation task, each with different start and goal zone locations. As to the specific LLM, I used the web-based version of Claude 3, a recent LLM released by Anthropic. I chose Claude 3 after reports (Enis and Hopkins, 2024) that it outperforms GPT-4 (Achiam et al., 2023) on many common benchmarks. The full initial prompt for scenario 2 can be found in Appendix A.1. The rest of the prompts, as well as scripts and raw data used to produce them can be found on GitHub.[1]

3.3 User Study Design

To validate my hypotheses, I created a short questionnaire, in which I asked participants to compare the baseline video explanation's interpretability against the original video. I then showed the video and the text explanation side-by-side, asking subjects to rate the interpretability of the combined explanation compared to the original video, and the text summary's faithfulness to the video. I present each question as an agree/disagree statement, using a five-point Likert scale (Sullivan and Artino, 2013) for all response options. To mitigate response bias (Kreitchmann et al., 2019), I phrase each statement in both the affirmative and the negative (e.g. "This explanation is helpful..." followed by "I find the explanation unhelpful...") and present both to users. Finally, users were shown this sequence of raw video, video with explanation, and video plus LLM-generated text explanation for three different scenarios. In all three scenarios, robots successfully navigate to their goal zones, but in two of the scenarios, collisions occur. The order of the scenarios was randomized before being shown to users. The questionnaire is public [2] and was shared using Qualtrics.

4 Results

In total, eight technically competent participants completed my user study, with a Bachelor's in Computer Science being the minimum achieved education level, and self-reporting a median familiarity of 4.5 compared to the average person (on a five-point Likert scale). This is beneficial for accurately establishing the faithfulness of the LLM-generated interpretation, but future work could extend this study to non-technical users to better assess interpretability.

Since I asked each non-demographic question as both an affirmative and negative statement to mitigate response bias, all the following results are rescaled such that marking "strongly agree" to the affirmative statement is a 5, while marking "strongly disagree" for the negative statement is also a 5. Then I sum responses for each affirmative/negative pair, meaning a "strongly agree" to the affirmative and a "strongly disagree" to the negative for the same statement is a 10, while the opposite is a 2. Finally, I zero-mean this range by subtracting 6 from each sum, making the strongest agreement with the affirmative a +4 and strongest disagreement a -4.

For each scenario, I asked a control question asking if a collision occurred, to ensure users under-

[1] https://github.com/kfu02/cs-7650-nlp-final-project

[2] https://gatech.co1.qualtrics.com/jfe/form/SV_3JXG5THBZwzr0X4

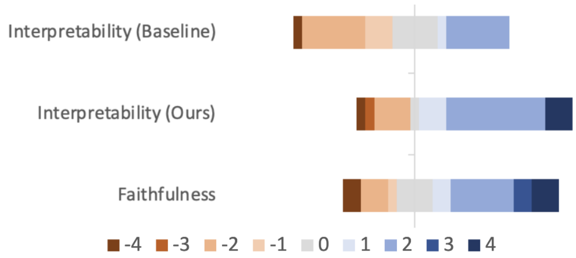

Figure 2: User ratings for interpretability of the baseline method and our method, and faithfulness ratings of textual explanations to the original input in our method. +4 represents the most positive sentiment, while -4 represents the most negative.

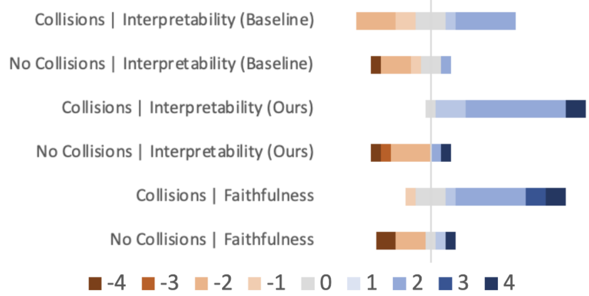

Figure 3: User ratings for interpretability of the baseline method and our method, and faithfulness ratings of textual explanations to the original input in our method. Data is subdivided into scenarios where collisions occurred and scenarios where they did not. +4 represents the most positive sentiment, while -4 represents the most negative.

stood the format of the video. Every user correctly identified that two of the scenarios had collisions (an average score of +4 for those two scenarios) and one did not (average score of -4).

Distributions of the user study results are shown in Figure 2. Analyzing the data across all three scenarios suggests that users believe our method, the combined video and text explanation, is more easily interpretable than the video-only explanation produced directly from GraphMask-Explainer's output. The data also suggest that that they largely found the text explanations faithful to the video. This would validate both hypotheses.

However, separating the data by the scenario types reveals more interesting trends. Specifically, since we asked the LLM to highlight collisions in each scenario, Figure 3 shows the response data separated into scenarios where a collision did occur and scenarios where a collision did not. We can see that our method improves interpretability greatly in the two scenarios where collisions did occur, but seemingly does not have any effect in the collision-free scenario. This can be explained by examining the faithfulness data: in the scenario without collisions, users largely rated the text interpretation as unfaithful to the video, suggesting that the LLM-provided explanation was faulty. This indicates there may be some bias in the initial prompt, which asked the LLM to specifically focus on the risk of collisions. Future work could invest more heavily into de-biasing technques or attempt to improve the explanations with a few-shot prompting strategy (Brown et al., 2020). That said, the strength of the interpretability improvement and faithfulness ratings on the two scenarios with collisions suggests that the method proposed here has potential as an improvement to long-horizon explainable AI.

5 Discussion

In this work, we showed that using LLMs to provide faithful, yet succinct interpretations of xAI data has potential, but may not yet be effective in all scenarios. I hope to extend this work further by improving the robustness of this method. I also hope to expand the user study to assessing the ease of interpretation for *non-technical* users, either by surveying them directly, and/or via an indirect approach such as measuring their average response time and accuracy on a questionnaire given our system's outputs.

6 External Collaborators

I am pursuing a thesis-option MSCS under the direction of Prof. Harish Ravichandar. I worked alone on this project instead of in the usual group of 2-4, because it may help contribute to one of the ongoing research directions in my lab. (This topic and work arrangement were approved by Prof. Wei Xu via email beforehand.) Two of the students in my lab, Siva Kailas and Shalin Jain, have been developing the xAI side of this project, specifically the use of GraphMaskExplainer on learned robot policies. My contributions are building a pipeline to use an LLM as a post-processing step to potentially improve explanation quality, determining which LLM and prompting strategy to use, and conducting a user study to assess interpretability and faithfulness. I also made suggestions during the training process to help improve the suitability of the output data for use in LLMs. Finally, I have written all parts of this report, as well as the associated presentation.

References

Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Anthony Brohan, Noah Brown, Justice Carbajal, Yevgen Chebotar, Xi Chen, Krzysztof Choromanski, Tianli Ding, Danny Driess, Avinava Dubey, Chelsea Finn, et al. 2023. Rt-2: Vision-language-action models transfer web knowledge to robotic control. *arXiv preprint arXiv:2307.15818*.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Marina Danilevsky, Kun Qian, Ranit Aharonov, Yannis Katsis, Ban Kawas, and Prithviraj Sen. 2020. A survey of the state of explainable ai for natural language processing. *arXiv preprint arXiv:2010.00711*.

Maxim Enis and Mark Hopkins. 2024. From llm to nmt: Advancing low-resource machine translation with claude. *arXiv preprint arXiv:2404.13813*.

Nils Feldhus, Leonhard Hennig, Maximilian Dustin Nasert, Christopher Ebert, Robert Schwarzenberg, and Sebastian Möller. 2022. Saliency map verbalization: Comparing feature importance representations from model-free and instruction-based methods. *arXiv preprint arXiv:2210.07222*.

Prashant Gohel, Priyanka Singh, and Manoranjan Mohanty. 2021. Explainable ai: current status and future directions. *arXiv preprint arXiv:2107.07045*.

Lei Huang, Weijiang Yu, Weitao Ma, Weihong Zhong, Zhangyin Feng, Haotian Wang, Qianglong Chen, Weihua Peng, Xiaocheng Feng, Bing Qin, et al. 2023. A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *arXiv preprint arXiv:2311.05232*.

Jaykumar Kakkad, Jaspal Jannu, Kartik Sharma, Charu Aggarwal, and Sourav Medya. 2023. A survey on explainability of graph neural networks. *arXiv preprint arXiv:2306.01958*.

Rodrigo Schames Kreitchmann, Francisco Abad, Vicente Ponsoda, Maria Nieto, and Daniel Morillo. 2019. Controlling for response biases in self-report scales: Forced-choice vs. psychometric modeling of likert items. *Frontiers in Psychology*, 10.

Zeyi Liu, Arpit Bahety, and Shuran Song. 2023. Reflect: Summarizing robot experiences for failure explanation and correction. *arXiv preprint arXiv:2306.15724*.

Pan Lu, Liang Qiu, Kai-Wei Chang, Ying Nian Wu, Song-Chun Zhu, Tanmay Rajpurohit, Peter Clark, and Ashwin Kalyan. 2022. Dynamic prompt learning via policy gradient for semi-structured mathematical reasoning. *arXiv preprint arXiv:2209.14610*.

Arjun Majumdar, Fei Xia, Dhruv Batra, Leonidas Guibas, et al. 2023. Findthis: Language-driven object disambiguation in indoor environments. In *7th Annual Conference on Robot Learning*.

Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. 2009. The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1):61–80.

Michael Sejr Schlichtkrull, Nicola De Cao, and Ivan Titov. 2020. Interpreting graph neural networks for nlp with differentiable edge masking. *arXiv preprint arXiv:2010.00577*.

Gail Sullivan and Anthony Artino. 2013. Analyzing and interpreting data from likert-type scales. *Journal of graduate medical education*, 5:541–2.

Mukund Sundararajan, Ankur Taly, and Qiqi Yan. 2017. Axiomatic attribution for deep networks. In *International conference on machine learning*, pages 3319–3328. PMLR.

Georgios Tziafas, Yucheng Xu, Arushi Goel, Mohammadreza Kasaei, Zhibin Li, and Hamidreza Kasaei. 2023. Language-guided robot grasping: Clip-based referring grasp synthesis in clutter. *arXiv preprint arXiv:2311.05779*.

Amber Xie, Youngwoon Lee, Pieter Abbeel, and Stephen James. 2023. Language-conditioned path planning. In *7th Annual Conference on Robot Learning*.

Zhitao Ying, Dylan Bourgeois, Jiaxuan You, Marinka Zitnik, and Jure Leskovec. 2019. Gnnexplainer: Generating explanations for graph neural networks. *Advances in neural information processing systems*, 32.

Chao Yu, Akash Velu, Eugene Vinitsky, Jiaxuan Gao, Yu Wang, Alexandre Bayen, and Yi Wu. 2022. The surprising effectiveness of ppo in cooperative multi-agent games. *Advances in Neural Information Processing Systems*, 35:24611–24624.

Tianyi Zhang, Faisal Ladhak, Esin Durmus, Percy Liang, Kathleen McKeown, and Tatsunori B Hashimoto. 2024. Benchmarking large language models for news summarization. *Transactions of the Association for Computational Linguistics*, 12:39–57.

Haiyan Zhao, Hanjie Chen, Fan Yang, Ninghao Liu, Huiqi Deng, Hengyi Cai, Shuaiqiang Wang, Dawei Yin, and Mengnan Du. 2024. Explainability for large language models: A survey. *ACM Transactions on Intelligent Systems and Technology*, 15(2):1–38.

A Appendix

A.1 Selected Full Prompt

Below is one of the full prompts given to the LLM for scenario 2 (which is depicted graphically in Figure 1). Note that the whitespace here is reformatted to fit the style guidelines. When presented to the LLM, there is only one space between each data value and spacing line (|). The other two full prompts can be found on GitHub. [3]

```
You will be presented with a table of data derived from observing a robot team. Each robot has attempted to learn a policy to navigate near its goal zone without colliding with another robot. This data is collected over a series of timesteps (denoted under the ”timestep” column).

A value of 3.4 under ”r0-r1 distance” means that at that timestep, robot 0 and robot 1 are 3.4 meters apart. Thus, you can assume that any value less than 0 here means these two robots have collided. A value of 1.0 under ”r0-r1 attention” means at that timestep, robot 0 is paying attention to the behavior of robot 1, while a value of 0.0 means it is not. A value of 0.5 under ”r0 dist to goal” means at that timestep, robot 0 is 0.5 m away from the center of its goal zone. Similar logic applies to all other pairings of robots.

Please interpret the policy’s learned behavior based on the input table. Support your interpretation with data from the table. If you notice that the robots have collided (meaning a value in one of the robot-robot distance columns is negative), please highlight this.

Here is the input data you must analyze.

Input table:
timestep | r0-r1 distance | r0-r2 distance | r1-r2 distance | r0-r1 attention | r0-r2 attention | r1-r0 attention | r1-r2 attention | r2-r0 attention | r2-r1 attention | r0 dist to goal | r1 dist to goal | r2 dist to goal
1 | 1.2789000833944868 | 1.4505268420471429 | 1.4497620938325315 | 1.0 | 1.0 | 1.0 | 0.0 | 1.0 | 1.0 | 1.2681253329225783 | 1.2719507537636825 | 0.895901428729746
2 | 1.2744100085652532 | 1.4376884112368524 | 1.4320414181856822 | 1.0 | 0.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.2570732240008933 | 1.268845592576181 | 0.8877121496594265
3 | 1.2659014839483806 | 1.4205222369721173 | 1.4026698401091948 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.2398421915711693 | 1.2627928412847453 | 0.8760035673443345
4 | 1.2586134214188642 | 1.3997023924165823 | 1.3620661838241788 | 1.0 | 0.0 | 1.0 | 1.0 | 0.0 | 1.0 | 1.2189337143585781 | 1.2550881921203785 | 0.8599923081051365
5 | 1.253648048212327 | 1.3754475014866518 | 1.3118047917407754 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.1959137636134138 | 1.246215342547186 | 0.8395472410769985
6 | 1.2489139968017187 | 1.3480597214392063 | 1.254029307714063 | 1.0 | 0.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.170776298017687 | 1.235951896313121 | 0.8157366364213391
7 | 1.243413557406108 | 1.3175464253153462 | 1.1908812458885405 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.1436396810184577 | 1.2243577581736476 | 0.7894019571802442
8 | 1.237142683473379 | 1.285254613014024 | 1.1239287220682501 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.1147413063128147 | 1.2120139025605277 | 0.7615770808000987
9 | 1.230592585807607 | 1.2521403261917663 | 1.0539348531224035 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0844537288423144 | 1.199529816219672 | 0.7326112475249067
10 | 1.223852416280833 | 1.2191355217629782 | 0.981886537906137 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0528288084964241 | 1.187601153586506 | 0.7029703052049923
11 | 1.2174117190283373 |
```

[3] https://github.com/kfu02/cs-7650-nlp-final-project

1.1864858406681527 | 0.9084591840204059 | 1.0 | 1.0 | 0.0 | 1.0 | 1.0 | 1.0 | 1.0199380765517092 | 1.176765333445883 | 0.6727005351566179
12 | 1.2105563428261048 | 1.1536608169413582 | 0.8349220113354352 | 1.0 | 1.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.9856789132369628 | 1.1668633381848965 | 0.6420009423669096
13 | 1.2024831141352785 | 1.1202995635304818 | 0.7615776332007571 | 0.0 | 1.0 | 1.0 | 1.0 | 0.0 | 1.0 | 0.9498684119392539 | 1.1577516184398102 | 0.6106600117905217
14 | 1.1922832754090094 | 1.086256827733352 | 0.6890529412525983 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.0 | 0.9123894837184392 | 1.148972279909311 | 0.5790208631819755
15 | 1.1800606415437869 | 1.0523656785688618 | 0.6172550029710184 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.0 | 0.8735440515509221 | 1.140478114651921 | 0.5473891211925936
16 | 1.1650263715900866 | 1.018060838569737 | 0.5464356006867113 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.8334632925330305 | 1.1317542356890034 | 0.5156121895378347
17 | 1.146076052055668 | 0.9816889780637916 | 0.47669656235695607 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.7923628587963977 | 1.1219720673884888 | 0.4827572889144192
18 | 1.1222958876492173 | 0.9419384665291015 | 0.4086186841003056 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.7505218051462597 | 1.1102248646107689 | 0.4484438091890666
19 | 1.0935755973677344 | 0.8985087617001486 | 0.3428287483773372 | 1.0 | 1.0 | 0.0 | 1.0 | 1.0 | 1.0 | 0.7083383725310948 | 1.0962786963176836 | 0.4124568947174965
20 | 1.0601829947316126 | 0.8521365771206577 | 0.2798529583787616 | 1.0 | 1.0 | 1.0 | 1.0 | 0.0 | 1.0 | 0.6660523102579856 | 1.0801140032422505 | 0.37523947020536097
21 | 1.021544622637463 | 0.8029799851414492 | 0.21948544913341556 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.6235391006183975 | 1.061088544844397 | 0.3369914687347441
22 | 0.9769628231486558 | 0.7501052777920549 | 0.16259671979775253 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.5808720513159503 | 1.0380533415966637 | 0.29789630746284856
23 | 0.9267547950630073 | 0.69492938525331 | 0.11010255982042912 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.5380261053146027 | 1.0113018589916662 | 0.2586185801523162
24 | 0.8720294181019117 | 0.6397940741828825 | 0.063504673682816 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.49529031890397374 | 0.9815672417109284 | 0.22020535869955574
25 | 0.8137096545935905 | 0.587578559646057 | 0.024321182053460108 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.4528251428531768 | 0.9495522313174775 | 0.18451582587951634
26 | 0.7527283279518491 | 0.54117563019319 | -0.0066932256115620065 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.41075877349120615 | 0.9160068176602181 | 0.1538334488984759
27 | 0.6895178612869568 | 0.5007097751522454 | -0.028436329122664847 | 0.0 | 1.0 | 0.0 | 1.0 | 1.0 | 1.0 | 0.3689384772560325 | 0.8813305055426143 | 0.1282076440778786
28 | 0.6315042806703908 | 0.45955336067155167 | -0.026179851408580354 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.3313210592796366 | 0.8461319637030621 | 0.10713192801401458
29 | 0.5754966498986653 | 0.42257152909228707 | -0.010459812923630338 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.3013792295431123 | 0.8110066090976078 | 0.09073329047268153
30 | 0.5165415534781455 | 0.3951007803604218 | 0.0051707130157028075 | 1.0 | 1.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.2691942235635824 | 0.7763939270756824 | 0.07841186134763031
31 | 0.4556909500651906 | 0.37653670465445793 | 0.02081216746679257 | 1.0 | 1.0 | 0.0 |

1.0 | 0.0 | 0.0 | 0.2363594296828455 | 0.7428293209075689 | 0.0696
32 | 0.3944027750588371 | 0.36590781698147706 | 0.03453141259970251 | 1.0 | 0.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.20385889727946632 | 0.7105943005681934 | 0.06290031796422019
33 | 0.33498227899636573 | 0.36240433549210466 | 0.043873075549590704 | 1.0 | 1.0 | 0.0 | 1.0 | 1.0 | 1.0 | 0.172860782134063 | 0.6804423267257851 | 0.05530795602804355
34 | 0.28117913962973357 | 0.3659439222314708 | 0.05338685045188253 | 1.0 | 1.0 | 0.0 | 1.0 | 1.0 | 1.0 | 0.14610355916267062 | 0.6524328164646533 | 0.04353412454615345
35 | 0.233907371524191 | 0.3706742399190945 | 0.06600354770526885 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.12354002895259188 | 0.6264131328444512 | 0.03326611849915767
36 | 0.19157727099537453 | 0.3711499386788431 | 0.07318205947840663 | 1.0 | 1.0 | 0.0 | 1.0 | 1.0 | 1.0 | 0.10373514351462575 | 0.6013717735311493 | 0.0269462427807663
37 | 0.15285918253410363 | 0.37369713659407183 | 0.08342251070083774 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.08624227501637466 | 0.5757897185605175 | 0.017920100446147058
38 | 0.11668682777916935 | 0.36801829903498756 | 0.08849042124513358 | 1.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0708057201079122 | 0.5486364278828011 | 0.013912943613772032
39 | 0.08248578478996334 | 0.36072067295731775 | 0.09552250315944184 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0577221794460327 | 0.5195516721174132 | 0.011766052864066182
40 | 0.04875327904976341 | 0.3445361276078263 | 0.09782282868673865 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.046472799756954605 | 0.48804292270250166 | 0.011167676125765824
41 | 0.014264231157895699 | 0.32625512354576597 | 0.10351248816699489 | 1.0 | 1.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.03761967570301477 | 0.45367399087891297 | 0.013284953895290717
42 | -0.020521454787739188 | 0.30019885620789283 | 0.10515353296558229 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.03075613759885984 | 0.41689357155034185 | 0.013892443989449804
43 | -0.04489860080513351 | 0.2793702934696856 | 0.10479976908858786 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.029138634147811388 | 0.3821242729793542 | 0.01603277892319357
44 | -0.03412127609648276 | 0.27523428798902844 | 0.08962574182389939 | 1.0 | 1.0 | 0.0 | 1.0 | 1.0 | 1.0 | 0.041895107112883724 | 0.3611100940156617 | 0.017495713760804388
45 | -0.01686054051465037 | 0.2754076503244575 | 0.07453111889917503 | 1.0 | 1.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.05727521278877976 | 0.3431938373572579 | 0.01854427135263071
46 | -0.018533650452488712 | 0.26746127025641736 | 0.072787226497332 | 1.0 | 0.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.06034583664180985 | 0.31901186498310685 | 0.020231164079212053
47 | -0.02918488005401669 | 0.2518702046113952 | 0.0776420669193116 | 1.0 | 1.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.055682133579811756 | 0.2892006396950048 | 0.0210261741645978
48 | -0.026531191144412117 | 0.23897717808187363 | 0.07970987638047568 | 1.0 | 1.0 | 1.0 | 0.0 | 1.0 | 1.0 | 0.05496216880728052 | 0.2576958284489681 | 0.0221
49 | -0.01775336838404458 | 0.22579187969275527 | 0.08254669061154842 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.054011295115003494 | 0.22333624873719 | 0.0226883229878279
50 | -0.01120105364250519 | 0.20900057109519093 |

0.08936045319335824 | 1.0 | 0.0 | 1.0 |
1.0 | 1.0 | 1.0 | 0.048834516481685374
| 0.18549288396054442 |
0.023469128658729535
51 | 0.00020413256794248047 |
0.18955988573101246 | 0.0947148550672497
| 1.0 | 0.0 | 1.0 | 1.0 |
1.0 | 1.0 | 0.045153626653902344
| 0.1464631352934929 |
0.023967686580060245
52 | 0.013587949175465783
| 0.16938172787124978 |
0.09922430985123921 | 1.0 | 0.0 | 1.0 |
1.0 | 0.0 | 1.0 | 0.04231181867989132
| 0.10938738501308091 |
0.024340295807569797

Interpretation: