How to improve an academic chart scientifically?

Student(s)

Group 16

- Motong Tian, 320180940301, tianmt18@lzu.edu.cn
- Jiang Zhou, 320180940681, zhoujiang18@lzu.edu.cn

1. Abstract

This report details the progress of our project, which tells us how to improve an academic chart scientifically. At the beginning we chose a visualization chart and explained why we chose it. Then our paper based on the source of the chart explained the specific content of the chart, and we told readers how to read the chart based on the content of the chart and cognitive theory. After this, we used Matplotlib to reproduce the picture. After completing the reproduction, we found that this chart needs improvement based on some basic principles and guidelines of visualization, so we improved the chart to make it conform to the theoretical knowledge of information visualization. At the end of the report, we summarized the entire project, explaining our gains and how to obtain better visualization.

2. Introduction

Influential Chart Source: Academic Paper

```
https://arxiv.org/abs/2011.12342
Quantization of Blackjack: Quantum Basic Strategy and Advantage
(Yushi Mura, Hiroki Wada)
```

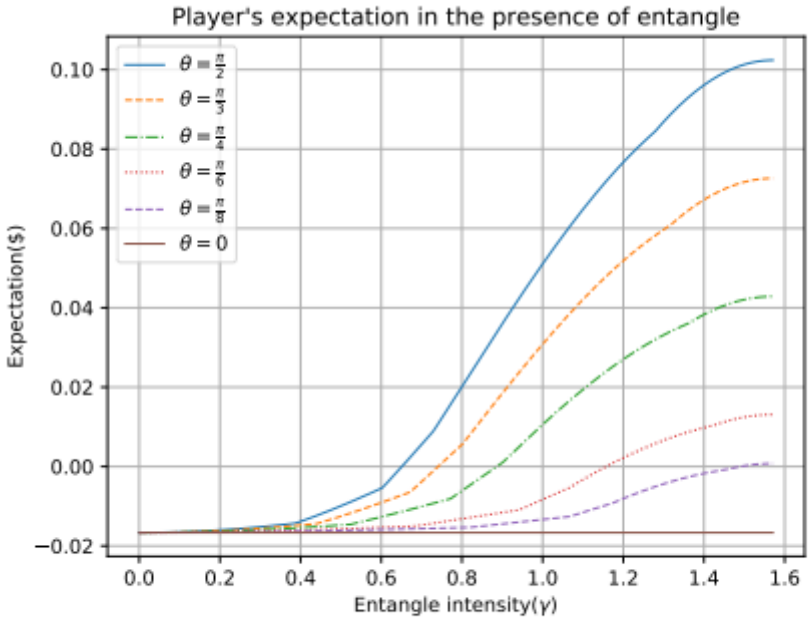


FIG. 4. Player's expectation in various parameters $(0 \leq \gamma, \theta \leq \frac{\pi}{2})$

Chart Importance:

This chart comes from an academic paper in the field of quantum physics, which is a preprint on the $arXiv$ website. Most of the articles on $arXiv$ will be submitted to academic journals, and most of the authors maintain a rigorous attitude towards the articles. Most of the papers that have been uploaded are waiting for submission and provided for peer discussion [2]. Therefore, the papers that have been uploaded and approved on this website will have a certain influence in the academic circles. These papers with academic influence need to be very accurate, whether the data or charts need to be be exact in every particular. We found this article and found an information visualization chart that could be improved.

Refer to the paper itself, the paper we found is about the application of game theory in quantum computer. In order to promote the practical application of quantum computers, it is necessary to reproduce some classical theories on quantum computers. The research direction of this paper is a hot topic, which will be widely discussed by the academic community. We can't discuss the preciseness of the experiment, and we can't know the experimental data, so we analyze the charts in the paper from the perspective of information visualization. The graph we selected violates some principles of visualization, which can be used to make this graph more perfect. And this chart also violates some guidelines to be followed in visualization, such as not paying attention to color blindness friendly. If we make some improvements, this excellent paper will be more perfect.

3. Interpretation Visualization

Background:

The two authors of this article use quantum computers to study gambling theory. A quantum blackjack is proposed and a quantum circuit is designed to replicate the classical blackjack. Furthermore, authors showed the player's expectation increases compared to the classical game, using quantum basic strategy, which is a quantum version of the popularly known basic strategy of blackjack [3].

In this paper, after completing the design of quantum circuit, we simulate Blackjack based on quantum basic strategy (QBS), and propose new parameters in the experiment. γ is the measure for the game's entanglement and they call γ entangle intensity. And θ is game parameter [3]. In the case of quantum entanglement, the player's expectation is considered. Then it shows the classical basic strategy (CBS) that players can choose in the non entanglement game and the quantum basic strategy (QBS) that they can choose in the entanglement game. The visual graph we choose is the display chart of the player's expectation with various values in entangle intensity γ and game parameter θ after the player chooses the quantum basic strategy. And the main function of this chart in the paper is to let readers understand the changes in expectations with angles, so as to draw the same conclusions as the authors.

Chart Details:

Original Chart:

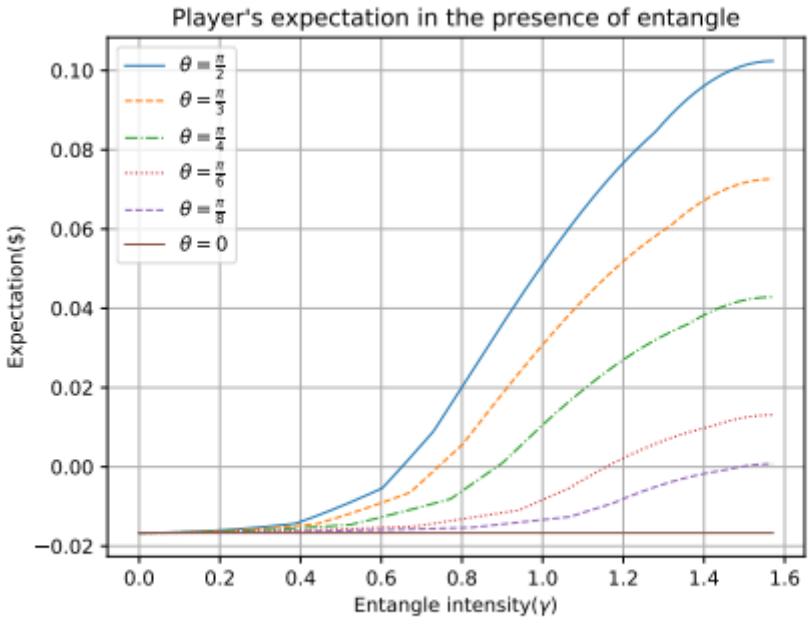


FIG. 4. Player's expectation in various parameters $(0 \leq \gamma, \theta \leq \frac{\pi}{2})$

First of all, the title of the chart is player's expectation in the presence of angle. The visual pattern of this chart is position in x-axis and y-axis. Then the label of the x-axis is **Entangle intensity**(γ), and the variables of the x-axis are consecutive floating-point numbers with a starting point of 0.0 and an interval of 0.2. The label of the y-axis is **Expectation**$(\$)$, and the variables of the y-axis are consecutive floating-point numbers with a starting point of -0.02 and an interval of 0.02. There are six lines in the graph, and

they show how expectations change with different parameters. The blue line represents the change of the expectation value with the entangle intensity when $\theta = \frac{\pi}{2}$; The orange line represents the change of the expectation value with the entangle intensity when $\theta = \frac{\pi}{3}$; The green line represents the change of the expectation value with the entangle intensity when $\theta = \frac{\pi}{4}$; The red line represents the change of the expectation value with the entangle intensity when $\theta = \frac{\pi}{6}$; The purple line represents the change of the expectation value with the entangle intensity when $\theta = \frac{\pi}{8}$; The brown line represents the change of the expectation value with the entangle intensity when $\theta = 0$.

From the perspective of chart content, it is clear from this chart, the expectation values are always the highest when $\theta = \frac{\pi}{2}$. And in $\theta = 0$ game, you can see any γ do not change the expectation value from the classical one. It shows that the upper and lower bounds in the figure are represented by solid lines, and the intermediate values are all broken lines of varying degrees. Otherwise, according to the experimental data, when $\gamma = \frac{\pi}{4}$, $\theta = \frac{\pi}{2}$, half entangled game, the expectation is $+1.8\%$ [3]. In these results, we can conclude the player has the advantage in the appropriate parameter game.

From the perspective of cognitive theory, this graph is consistent with cognitive theory. There are six labels in this picture, and different colors correspond to different angles. These contents are the working memory that you need to have when looking at the picture. As attention sources, lines with different colors can attract readers' attention. It can be seen that there is no unwanted cognitive tunneling in the figure, which is the key information related to the research. What's more, this picture also contains procedural knowledge.

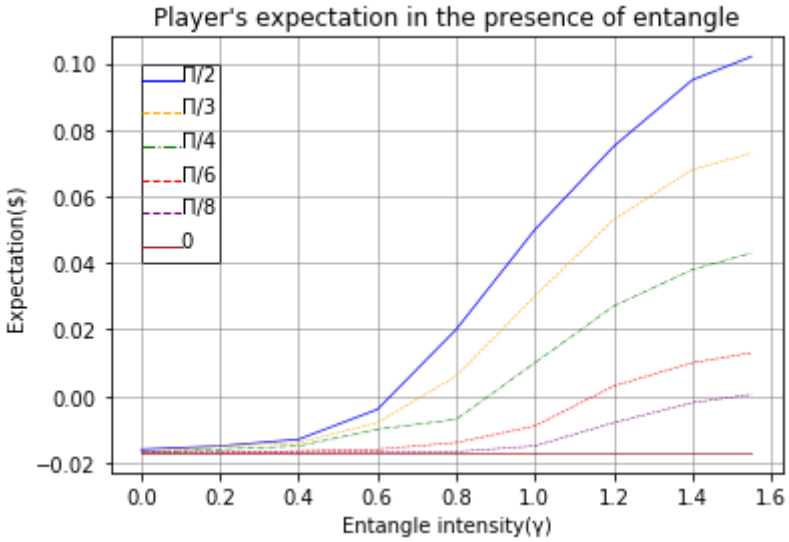
- Declarative. This chart first provides the contents of the title and label. This allows the reader to read the map with certain premise knowledge, without confusion about the content of the map.
- Consolidation: Declarative + Procedural. Over time the premise knowledge is matched and replaced by the specific data in the research.
- Procedural. Eventually, it is difficult for readers to remember the premise knowledge all the time. In order to understand the specific data, readers need to memorize the contents of previous statements repeatedly.

The above is all the details in the chart. From the previous analysis, we can find that there is something to be improved in this diagram. Next, we will use **Matplotlib** to restore this chart, and then we will discuss the modification.

4. Implementation and Improvement

The detailed execution code is in ***visulization.ipynb*** [4].

Replicate the information visualization：



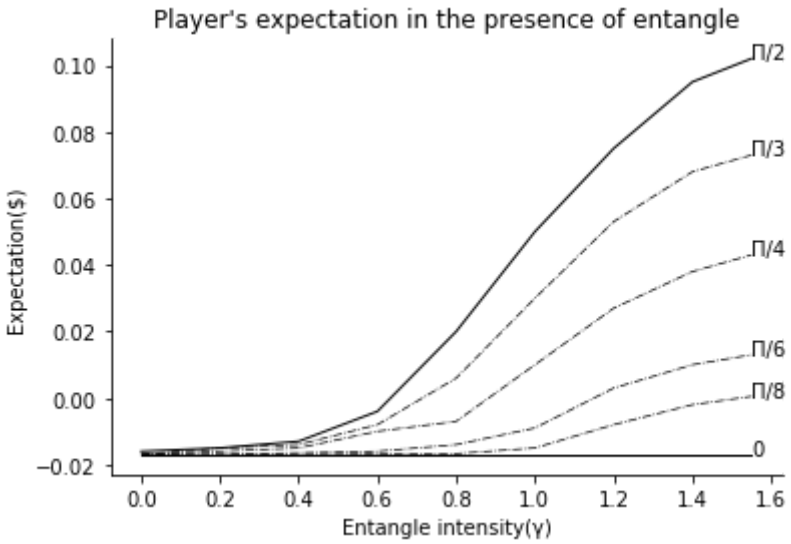
This chart uses *Jupyter Notebook* as a tool, mainly using the **Matplotlib** package in Python. By creating our own data, we have basically completely restored the charts in the paper. Next, we're going to discuss the areas in this chart that can be improved based on some theoretical knowledge of data visualization.

Improve the information visualization：

First of all, starting from the ontology of chart type, this chart belongs to a line chart, and this chart fully meets the requirements of line chart (including visual pattern and data types). Then, from the perspective of user tasks, this figure only has two tasks: overview and relate, but it meets the needs of the paper. So none of these needs to be modified. Besides, from the perspective of five Tufte's visualization principles, this chart shows all the data and clearly shows the trend and special value of the data (reasonable data-ink ratio). But this chart does not achieve *Erase Non-Data-Ink*, but this chart does not have *Redundant Data-Ink*. So our improvement plan is to remove unnecessary edges and unnecessary grid lines to make the chart look clearer and easier to read. In addition, considering the principles for graphical integrity, the lie factor of this chart is reasonable, and the label also indicates the meaning of each line in the chart (although the label is inappropriate), and the chart shows that the change in the data did not cause Ambiguous content, in short, there is no problem with the integrity of this chart.

What's more, there is a very obvious irrationality in this chart. The color-blind readers are not taken into consideration. Red and green and blue and orange are used together in the chart. This will cause great distress to color-blind people. The most important thing is that the use of color violates two very important practical guidelines for Effective visualizations: *Never use different colors to represent the same kind of data* and *Use visual variables (color, shape, shade) only for data variation*. Lines of different colors indicate that the angle value itself has nothing to do with the data variation. And not only color-blind readers, readers who are not color-blind can not easily distinguish the relationship between these colors and angles. But this chart follows *Use bold type/lines only to emphasize something*, and only uses solid lines to emphasize the most value. So we decided to abandon color and use black lines to represent data changes. Meanwhile, we also considered the cognitive theory and decided to label elements directly, avoiding indirect look-up. Finally, our improvement plan was to replace the colored lines with black lines and directly mark the angle values on the lines.
After our improvement, the final effect of this chart can not only meet the needs of the paper, but also almost completely consistent with the theory of information visualization.

Improved Chart:



Let's look at this improved chart. This chart has almost fully met the requirements of information visualization. When all of them are represented by black lines, we can clearly see that the solid line highlights the change of the maximum value and the constant minimum (zero) through the virtual and solid lines. And readers do not need too much working memory, can directly identify the angle represented by each line. The final chart also meets the requirements of the paper, allowing readers to see the changes and know what the most expected situation is.

5. Conclusion

Looking back at the entire project, we started by looking for papers on the $arXiv$ website. For each paper we clicked on, we would carefully check the visualization chart in it, and consider whether there could be improvements, until we found this article on quantum theory paper. The chart we choose needs to be relatively accurate and complete in itself, and there is room for improvement. Moreover, the most important thing is that the graph itself can be reproduced without experimental data. Fortunately, we found the required graph and the source is an excellent paper. After that, we successfully reproduced the graph and improved the graph based on the principles and guidance of information visualization. We think the final result is reasonably good.

This project has made us very rewarding. We not only read many visualizations in the literature, but also practiced the realization and improvement of visualization. It is the most important thing that we understand the importance of better visualization. We found that even rigorous academic papers would violate some principles of information visualization, which surprised us. Before we studied visualization, we did not expect that the charts in the paper would have errors. Information visualization brings us a new perspective on problems, data is provided to professionals, and visual charts are the bridge to communicate with readers. This new way of thinking is the wealth of our life.

In the end, we summarized how to have better visualization. Good visualization must meet all the basic principles of visualization, whether it is the five visualization principles, the principles of graphical integrity, or the practical guidelines. These principles are the basic requirements for achieving visualization. After completing these basic requirements, add some expanded content to the chart for user tasks, so that the

information visualization achieved is generally better. All in all, good visualization needs to be prepared in advance, and data visualization can only be done after considering all the requirements, so that it can be foolproof.

We shared our project and article in the GitHub community to spread visualization to the public.

https://github.com/tianmotong/info [5]

References

[1] https://luiscruz.github.io/course_infovis/projects/project2 “Project 2”

[2] https://www.editage.com/insights/can-you-explain-what-is-an-arxiv-publication "Q: Can you explain what is an arXiv publication?"

[3] https://arxiv.org/abs/2011.12342 "Quantization of Blackjack: Quantum Basic Strategy and Advantage"

[4] https://matplotlib.org/3.1.1/api/index.html "API Overview - Matplotlib"

[5] https://github.com/tianmotong/info "tianmotong / info"