University of Sheffield

Explainable Artificial Intelligence of Gomoku

Lu Wang

Supervisor: Emma Norling

A report submitted in fulfilment of the requirements
for the degree of MSc in Advanced Computer Science

in the

Department of Computer Science

May 9, 2019

Declaration

All sentences or passages quoted in this report from other people's work have been specifically acknowledged by clear cross-referencing to author, work and page(s). Any illustrations that are not the work of the author of this report have been used with the explicit permission of the originator and are specifically acknowledged. I understand that failure to do this amounts to plagiarism and will be considered grounds for failure in this project and the degree examination as a whole.

Name: ______________________________

Signature: ______________________________

Date: ______________________________

Abstract

Lorem ......

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Lorem ipsum ....

1.1 Aims and Objectives

Lorem ipsum ......

1.2 Overview of the Report

Lorem ipsum ......

Chapter 2

Literature Survey

Lorem ipsum dolor.....

2.1 Conventional Gomoku AI

Game tree searching based on minimax tree (See Figure 2.1), which is combined with an evaluation function of leaf board situations, is often used to develop an AI agent for Gomoku. This algorithm would generate a partial game tree which make current board situation as a root node. The depth of game tree represents how many moves of two players. When the board evaluation function has been selected an optimal leaf board situation among all leaf nodes, game tree searching algorithm will return the position of next move with tracking back the path from leaf node to root node. Due to the growing search space complexity with bigger search depth, history heuristic and alpha-beta search (Schaeffer, 1989) is used to speed up the game tree search. What is more, to reduce the search space, Allis et al. (1993) proposed a method called threat-space search, which only considers the threats of winning threat sequences and makes searching process as efficiently as possible, allowing to expand the search depth of game tree within limited time.

2.2 XAI Architecture

Explainable artificial intelligence architecture needs to be specifically designed for different applications. For complex AI systems of simulation and game, Core et al. (2006) present an architecture, as Figure 2.2 shows, to explain the behaviour of simulated entities. On the left side of the diagram, AI behaviors, such as some tasks in the game, are extracted from Simulation Environment. Since scenario and log do not include all the essential information, XAI database could be searched according scenario and logs to provide more specific information. On the right side of the Figure 2.2, the dialogue manager is used to organize the XAIs response. Firstly, reasoner is applied to collect related information from XAI database and then Natural Language Generator, which is coded by XSL templates, produce English response. The work of Core et al. (2006) is a typical XAI system for the early AI, which are

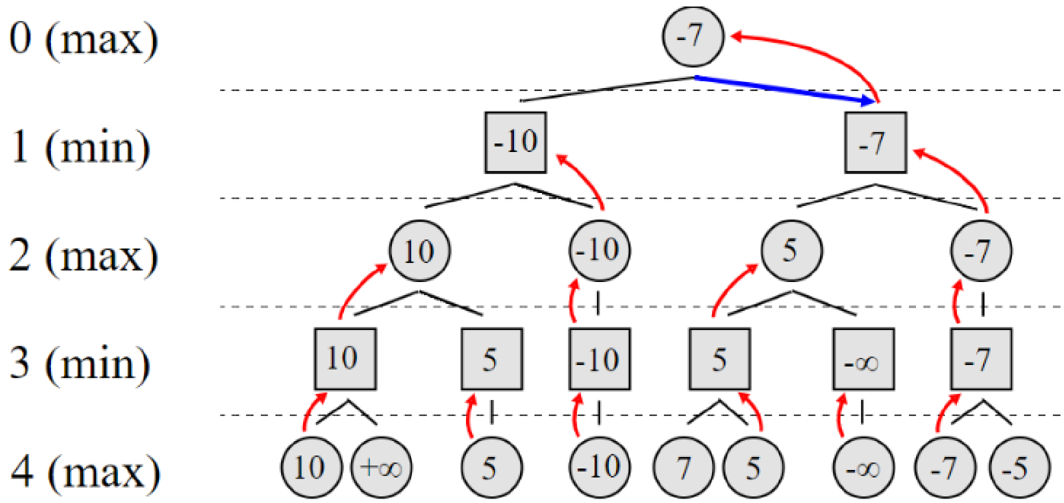

Figure 2.1: MiniMax Tree (Nuno Nogueira, 2006)

transparent systems.

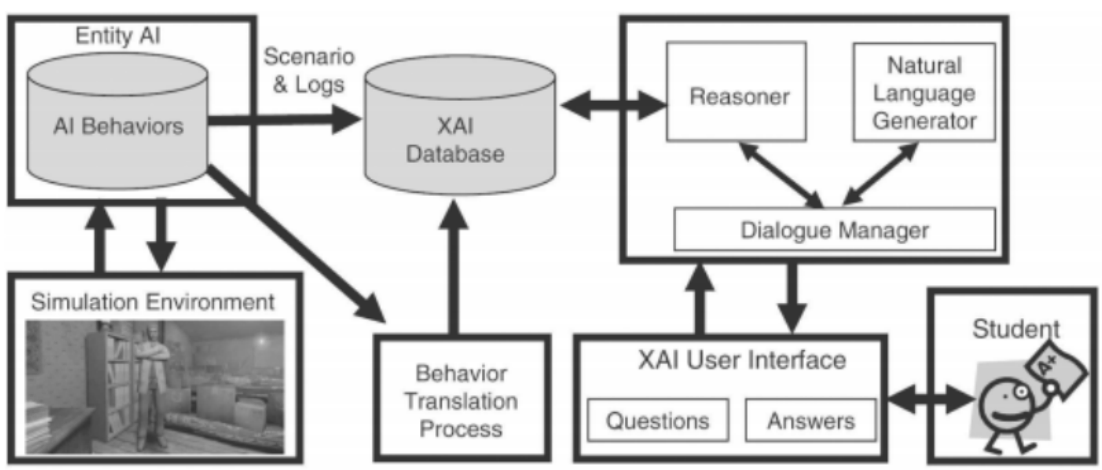

Figure 2.2: XAI Architecture for simulation game (Core et al., 2006)

For classification AI system using deep learning technique, which could be considered as a Black-box System, explaining the decision of classification is much more challenging than early AI system. In the domain of deep visual recognition, the model, Hendricks et al. (2016) propose, is a combination of classification and explanation (see Figure 2.3). The deep fine-grained classifier is used to extract visual features. For sentence generation, LSTM is used to generate sequence.

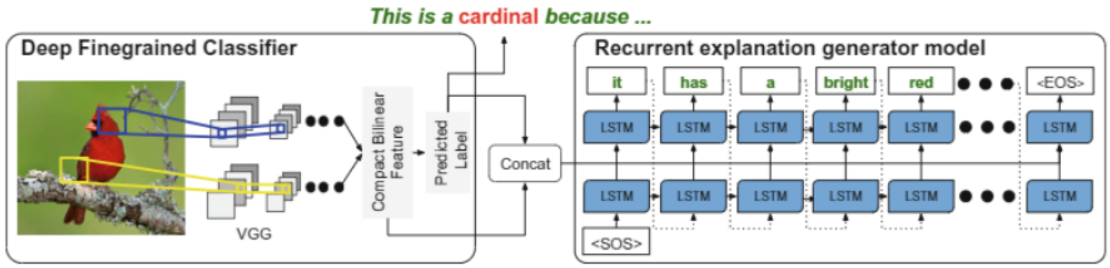

Figure 2.3: XAI Architecture for deep visual recognition (Hendricks et al., 2016)

The model of Hendricks et al. (2016) learns from the fine-grained visual descriptions (Figure 2.4, bottom left) and visual features (Figure 2.4, top left). Discriminative Loss function based on reinforcement learning could encourage to generate sentence that includes class specificity. In addition, relevance loss function (Figure 2.4, bottom right), which is defined for a specific image and caption, is applied to generate sentences that are image relevant.

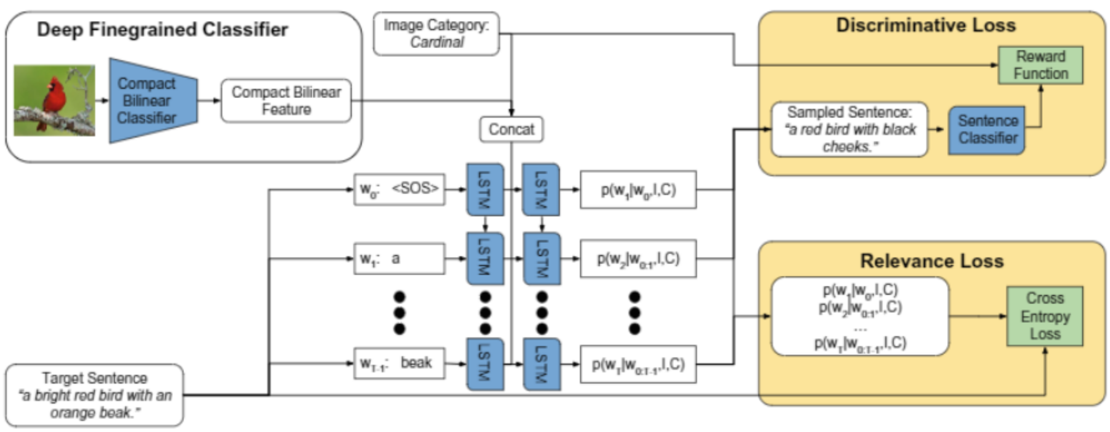

Figure 2.4: XAI Architecture for deep visual recognition (Hendricks et al., 2016)

In the critical decision-making domain, such as human life, AI application is often impeded since not be able to provide a basis inference to explain the reason of decision. To address this issue, Fuji et al. (2019) develop a technology which combines Deep Tensor neural network, a neural classifier originated by Fujitsu for graph data, with knowledge graph technology. The reason of using knowledge graph is that, in the domains of medicine, biology, chemistry et al., there are lots of datasets which could be encoded by graph to presents the interactions between different entities, e.g., diseases, genes, drugs. For this kind of technology, as shown in Figure 2.5, its first step is to use Deep Tensor to recognize the partial graphs which has the greatest influence in the in the inference result. In next step, it builds a correlation between partial graphs and nodes in knowledge graph. Then, it extracts those information from knowledge graph and presents them to user.

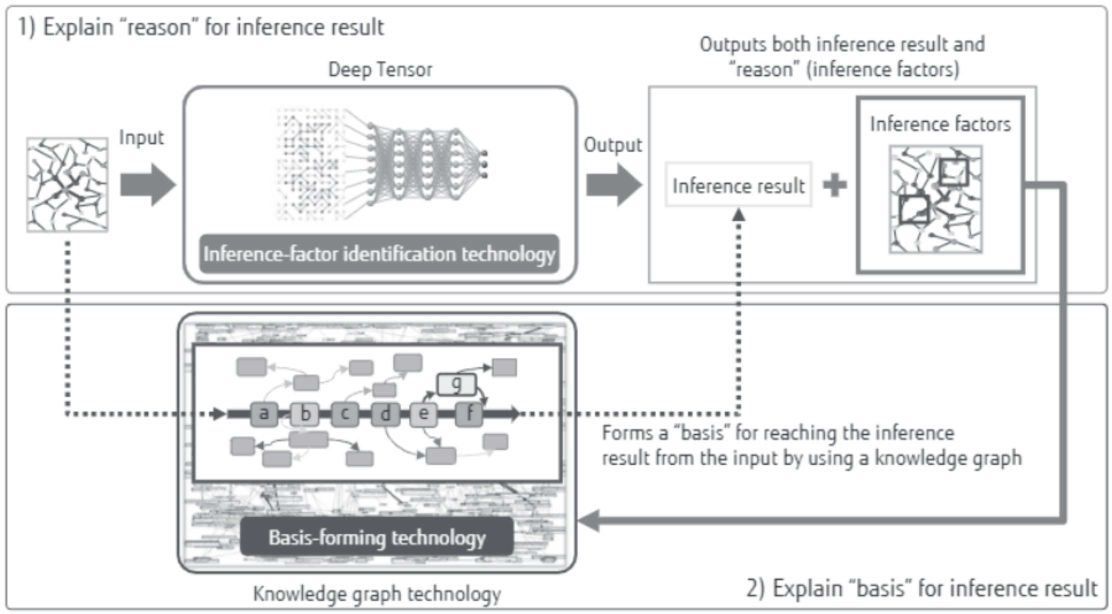

Figure 2.5: XAI Architecture for decision-making (Hendricks et al., 2016)

2.3 Explaining Combinatorial Search

Since machine/deep learning is treated as a black box, most of research of Explainable Artificial Intelligence (XAI) focus on this domain. However, the systems based on deep search also face the same issue (Weld et al., 2018). Especially for large space search, it seems really essential to explain its decisions for user. In these cases, what kinds of question need to be answered and how to response is a crucial problem. To address this issue, Fox et al. (2017) present a discussion of this problem in AI planning domain. They give a list of some questions which characterise what AI planner needs to explain.
Q1: Why did you do that?
For the first kind of question, Fox et al. (2017) state that because some plan is long and complex, questioner can not immediately be aware that later action should be supported by the target action. Hence, the example like action A is in the plan to allow this application of action B could be used as an answer. However, in some case, this causality is not evident, as some practical issue need professional knowledge to explain the causal relationship.
Q2: Why did not you do something else?
The question Q2 is same as the intention of Q1. However, to response Q2, an alternative action need to be carried out (Fox et al., 2017). And the answer to this question, in essence, is a presentation of the shortage of the plan which applies the alternative action. Therefore, the answer could be described as one plan is cheaper or shorter than another one (Fox et al., 2017). To execute the plan suggested by human, the AI planner need to modify its algorithm to search the state which matches the plan of human. There are two possible outcomes: one is that planner finds the plan which achieve a different or same goal (Figure 2.6, left)

according to the suggestion of human. The other outcome is that planner fails to find an alternative plan, as shown in Figure 2.6 right.

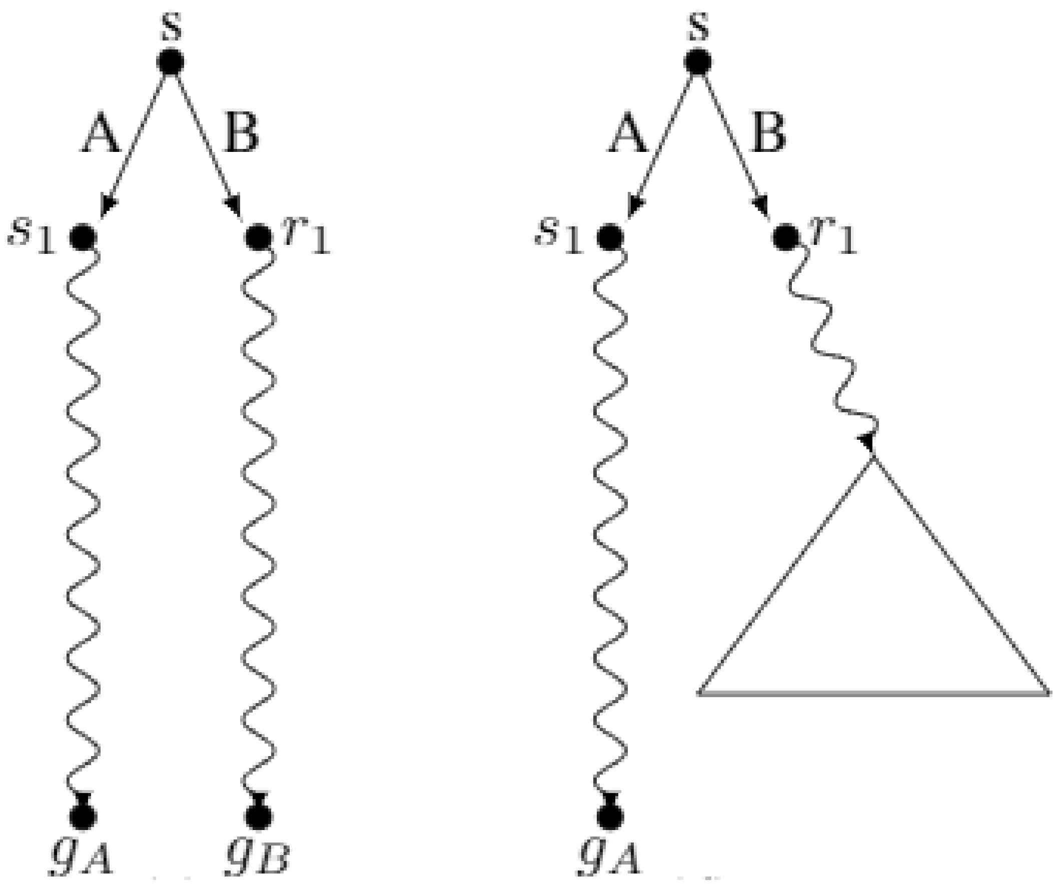

Figure 2.6: XAI Architecture for decision-making (Hendricks et al., 2016)

Q3:Why can not you do that?
Fox et al. (2017) consider that this kind of question would be asked when AI planner can not find a plan for a problem. There are two specific reasons for this question. One is that the current state could not satisfy the precondition of the action user suggests. The other reason is that if apply that action, planner would not achieve the goal from the resulting state. For the first case, Fox et al. (2005) propose a validator VAL, a validating models makes event triggered when its preconditions is in the state of true, to provide the explanation. For the second case, however, it is a challenging task to provide the explanation. Because, the question: why the goal cannot be achieved at all (Fox et al., 2017) is a planning problem which is unresolvable. But there are still some promising directions which can be found in the work of proving plan non-existence.

2.4 Summary

Lorem . ...

Chapter 3

Analysis

Lorem ipsum ....

3.1 Project Requirements

The basic requirement of Explainable artificial intelligence for Gomoku AI based on game tree search is that XAI system could answer two kinds of questions:
Q1: Why did you do that?
The essence of this kind of question is that XAI system could automatically generate the related explanation for every move decided by program.The crucial part of explaining is to find the threat sequences (e.g., if the opponent constructs a four, in next move, he threatens to win) of both players. For attacking moves, XAI should consider how to win the game; how to form threat sequences as possible as. When it is on defense, XAI needs to explain why this move could be used to stop the opponent from forming threaten sequence.
Q2: Why didnt you do that?
When user raises this kind of question, XAI needs to execute the move suggested by user and compare it with the move decided by program. Since the AI always consider a few steps advance and then select the best move, it may be really unfriendly for user when present huge simulate moves of both players. Therefore, XAI need to make it as briefly as possible. Additionally, in this part, it is really easy to begin an argument between the user and XAI. But to complete a nature argument dialogue system is a extremely hard task. So it could be substituted by the simple scheme: let user click the button to select whether they are satisfied the explanation. If user are not satisfied, XAI could strengthen the depth of search and give more detailed explanation.

For the requirements of the quality of explanation, XAI system could use the pictures, gif and text to complete the explaining task. To generate explaining context, setting up template is a suitable method but it is a bit of a mechanical. A better method is to apply the technique called long short-term memory (LSTM) which is an artificial recurrent neural network learning to store information and generate natural language (Hochreiter et.al., 1997). LSTM is a deep learning technique which needs enormous training data. However, so far, I have not find an existing Gomoku corpora which could be used to train deep learning model. Therefore, the basic requirement for generating explaining text is to use template. Optional bonus is to artificially build an Gomoku corpora based on knowledge graph and implement LSTM model.

There are always various different expressions for the same kind of question. Hence, correctly identifying the aim of users question is essential for XAI system. A mechanical method is only providing two choices of question: Button 1 represents Q1; Button 2 represents Q2. This method is easy but seems not enough intelligent. Applying machine learning technology, e.g., structured perceptron method with the support of query training data, to classify different questions. Additionally, user might ask some open question, e.g., raise query about the rules of Gomoku. Therefore, implementing a question answering system over RDF dataset (See Figure 3.1) based on Gomoku knowledge could be an optional bonus feature for XAI system.

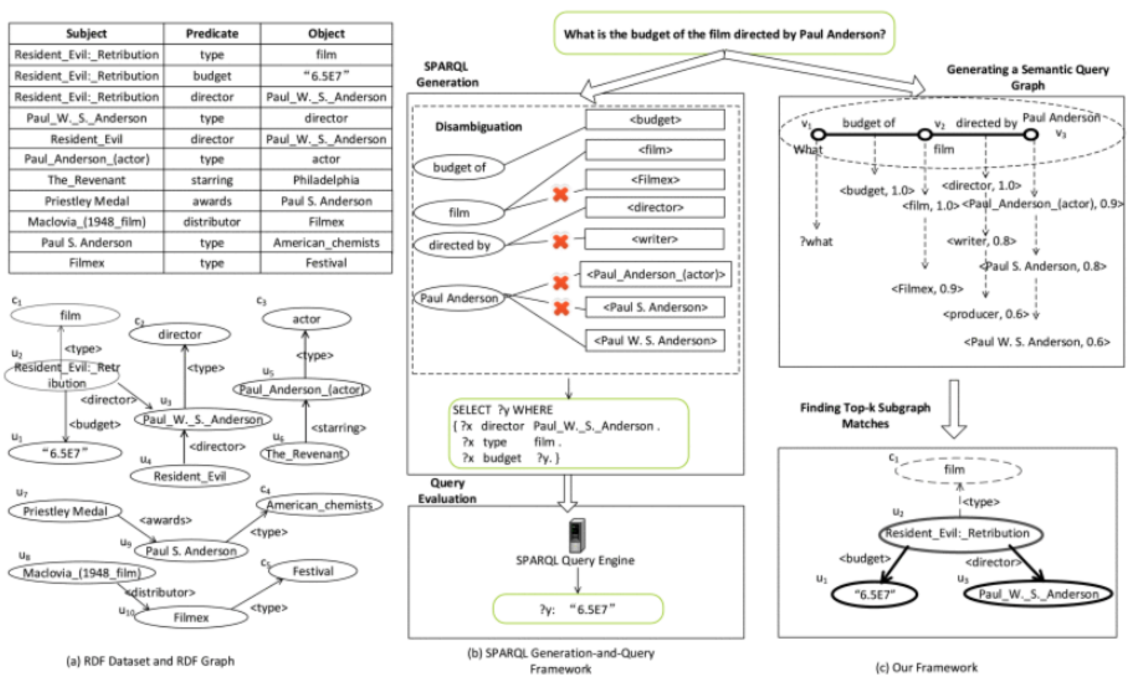

Figure 3.1: Question answering over RDF dataset (Hu et.al., 2018)

Since Explainable AI is not directly measurable, the evaluation of XAI rely instead on measurable outcomes (Luca and Feddy, 2019). In addition, the aim of Explainable AI is to explain the behaviour of AI to human. Therefore, human based evaluation, such as usefulness, relevance, is essential. When Invite a group of people to test the performance of XAI, there are three aspects need human to evaluate. Firstly, whether XAI has correctly understanding the type of question: Q1 or Q2; Secondly, whether the content of explanation is accurate and useful. Finally, whether the answer of XAI is nature and user-friendly.

3.2 Ethical, Professional and Legal Issues

Lorem ipsum ....

Chapter 4

Planning

4.1 Risk Analysis

Answering Question 2: Why didnt you do that? is likely a hard question. The Gomoku AI based on game search tree always consider a few steps advance and then select the best move. Since the complexity of search space grows exponentially with deeper search, we always use some algorithm, such as alpha-beta pruning, to speed up. However, in order not to find the move suggested by user, XAI system have to violent search all the possible moves which is really time consuming. Therefore, I assign two weeks for this task to make sure I have enough time to solve this problem.

Applying long short-term memory (LSTM) recurrent neural network and building knowledge graph of Gomoku are new things for me. I am not sure whether those technology could be implemented successfully. Hence, I prepare to make setting up template for each kind of question as an alternative technique to ensure XAI system run successfully.

4.2 Project Plan

Task 1 (Week 1):
construct the code of Gomoku AI based on game tree search.
Task 2 (Week 2):
modify the algorithm to make Gomoku XAI generate explanation for Q1:Why did you do that?
Task 3 (Week 3 and 4):
modify the algorithm to make Gomoku XAI answer Q2: Why didnt you do that?
Task 4 (Week 5 and 6):
build knowledge graph of Gomoku.
Task 5 (Week 7 and 8):
try to combine the technique of knowledge graph amd LSTM model to make the answer more nature
Task 6 (Week 9 and 10):

invite people to experience Gomoku XAI and modify the program and add some new functions according to the feedback.
Task 7(Week 11 and 12):
write the report.

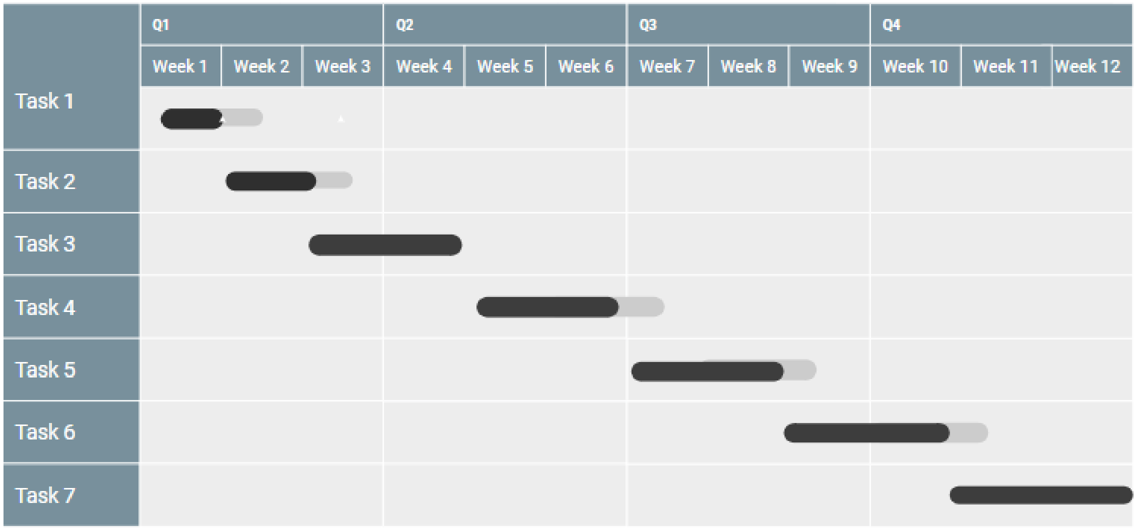

Figure 4.1: Gantt Chart

Chapter 5

Conclusions

Lorem ipsum ......

Bibliography

[1] Allis, L.V., Herik, H.J. and Huntjens, M.P.H., 1993. Go-moku and threat-space search. University of Limburg, Department of Computer Science.

[2] Samek, W., Wiegand, T., Mller, K. R. (2017).Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *In AAAI* .(pp. 1766-1773).

[3] Fuji, M., Morita, H., Goto, K., Maruhashi, K., Anai, H. and Igata, N., 2019. Explainable AI Through Combination of Deep Tensor and Knowledge Graph. *FUJITSU SCIENTIFIC TECHNICAL JOURNAL, 55(2),*.pp.58-64.

[4] Fox, M., Long, D. and Magazzeni, D., 2017. Explainable planning. *arXiv preprint arXiv:1709.10256.*.

[5] Fox, M., Howey, R., Long, D. (2005, July). Validating plans in the context of processes and exogenous events. In *AAAI.*(Vol. 5, pp. 1151-1156).

[6] Hendricks, L.A., Akata, Z., Rohrbach, M., Donahue, J., Schiele, B. and Darrell, T., 2016, October. Generating visual explanations. In *European Conference on Computer Vision.* Springer, Cham.

[7] Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. *Neural computation, 9(8),* .pp.1735-1780.

[8] Hendricks, L.A., Akata, Z., Rohrbach, M., Donahue, J., Schiele, B. and Darrell, T., 2016, October. Generating visual explanations. In *European Conference on Computer Vision.* Springer, Cham.

[9] Hu, S., Zou, L., Yu, J.X., Wang, H. and Zhao, D., 2018. Answering natural language questions by subgraph matching over knowledge graphs. *IEEE Transactions on Knowledge and Data Engineering, 30(5),*. pp.824-837.

[10] Luca C., Freddy L., 2019. Tutorial on Explainable AI. *AAAI https://xaitutorial2019.github.io/* . pp.37

[11] Schaeffer, J., 1989. The history heuristic and alpha-beta search enhancements in practice. *IEEE Transactions on Pattern Analysis Machine Intelligence, (11),*. pp.1203-1212.

[12] Weld, D.S. and Bansal, G., 2018. The challenge of crafting intelligible intelligence. *arXiv preprint arXiv:1803.04263.*.

Appendices

Appendix A

An Appendix of Some Kind

Lorem ipsum ......

Appendix B

Another Appendix

Lorem ipsum ...