# PLAYING YAHTZEE WITH DEEP REINFORCEMENT LEARNING - A SYSTEMATIC COMPARISON OF DIFFERENT APPROACHES

A PREPRINT

**Markus T. Dutschke***
auXolutions
Erlangen, Germany
markus.dutschke@auXolutions.de

**Elias D. Striatum**
Department of Electrical Engineering
Mount-Sheikh University
Santa Narimana, Levand
stariate@ee.mount-sheikh.edu

January 14, 2020

## ABSTRACT

In this paper we present an open source Q-learning algorithm for the dice game yahtzee. We implemented a variation of the Q-learning algorithm as used by Mnih [**?**], which he used for playing Atari games. The specific obstacles of yahtzee are thereby to handle two different types of possible actions: 1) choose what dice to re-roll; 2) choose a category on the score board, the significantly larger number possible actions of type 1 compared to an Atari game controller and the randomness in the response of the game to the players actions of type 1. By presenting different implementations of increasing complexity, we give the reader an overview of different concepts to improve the performance of Q-learning for certain situations and evaluate their performance in the specific use case. Among those concepts are different exploration strategies, concepts to handle randomness and a technique for the efficient handling of the two decision types. The most successful implementation achieves superhuman performance within a few thousand training cycles.

***Keywords*** Q-learning · neural networks · exploration strategies · replay memory

## 1 Introduction

The complete source code of this project is publicly available at

`https://github.com/markusdutschke/yahtzee`

Since Mnihs famous publication 'Playing Atari with deep reinforcement learning' [**?**] strong research interest has evolved around the possibilities of Q-learning in combination with neural networks. Thereby computer and board games turned out to be an excellent playground for this research, due to their complex character, their easy reproducibility and the clear definition of the systems rules.

Especially the dice game Yahtzee has a set of interesting properties, which makes it especially suitable for a Q-learning test system.

- Yahtzee is a broadly known game.
- Even after several hundred games, Yahtzee is still challenging for human player.
- There is a mixture of randomness and strategy involved.
- Yahtzee is exactly solved, but this is far beyond the score of the abilities of a human brain.

*Use footnote for providing further information about author (webpage, alternative address)—*not* for acknowledging funding agencies.

# 2 The dice game Yahtzee

## 2.1 Rules

## 2.2 Code

Relate the rule with the different classes in the code. Especially: Dice, ScoreBoard, Game and AbstractPlayer

## 2.3 Existing literature

# 3 Q-learning

This chapter contains all the theoretical background of the code.

## 3.1 Background

## 3.2 Handling the two decision types

## 3.3 Information encoding

## 3.4 Exploration

## 3.5 Concepts to handle a stochastic system response

# 4 Implementations

This chapter is just a description of the different implementations and their performance.

## 4.1 Naive Implementations

- random implementation - greedy implementation with and without re-roll

## 4.2 AI player Version 0

## 4.3 AI player Version 1

## 4.4 AI player Version 2

# 5 Benchmark

In this chapter the benefit of different Q-learning concepts are quantitatively benchmarked. The implementation of these benchmarks can be found in the functions bench... in main.py with player implementations in botBench.py

## 5.1 Information encoding

Different encodings. Not yet sure, what to compare. Maybe: rgrSC with - 13 inputs (-1 for empty, otherwise score) - 26 inputs (first 13: 0 for empty, second 13: 0 or 1 for empty and used) - a good encoding (check maybe player v2)

## 5.2 Exploration

- epsilon greedy - softmax - minMaxRat

## 5.3 Concepts to handle a stochastic system response

- implicitly in MLP regressor (v0) - explicitly in mlprgr with pretraining and benchmarking (this is v1) - exactly by lookup table (v2)

# 6 Conclusion

Collection of key facts, whatever turned out to bring the most significant improvement.

# 7 NOW FOLLOWS THE TEMPLATE

# 8 Headings: first level

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

See Section 8.

## 8.1 Headings: second level

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

$$\xi_{ij}(t) = P(x_t = i, x_{t+1} = j | y, v, w; \theta) = \frac{\alpha_i(t) a^{w_t}_{ij} \beta_j(t+1) b^{v_{t+1}}_j (y_{t+1})}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i(t) a^{w_t}_{ij} \beta_j(t+1) b^{v_{t+1}}_j (y_{t+1})} \tag{1}$$

### 8.1.1 Headings: third level

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

**Paragraph** Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

# 9 Examples of citations, figures, tables, references

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

[1, 2] and see [3].

The documentation for `natbib` may be found at

`http://mirrors.ctan.org/macros/latex/contrib/natbib/natnotes.pdf`

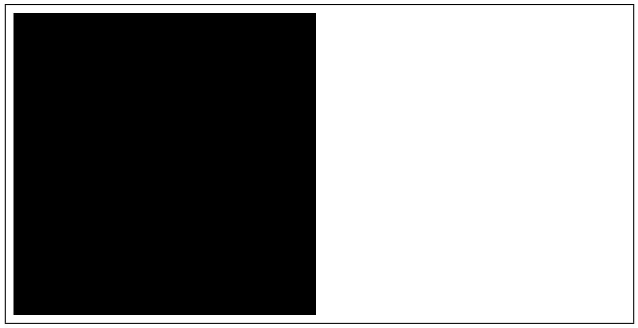

Figure 1: Sample figure caption.

Table 1: Sample table title

| Part | | |
|---|---|---|
| Name | Description | Size ($\mu$m) |
| Dendrite | Input terminal | $\sim$100 |
| Axon | Output terminal | $\sim$10 |
| Soma | Cell body | up to $10^6$ |

Of note is the command `\citet`, which produces citations appropriate for use in inline text. For example,

```
\citet{hasselmo} investigated\dots
```

produces

> Hasselmo, et al. (1995) investigated…

`https://www.ctan.org/pkg/booktabs`

### 9.1 Figures

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

See Figure 1. Here is how you add footnotes. [2] Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetuer eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

### 9.2 Tables

Etiam euismod. Fusce facilisis lacinia dui. Suspendisse potenti. In mi erat, cursus id, nonummy sed, ullamcorper eget, sapien. Praesent pretium, magna in eleifend egestas, pede pede pretium lorem, quis consectetuer tortor sapien facilisis magna. Mauris quis magna varius nulla scelerisque imperdiet. Aliquam non quam. Aliquam porttitor quam a lacus. Praesent vel arcu ut tortor cursus volutpat. In vitae pede quis diam bibendum placerat. Fusce elementum convallis neque. Sed dolor orci, scelerisque ac, dapibus nec, ultricies ut, mi. Duis nec dui quis leo sagittis commodo.

See awesome Table 1.

### 9.3 Lists

- Lorem ipsum dolor sit amet

[2] Sample of the first footnote.

- consectetur adipiscing elit.
- Aliquam dignissim blandit est, in dictum tortor gravida eget. In ac rutrum magna.

## References

[1] George Kour and Raid Saabne. Real-time segmentation of on-line handwritten arabic script. In *Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on*, pages 417–422. IEEE, 2014.

[2] George Kour and Raid Saabne. Fast classification of handwritten on-line arabic characters. In *Soft Computing and Pattern Recognition (SoCPaR), 2014 6th International Conference of*, pages 312–318. IEEE, 2014.

[3] Guy Hadash, Einat Kermany, Boaz Carmeli, Ofer Lavi, George Kour, and Alon Jacovi. Estimate and replace: A novel approach to integrating deep neural networks with existing applications. *arXiv preprint arXiv:1804.09028*, 2018.