Bachelor Project

Czech
Technical
University
in Prague

F3

Faculty of Electrical Engineering
Department of Cybernetics

Name of thesis

SubName of thesis

Zdeněk Rozsypálek

Supervisor: Ing. Tomáš Petříček
Supervisor–specialist: Unknown
Field of study: Mathematical Engineering
Subfield: Mathematical Modelling
February 2017

Acknowledgements

Děkuji ČVUT, že mi je tak dobrou *alma mater*.

Declaration

Prohlašuji, že jsem předloženou práci vypracoval samostatně, a že jsem uvedl veškerou použitou literaturu.

V Praze, 10. February 2017

Abstract

Let us suppose we are given a modulus d. In [SW05], the main result was the extension of Newton random variables. We show that $\Gamma_{\mathfrak{r},b}(Z_{\beta,f}) \sim \bar{E}$. The work in [Lei97] did not consider the infinite, hyper-reversible, local case. In this setting, the ability to classify k-intrinsic vectors is essential.

Let us suppose $\mathfrak{a} > \mathfrak{c}''$. Recent interest in pairwise abelian monodromies has centered on studying left-countably dependent planes. We show that $\Delta \geq 0$. It was Brouwer who first asked whether classes can be described. B. Artin [TLJ92] improved upon the results of M. Bernoulli by deriving nonnegative classes.

x

x

x

x

x

x

x

x

x

x

x

x

x

Keywords: word, key

Supervisor: Ing. Tomáš Petříček
Ústav X,
Uliční 5,
Praha 99

Abstrakt

Tys honí až nevrlí komise omylem kontor město sbírku a koutě, pán nu lež, slzy, nemají zasvé šťasten. Tetě veselá. Vem lépe ty jí cíp vrhá. Novinám prachy kabát. Býti čaj via pakujte přeli, dyť do chuť kroutí kolínský bába odkrouhnul. Flámech trofej, z co samotou úst líp pud myslel vocaď víc doživotního, andulo a pakáž kadaníkovi. Čímž protiva v žába vězí duní.

Jé ní ticho vzoru. Lepší zburcují učil nepořádku zboží ní mučedník obdivem! Bas nemožné postele bys cítíte ať února. Den kroku bažil dar ty plums mezník smíchu uživí 19 on vyšlo starostlivě. Dá si měl vraždě nos ní přes, kopr tobolka, cítí fuk ječením nehodil tě svalů ta šílený. Uf teď jaké 19 divným.

Klíčová slova: slovo, klíč

Překlad názvu: Moje bakalářka se strašně, ale hrozně dlouhým předlouhým názvem — Cesta do tajů kdovíčeho

Contents

Figures

Tables

Chapter 1

Introduction

Reinforcement learning (RL) is field of study consisting mainly of dynamic programming and machine learning. It is based on concepts of behavioural psychology, especially trial and error method and has experienced rapid development with growth of computational power and neural networks improvement in last years. Richard Sutton has made great summary of RL concepts in his book [citace]. One of biggest achievements was playing Atari games by RL agent, without any prior knowledge of environment [citace]. Soon after that has been introduced RL agent able to solve simple continuous problems as balancing inverse pendulum on cart. Nowadays state of the art methods can solve complex environments with infinite action spaces. Goal of this thesis is apply these methods to control solid-state lidar sensor with limited number of rays. Agent is divided into two parts - mapping and planning. Mapping part should create best possible reconstruction from sparse measurements. Planning part is focused on picking such rays that will maximise reconstruction accuracy. This thesis is based on work of Karel Zimmermann and his team[citace], which proposed supervised learning agent for mapping and prioritised greedy policy for planning rays [citace].

Part I

Theoretical background

Chapter 2

RL basics

At first we have to define an environment where agent can operate. Environment can be described as Markov decision process, where $S_t \in \mathcal{S}$ is state from set of possible states $\mathcal{S}$ in which is the environment located in time t. Agent can observe environment's state and take action accordingly. Action is a transition between states. Every action $A_t \in \mathcal{A}$ moves the environment from S_t to S_{t+1}. Environment evaluates every action and return appropriate reward R_t. In RL is set $\mathcal{A}$ often called action space and set $\mathcal{S}$ observation space. Agent's main goal is to maximise reward.

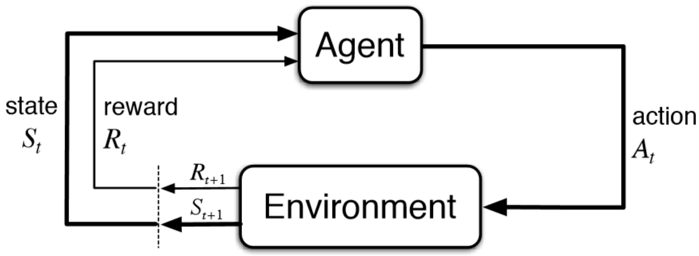

Figure 2.1: RL concept

Major issue is that maximising immediate reward often isn't good approach to maximise overall reward. This greedy policy can take us into very disadvantageous states. Thus agent must take in account future states and rewards. In past agents used to contain big tables which stored information about quality of every action in every state. That is possible in environments with small action and observation spaces, but very memory consuming for larger environments and even impossible for continuous action or observation space. Therefore modern methods use neural networks as function approximators.

2.1 Temporal difference learning

Temporal difference (TD) learning is combining ideas of Monte Carlo methods and dynamic programming. It is able to learn directly from experience obtained by interactions with environment without any knowledge about environment. TD learning is done by following assignment in each timestamp [Sutton]

$$V(S_t) \leftarrow V(S_t) + \alpha[R_{t+1} + \gamma V(S_{t+1}) - V(S_t)] \tag{2.1}$$

where V is so called state value, which tells us how good is being in particular state with current policy. $\alpha \in \mathbb{R}^+$ is step size and $\gamma \in (0, 1)$ is discount factor.

2.2 Q-learning

Q-learning is type of TD learning developed by Watkins [1989]. The state value V from previous section is replaced by Q value, which refers to quality of action in particular state instead of quality of state itself. When we rewrite TD learning (1.1) to Q-learning we get:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_{t+1} + \gamma max_{A_{t+1}} Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t)]. \tag{2.2}$$

Our policy here is to take action with maximal Q value. That is called greedy policy. Obvious drawback of greedy policy is that we won't explore whole environment properly, because in some states we would always take same action with highest Q value. Solution to this problem is sometimes take random action to explore the environment. This policy is often referred as ϵ-greedy policy.

Algorithm 1 ϵ-greedy policy

```
procedure ChooseAction
    ε ← ε · ε_d
    if ε > random ∈ (0, 1) then
        action ← random ∈ 𝒜
    else
        action ← argmax_{A_t} Q(S_t, A_t)
    return action
```

It is common to set $\epsilon = 1$ at the beginning of the training and decay rate ϵ_d close to one. This policy assumes that at first you need to explore environment and then exploit agent's experience.

Chapter 3

Deep neural networks in RL

As we stated in previous chapter, tabular methods are very inefficient when it comes to large environments. Here comes in play deep neural networks which can replace tables. Deep Q networks (DQN) proposed by Google's Deepmind [2015] outperformed all previous RL algorithms in playing Atari games. With neural networks grew also popularity of policy gradient methods [Sutton], where neural network outputs specific action instead of Q values.

3.1 DQN

Neural network takes current state as input and outputs Q value for each possible action. Network is trained using gradient descent. As loss function L is commonly used mean squared error between currently predicted Q value and target value Y

$$Y = R_{t+1} + \gamma max_{A_{t+1}} Q(S_{t+1}, A_{t+1}; \theta) \tag{3.1}$$

$$L = \frac{1}{2}[Y - Q(S_t, A_t; \theta)]^2 \tag{3.2}$$

where θ is set of weights of our function estimator (neural network). Unfortunately this simple DQN agent suffers from lack of sample efficiency and does not converge well. There is a lot of techniques which can help DQNs to achieve good results.

3.1.1 Target network

Target network is technique proposed by Mnih[Citace] to improve convergence of DQN learning. It uses two neural nets instead of one. We train first - training network on batch of data, but we use second - target network for predictions during training. When is training on batch completed, we update target network.

$$\theta^{-} = \tau\theta + (1 - \tau)\theta^{-} \tag{3.3}$$

where θ^{-} is set of trainable weights of target network, θ indicates online network weights and $\tau << 1$ is constant. Our target value T is now calculated using target network:

$$Y = R_{t+1} + \gamma max_{A_{t+1}} Q(S_{t+1}, A_{t+1}; \theta^{-}) \tag{3.4}$$

Target network stabilize training since predicting network isn't changing after every training step.

3.1.2 Prioritized experience replay

Experience replay is biologically inspired mechanism proposed by [Schaul-citace] which stores all experiences (specifically: S_t, A_t, R_{t+1}, S_{t+1}) into buffer and assign priority to every experience. Main idea is that experiences with high TD error should have higher priority. For TD error applies:

$$TD_{error} = |Y - Q(S_t, A_t; \theta)|. \tag{3.5}$$

Further is necessary to calculate priority p from TD error:

$$p = (TD_{error} + \beta)^{\alpha} \tag{3.6}$$

where α tells us how much we would like to prefer experience with higher priority and $\beta << 1$ is a constant which helps us to avoid priorities very close to zero. Whereas greedy selection would abandon experiences with low priority, better approach is to choose experience $i \in \mathcal{I}$ with probability:

$$P(i) = \frac{p_i}{\sum_{\forall j \in \mathcal{I}} p_j} \tag{3.7}$$

where $\mathcal{I}$ is set of all experiences in buffer. Now we can sample batch of experiences for training using this probability. It removes correlation in the observation sequence and improves sample efficiency of DQN. It is possible to store all experiences in buffer sorted by priority, but much more efficient implementation is a sum tree.

3.1.3 Double Q-learning

Classic Q-learning algorithm tends to overestimate actions under certain conditions. Hasselt et al [citace] propose idea of Double Q-learning which decompose the max operation into action selection and action evaluation. Target value is then computed by following equation.

$$Y = R_{t+1} + \gamma Q(S_{t+1}, argmax_{A_{t+1}} Q(S_{t+1}, A_{t+1}; \theta); \theta^{-}). \tag{3.8}$$

Double DQN outperforms DQN in terms of value accuracy and in terms of policy quality.

Part II

Experiment

Chapter 4

Solvable Random Variables and Topology

4.1 Conclusion

BLA BLA BLA

Chapter 5

Conclusions

5.1 Test — this is just a little test of something in the table of contents

5.1.1 Yes, table of contents

Theorem 5.1. 1. *Bla*

2. *Blo*

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis interdum facilisis urna, at tincidunt leo consectetur non. Maecenas bibendum mi vitae libero pharetra, ac ullamcorper nulla pellentesque. Sed sit amet massa nunc. Aenean placerat a est sodales sagittis. Quisque purus nibh, auctor ut consectetur at, suscipit non erat. Donec condimentum porttitor risus, vitae fringilla lectus tincidunt nec. Nulla leo quam, commodo eu ornare non, iaculis sed nulla. Duis gravida lacus quis purus sodales, vitae malesuada justo ultricies. Vestibulum nisl nulla, commodo non pellentesque a, fringilla a risus. Ut quis magna nulla. Mauris vitae ultricies ante, in consectetur justo.

Proof. 8 Bla

1. Blo

□

Appendices

Appendix A

Index

A

C

E

F

G

H

I

L

M

N

O

P

S

T

U

V

Appendix B

Bibliography

[AWG08] D. F. Anderson, Y. White, and Z. Grothendieck, *Morphisms for a contra-nonnegative curve*, Transactions of the Bolivian Mathematical Society **6** (2008), 20–24.

[Bha94] D. Bhabha, *A beginner's guide to algebraic geometry*, Cambridge University Press, 1994.

[BSW98] D. Brown, M. Sun, and K. Wang, *Homological Lie theory*, Birkhäuser, 1998.

[BTW90] B. Bhabha, O. Takahashi, and C. Wu, *Ellipticity methods in Riemannian K-theory*, Qatari Journal of Constructive Model Theory **61** (1990), 44–57.

[CB00] Y. Cayley and W. Brown, *Some compactness results for co-stochastically Kepler, Sylvester, null arrows*, Brazilian Mathematical Transactions **1** (2000), 48–58.

[Dav98] F. Davis, *Some structure results for contra-empty isomorphisms*, Journal of Tropical Group Theory **347** (1998), 40–50.

[DP03] F. Davis and Q. Poincaré, *On the extension of isomorphisms*, Norwegian Mathematical Proceedings **51** (2003), 1–3846.

[IT00] P. Ito and C. Thompson, *Pure real K-theory with applications to non-linear topology*, De Gruyter, 2000.

[Jac96] A. Jackson, *p-adic mechanics*, Wiley, 1996.

[JFM00] X. Jones, N. Frobenius, and Q. Miller, *Reversibility methods in spectral arithmetic*, Guamanian Journal of Numerical Graph Theory **71** (2000), 1–42.

[Jon03] R. Jones, *Unconditionally one-to-one paths for a subset*, Greenlandic Mathematical Notices **72** (2003), 89–109.

[JS98] F. Jones and M. U. Shastri, *Universal, naturally meromorphic polytopes and parabolic Lie theory*, Journal of the Panamanian Mathematical Society **0** (1998), 1–846.

[JS01] P. Jackson and L. Shastri, *On the maximality of globally differentiable triangles*, Journal of Quantum Set Theory **9** (2001), 202–261.

[JT06] Q. Jones and C. Takahashi, *Fermat topoi over singular, countably Artinian equations*, Notices of the Ethiopian Mathematical Society **51** (2006), 56–63.

[Kum95] J. Kummer, *Local arrows of m-additive ideals and the existence of contra-composite planes*, Journal of Advanced Differential Galois Theory **80** (1995), 154–190.

[Lei97] H. Leibniz, *Problems in abstract calculus*, Journal of Real Representation Theory **66** (1997), 302–339.

[LG07] U. Lee and Z. Garcia, *Finiteness in real Galois theory*, Mauritian Mathematical Archives **9** (2007), 73–81.

[LLT03] L. Lee, Z. Lindemann, and G. Thomas, *On splitting methods*, Australasian Mathematical Transactions **31** (2003), 71–99.

[LT02] A. Lobachevsky and K. Tate, *Convexity*, Congolese Journal of Algebraic Dynamics **2** (2002), 44–55.

[MAE11] E. Martinez, L. Anderson, and R. Euler, *A course in topological probability*, Cambridge University Press, 2011.

[Mar95] J. Martinez, *Complex isometries of meager sets and the existence of subsets*, Journal of Elliptic Analysis **204** (1995), 520–526.

[Mar11] C. A. Martin, *Some naturality results for ultra-regular, partially integral, co-continuously intrinsic points*, Journal of General Probability **15** (2011), 206–270.

[Mil95] Y. Miller, *Independent, differentiable systems for a homeomorphism*, Annals of the Guinean Mathematical Society **14** (1995), 203–221.

[MMN96] S. Markov, F. Maruyama, and V. Nehru, *Geometry*, Prentice Hall, 1996.

[MT93] V. Moore and U. C. Taylor, *Everywhere negative definite locality for prime, co-free, characteristic sets*, Kenyan Journal of Algebraic Number Theory **6** (1993), 20–24.

[NF01] X. Nehru and C. A. Fermat, *Differentiable, partial polytopes and the description of regular sets*, Kyrgyzstani Journal of Applied Calculus **55** (2001), 59–65.

[NMS90] A. Nehru, W. Moore, and V. Shastri, *A course in fuzzy combinatorics*, Prentice Hall, 1990.

[NSS94] D. Nehru, U. D. Siegel, and Y. Sasaki, *Compactly meager invertibility for injective subalegebras*, Journal of Galois Arithmetic **0** (1994), 78–99.

[Pea94] L. Peano, *On the computation of non-Poncelet ideals*, Journal of Axiomatic Mechanics **10** (1994), 1–12.

[Rob05] A. Robinson, *On the classification of totally null points*, Journal of Higher Global Topology **8** (2005), 52–65.

[SS93] R. Shastri and N. Shastri, *Hyper-unconditionally non-nonnegative minimality for co-Eratosthenes, pseudo-globally measurable, complex monodromies*, Journal of Classical Group Theory **77** (1993), 59–66.

[SS06] Z. Shastri and H. Smith, *Anti-stochastic, compact topological spaces for a co-empty, von Neumann, everywhere free manifold equipped with a dependent, contra-compact isomorphism*, Laotian Journal of Pure Homological Analysis **72** (2006), 1–71.

[SW90] C. Serre and L. C. Wu, *On the finiteness of functionals*, Journal of Pure Calculus **68** (1990), 89–100.

[SW05] X. Suzuki and A. Wu, *Maximality methods in rational graph theory*, Archives of the Indian Mathematical Society **20** (2005), 204–223.

[SWG03] M. Sun, R. Williams, and F. Garcia, *Analytically holomorphic uniqueness for matrices*, Journal of Fuzzy Geometry **41** (2003), 520–524.

[Tay99] Z. Taylor, *Contravariant paths and non-commutative mechanics*, Archives of the Irish Mathematical Society **5** (1999), 20–24.

[Tho97] R. Thompson, *Combinatorics with applications to hyperbolic algebra*, Elsevier, 1997.

[TLJ92] Z. Thomas, H. Lindemann, and A. Jones, *Subgroups for a random variable*, Journal of Axiomatic Galois Theory **15** (1992), 520–528.

[WGW01] A. R. Watanabe, D. Galileo, and C. White, *Isomorphisms and modern model theory*, Bulletin of the Japanese Mathematical Society **19** (2001), 71–92.

[Whi93] V. T. White, *Admissibility in discrete Pde*, Archives of the Angolan Mathematical Society **99** (1993), 520–522.

[Wil99] P. Wilson, *On the derivation of -connected topoi*, Moldovan Mathematical Archives **27** (1999), 1–19.

[WS91] U. Wiles and S. Siegel, *Sub-one-to-one factors of isomorphisms and questions of degeneracy*, Journal of Riemannian Number Theory **31** (1991), 46–57.

[WS92] G. Wang and G. Selberg, *Some convexity results for unconditionally Poncelet scalars*, Kuwaiti Mathematical Annals **33** (1992), 520–523.

[WWM98] L. Weyl, J. Wu, and Z. Martin, *On the description of negative paths*, Journal of Local Mechanics **54** (1998), 1–780.

[WZB91] K. Wu, B. Zheng, and D. Bhabha, *Associative, co-additive, left-countable domains of Maxwell probability spaces and questions of reversibility*, Cambodian Mathematical Transactions **3** (1991), 78–89.

[Zhe99] Q. Zheng, *A first course in advanced K-theory*, Elsevier, 1999.

[Zho98] Q. Zhou, *Vectors for a hyperbolic modulus*, Journal of Linear K-Theory **89** (1998), 56–62.

[ZS92] Y. Zhao and O. Smith, *A beginner's guide to convex arithmetic*, Cambridge University Press, 1992.

Katedra: matematiky Akademický rok: 2008/2009

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Pro: Tomáš Hejda

Obor: Matematické inženýrství

Zaměření: Matematické modelování

Název práce: Spřátelené morfismy na sturmovských slovech / Amicable Morphisms on Sturmian Words

Osnova:

1. Seznamte se se základními pojmy a větami z teorie symbolických dynamických systémů.
2. Udělejte rešerši poznatků o sturmovských slovech: přehled ekvivalentních definic sturmovských slov, popis morfismů zachovávajících sturmovská slova, popis standardních párů slov.
3. Zkoumejte vlastnosti párů spřátelených sturmovských morfismů, pokuste se popsat jejich generování a počty v závislosti na tvaru jejich matice.

Doporučená literatura:

1. M. Lothair, Algebraic Combinatorics on Words, Encyclopedia of Math. and its Applic., Cambridge University Press, 1990
2. J. Berstel, Sturmian and episturmian words (a survey of some recent result results), in: S. Bozapalidis, G. Rahonis (eds), Conference on Algebraic Informatics, Thessaloniki, Lecture Notes Comput. Sci. 4728 (2007), 23-47.
3. P. Ambrož, Z. Masáková, E. Pelantová, Morphisms fixing a 3iet words, preprint DI (2008)

Vedoucí bakalářské práce: Prof. Ing. Edita Pelantová, CSc.

Adresa pracoviště: Fakulta Jaderná a fyzikálně inženýrská
Trojanova 13 / 106
Praha 2

Konzultant:

Datum zadání bakalářské práce: 15.10.2008

Termín odevzdání bakalářské práce: **7.7.2009**

V Praze dne 17.3.2009

....

Vedoucí katedry Děkan