

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Matjaž Pogačnik

**Spodbujevano učenje na impulznih
nevronskih mrežah**

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJSKI PROGRAM
PRVE STOPNJE
RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: prof. dr. Zoran Bosnić

Ljubljana, 2025

To delo je ponujeno pod licenco *Creative Commons Priznanje avtorstva-Deljenje pod enakimi pogoji 2.5 Slovenija* (ali novejšo različico). To pomeni, da se tako besedilo, slike, grafi in druge sestavine dela kot tudi rezultati diplomskega dela lahko prosto distribuirajo, reproducirajo, uporabljajo, priobčujejo javnosti in predelujejo, pod pogojem, da se jasno in vidno navede avtorja in naslov tega dela in da se v primeru spremembe, preoblikovanja ali uporabe tega dela v svojem delu, lahko distribuira predelava le pod licenco, ki je enaka tej. Podrobnosti licence so dostopne na spletni strani creativecommons.si ali na Inštitutu za intelektualno lastnino, Streliška 1, 1000 Ljubljana.



Izvorna koda diplomskega dela, njeni rezultati in v ta namen razvita programska oprema je ponujena pod licenco GNU General Public License, različica 3 (ali novejša). To pomeni, da se lahko prosto distribuira in/ali predeluje pod njenimi pogoji. Podrobnosti licence so dostopne na spletni strani <http://www.gnu.org/licenses/>.

Besedilo je oblikovano z urejevalnikom besedil L^AT_EX.

Kandidat: Matjaž Pogačnik

Naslov: Spodbujevano učenje na impulznih nevronske mrežah

Vrsta naloge: Diplomski naloga na univerzitetnem programu prve stopnje
Računalništvo in informatika

Mentor: prof. dr. Zoran Bosnić

Opis:

Besedilo teme diplomskega dela študent prepíše iz študijskega informacijskega sistema, kamor ga je vnesel mentor. V nekaj stavkih bo opisal, kaj pričakuje od kandidatovega diplomskega dela. Kaj so cilji, kakšne metode naj uporabi, morda bo zapisal tudi ključno literaturo.

Title: Reinforcement learning on spiking neural networks

Description:

opis diplome v angleščini

Na tem mestu zapišite, komu se zahvaljujete za pomoč pri izdelavi diplomske naloge oziroma pri vašem študiju nasploh. Pazite, da ne boste koga pozabili. Utegnil vam bo zameriti. Temu se da izogniti tako, da celotno zahvalo izpustite.

Kazalo

Povzetek

Abstract

1	Uvod	1
1.1	Motivacija	1
1.2	Cilji	2
2	Pregled področja in sorodnih del	3
3	Metodologija in uporabljena orodja	5
4	Modeliranje nevronov in sinaps	7
5	Simulacija in vizualizacija impulznih nevronske mreže	9
6	Spodbujeno učenje na impulznih nevronske mreže	11
6.1	Sinaptična plastičnost odvisna od nagrajevanja in časovne razporeditve impulzov	11
6.2	Problem oddaljene nagrade	12
6.3	Pripisovanje odgovornosti	12
6.4	TD (angl. <i>Temporal Difference</i>) učenje	12
7	Nevronska vezja	15
8	Igranje igre Pong	17

9 Rezultati	19
10 Zaključek	21
Članki v revijah	23
Celotna literatura	25

Seznam uporabljenih kratic

kratica	angleško	slovensko
SNN	Spiking neural network	Impulzna nevronska mreža
R-STDP	Reward modulated spike timing dependent plasticity	Sinaptična plastičnost odvisna od nagrajevanja in časovne razporeditve impulzov
TD	Temporal difference	Temporalna razlika

Povzetek

Naslov: Spodbujevano učenje na impulznih nevronske mrežah

Avtor: Matjaž Pogačnik

V vzorcu je predstavljen postopek priprave diplomskega dela z uporabo okolja L^AT_EX. Vaš povzetek mora sicer vsebovati približno 100 besed, ta tukaj je odločno prekratek. Dober povzetek vključuje: (1) kratek opis obravnavanega problema, (2) kratek opis vašega pristopa za reševanje tega problema in (3) (najbolj uspešen) rezultat ali prispevek diplomske naloge.

Ključne besede: računalnik, računalnik, računalnik.

Abstract

Title: Reinforcement learning on spiking neural networks

Author: Matjaž Pogačnik

This sample document presents an approach to typesetting your BSc thesis using L^AT_EX. A proper abstract should contain around 100 words which makes this one way too short.

Keywords: computer, computer, computer.

Poglavje 1

Uvod

1.1 Motivacija

Impulzne nevronske mreže so v veliki večini implementacij poskus modeliranja bioloških značilnosti nevronov in sinaps v možganih. Kot izjemno močan računski stroj, so možgani navdih za mnoge moderne koncepte v umetni inteligenci. Najbolj očiten tak primer so nevronske mreže, vendar se po mehanizmi prisotnih v možganih lahko zgledujemo tudi pri metodah spodbujevanega učenja.

Delovanje možganov je kljub mnogim raziskavam še vedno precej slabo razumljeno, njihovo računalniško modeliranje pa je v času pisanja še precej mlado področje. Odkrivanje kakršnihkoli mehanizmov in vzorcev, ki se pojavijo med delovanjem in učenjem impulznih nevronske mreže ter uporaba teh pri modeliranju mehanizmov, za katere vemo, da so prisotni v možganih, predstavlja velik doprinos tako k področju računske nevroznanosti kot tudi psihoanalizi in drugim sorodnim področjem. V neposredni povezavi s psihoanalizo, raziskovanje impulznih nevronske mreže predstavlja raziskovanje temeljnih vprašanj o človeškem doživetju in delovanju možganov nasploh.

1.2 Cilji

V tej diplomski nalogi razvijemo rešitve kompleksne naloge s pomočjo impulznih nevronske mreže in spodbujevanega učenja. V prvi fazi so predstavljeni in ovrednoteni različni modeli bioloških značilnosti nevronov in sinaps 4. Na podlagi izbranih modelov potem izdelamo ogrodje za simulacijo in vizualizacijo delovanja impulznih nevronske mreže 5. V nadaljevanju so na le-teh uporabljeni algoritmi spodbujevanega učenja na impulznih nevronske mrežah kot so model sinaptične plastičnosti odvisne od nagrajevanja in časovne razporeditve impulzov (angl. R-STDP, primer Izhikevich EM [3]), s pomočjo česar je modeliran primer klasičnega pogojevanja [4] 6. V nadaljevanju klasično pogojevanje in učenje na podlagi nagrade uporabimo za rešitev kompleksnejše naloge - igranja igre Pong 8. Za ta namen modeliramo nevronska vezja in določene mehanizme iz človeškega dopaminskega sistema 7. Med rezultate vključimo primerjavo učenja na podlagi R-STDP, TD in naprednejšega modela Actor-Critic ter druga trenutno obstoječa orodja in implementacije spodbujevanega učenja na impulznih nevronske mrežah 6.4.

Poglavje 2

Pregled področja in sorodnih del

Na temo impulznih nevronske mreže v Sloveniji v času pisanja še ni bila napisana nobena diplomska ali magistrska naloga, doktorska dizertacija ali znanstveni članek, kar je dodatna motivacija za pisanje diplomske naloge na to temo. Impulzne nevronske mreže zaradi zahtevnosti učenja (v času pisanja) niso kaj dosti uporabljene, čedalje bolj uprabna metoda v umetni inteligenci pa je spodbujevano učenje, ki je tudi prevladujoča in biološko podprta metoda za učenje impulznih nevronske mreže.

Na temo spodbujevanega učenja je na voljo več slovenskih znanstvenih del. Pri spodbujevanem učenju predstavimo svoj sistem kot agenta, ki izvaja aktivnosti nad okoljem, ki mu kot odziv vrača nagrado in novo stanje okolja. Med drugim je uporabno v problemih kot so navigacija in reševanje problemov z roboti, na temo česar je bil objavljen članek pod avtorstvom prof. dr. Danijela Skočaja, rednega profesorja na FRI in dr. Mateja Dobrevskega [2]. Objavljenih je tudi več diplomskih in magistrskih nalog bolj simuliranih problemov, kot so uporaba spodbujevanega učenja za simulacijo psa ovčarja Štromajer T [7], reševanje problemov sorodnih problemu vozička s palico Svete A [6], igranje iger Šutar M [5] in uporaba TD [8] (angl. Temporal Difference) učenja v Monte Carlo preiskovanju dreves Deleva A [1]. V

vseh navedenih primerih se zgledujemo po raznolikem procesiranju podatkov iz zunanjega okolja, kjer je v robotiki in podatkih iz resničnega sveta prisoten tudi šum, ki je tako potrebna, kot tudi težavna komponenta pri učenju impulznih nevronske mreže.

Pri spodbujanem učenju, sploh v resničnem svetu, imajo impulzne nevronske mreže lahko določene prednosti. Impulzne nevronske mreže namreč naravno upoštevajo časovno komponento in procesirajo sekvenčne podatkovne tokove. Ker so dogodki v teh mrežah v osnovi samo propagiranje impulzov sosednjim nevronom v naslednjem časovnem intervalu, je računanje lahko učinkovito in preprosto. Zaradi tega se pojavljajo tudi trdo-ožičene implementacije impulznih nevronske mreže. V delu Wunderlich T, et al. [10] je raziskana uporaba TD učenja na trdo-ožičeni impulzni nevronske mreži, kjer je končna naloga igranje igre Pong. Tudi v tej diplomski nalogi bo končna naloga enaka, vendar bodo za to uporabljeni računalniški in ne trdo-ožičeni modeli ter naprednejši učni algoritmi osnovani na spodbujanem učenju.

V tej diplomski nalogi je poudarek na simulacijah in snovanju algoritmov za spodbujano učenje na impulznih nevronske mrežah ter modeliranju različnih bioloških procesov in možganskih nevronske vezij. Pri tem je dober zgled delo Izikhevich EM [3], ki poleg modela nevronov in sinaps vpeljuje še način pripisovanja odgovornosti sinapsam za določeno aktivnost nevronske mreže. V postopku nadgradnje algoritmov učenje poteka tudi na osnovi TD učenja in njegovi biološko bolj neposredni implementaciji Actor-Critic, Wiebke P, et al. [9]. V tem postopku implementiramo dejansko nevronske vezje odgovorno za nagrajevanje, kot je bilo to raziskano v človeških možganih.

Poglavje 3

Metodologija in uporabljena orodja

Simulacija, vizualizacija in evalvacija komponent je implementirana s pomočjo orodja Matlab. Rešitev kompleksnejšega problema, ki zahteva tudi implementacijo igre Pong je razvita v Pythonu. V sklopu te naloge uporaba impulznih nevronske mreže zunaj simuliranega okolja, na trdo-ožičenih nevronskih čipih ali na robotih ni pokrito, zato posebna oprema za ta namen ni bila uporabljena. Predvsem za namene primerjave obstoječih sistemov s sistemom iz diplomske naloge, so uporabljene tudi nekatere obstoječe knjižnice za simulacijo impulznih nevronske mreže, na primer Nest. Sistemi, razviti v diplomski nalogi so poleg medsebojne primerjave ovrednoteni tudi z drugimi trenutno obstoječimi orodji in implementacijami spodbujevanega učenja na impulznih nevronske mrežah. Rešitev je ovrednotena tudi po metrikah podobnih tem uporabljenim v diplomski nalogi Svete A [6].

Poglavje 4

Modeliranje nevronov in sinaps

1. Predstavitev bioloških značilnosti nevronov in sinaps, ki bodo modelirane v nadaljevanju.
 - **Trajanje: 1/2 tedna**
 - Predpogoji: /
2. Predstavitev različnih pristopov in pravil (Hebbian rule, Izhikhevich, Hodgkin-Huxley etc.)
 - **Trajanje: 1 teden**
 - Predpogoji: aktivnost 5.1

Poglavje 5

Simulacija in vizualizacija impulznih nevronske mreže

1. Vizualizacija mehanizmov znotraj impulznih nevronske mreže.
 - **Trajanje: 1 teden**
 - Predpogoji: aktivnosti poglavja 4
2. Izbira diagramskih tehnik za vizualizacijo aktivnosti in interakcij med nevroni.
 - **Trajanje: 1/2 teden**
 - Predpogoji: aktivnost 6:1
3. Izdelava ogrodja, ki je uporaben za poljubno nalogo in bo v nadaljevanju uporabljen pri igranju igre Pong.
 - **Trajanje: 2 tedna**
 - Predpogoji: aktivnost 6:1

Poglavje 6

Spodbujevano učenje na impulznih nevronskih mrežah

6.1 Sinaptična plastičnost odvisna od nagrajevanja in časovne razporeditve impulzov

1. Implementacija sinaptične plastičnosti odvisne od nagrajevanja in časovne razporeditve impulzov (angl. *Reward Modulated Spike Timing Dependent Plasticity - R-STDP*).

- Trajanje: 1 teden
- Predpogoji: aktivnosti poglavja 4

2. Rešitev preprostega problema s pomočjo spodbujevanega učenja na podlagi R-STDP.

- Trajanje: 3 tedne
- Predpogoji: aktivnost 7.1:1

6.2 Problem oddaljene nagrade

1. Predstavitev problema oddaljene nagrade (angl. *Distal reward problem*) in predlagana rešitev ter implementacija.
 - Trajanje: 1 teden
 - Predpogoji: aktivnosti poglavja 6.1

6.3 Pripisovanje odgovornosti

1. Predstavitev problema pripisovanja odgovornosti (angl. *Credit assignment*) in predlagana rešitev ter implementacija.
 - Trajanje: 1 teden
 - Predpogoji: aktivnosti poglavja 4

6.4 TD (angl. *Temporal Difference*) učenje

1. Implementacija osnovnega TD algoritma in primer klasičnega pogojevanja za preprost problem (npr. primer zvonec-hrana).
 - Trajanje: 1,5 tedna
 - Predpogoji: aktivnosti poglavja 4
2. Uporaba pripisovanja odgovornosti na implementiranih TD algoritmih
 - Trajanje: 3 tedne
 - Predpogoji: aktivnosti poglavja 6.3, 6.2 in aktivnost 7.4:1

6.4.1 Model Actor-Critic

1. Implementacija algoritma Actor-Critic.
 - Trajanje: 1 teden

-
- Predpogoji: aktivnosti 7.4:1,2

Poglavje 7

Nevronska vezja

1. Modeliranje človeških dopaminskih mehanizmov, dopaminskih receptorjev in odzivov na uspešno izvedene akcije.
 - **Trajanje: 2 tedna**
 - Predpogoji: aktivnosti poglavja 4 in 7.1-7.4.
2. Nov pristop k implementaciji nevronske vezji.
 - **Trajanje: 4 tedni**
 - Predpogoji: aktivnost 7.5:1.

Poglavje 8

Igranje igre Pong

1. Predstavitev naloge in pretvorba problema na problem rešljiv z spodbujanjem učenjem na impulznih nevronske mrežah. Definicije nagrad, vhodov in izhodov iz sistema ter izbira algoritma.

- **Trajanje: 2 tedna**
- Predpogoji: aktivnosti poglavja 6, 5

Poglavje 9

Rezultati

1. Primerjava različnih metod učenja: R-STDP, TD in Actor-Critic.
 - **Trajanje: 1/2 tedna**
 - Predpogoji: aktivnosti poglavja 6
2. Analiza obstoječih orodij in implementacij spodbujevanega učenja na impulznih nevronskih mrežah ter njihova primerjava z lastno rešitvijo
 - **Trajanje: 1 tedna**
 - Predpogoji: aktivnosti poglavja 8

Poglavje 10

Zaključek

1. Ovrednotenje impulznih nevronske mreže in spodbujevanja učenja na le-teh za rešitev različnih "real world" problemov.

- **Trajanje: 1/2 tedna**
- Predpogoji: aktivnosti poglavja 9

2. Komentar in povezava rezultatov in metod z dosedanjim razumevanjem človeške zavesti in dopaminskega sistema.

- **Trajanje: 1 teden**
- Predpogoji: aktivnosti poglavja 9

Članki v revijah

- [2] Dobrevski M, Skočaj D. “Deep reinforcement learning for map-less goal-driven robot navigation”. V: *International Journal of Advanced Robotic Systems*. 2021 18.1 (2021). DOI: 10.1177/1729881421992621.
- [3] Izhikevich, E. M. “Solving the distal reward problem through linkage of STDP and dopamine signaling”. V: *Cerebral cortex (New York, N.Y. : 1991)* 17.10 (2007). DOI: 10.1093/cercor/bhl152.
- [9] Wiebke P, et al. “An Imperfect Dopaminergic Error Signal Can Drive Temporal-Difference Learning”. V: *PLoS computational biology* 7.5 (2011). DOI: 10.1371/journal.pcbi.1001133.
- [10] Wunderlich T, et al. “Demonstrating Advantages of Neuromorphic Computation: A Pilot Study”. V: *Frontiers in neuroscience* 13.260 (2019). DOI: 10.3389/fnins.2019.00260.

Celotna literatura

- [1] Deleva A. “TD learning in Monte Carlo tree search : masters thesis”. Magistrska naloga. Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, 2015.
- [2] Dobrevski M, Skočaj D. “Deep reinforcement learning for map-less goal-driven robot navigation”. V: *International Journal of Advanced Robotic Systems*. 2021 18.1 (2021). DOI: 10.1177/1729881421992621.
- [3] Izhikevich, E. M. “Solving the distal reward problem through linkage of STDP and dopamine signaling”. V: *Cerebral cortex (New York, N. Y. : 1991)* 17.10 (2007). DOI: 10.1093/cercor/bhl152.
- [4] *Klasično pogojevanje*. URL: https://sl.wikipedia.org/wiki/Klasi%C4%8Dno_pogojevanje (pridobljeno 20. 3. 2025).
- [5] Šutar M. “Uporaba predvidevanja akcij nasprotnika pri učenju inteligentnega agenta”. Diplomaska naloga. Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, 2023.
- [6] Svete A. “Posplošitev problema vozička s palico na zahtevnejše domene”. Diplomaska naloga. Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, 2020.
- [7] Štromajer T. “Using machine learning to train a shepherd dog”. Magistrska naloga. Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, 2022.
- [8] *Temporal difference learning*. URL: https://en.wikipedia.org/wiki/Temporal_difference_learning (pridobljeno 20. 3. 2025).

- [9] Wiebke P, et al. “An Imperfect Dopaminergic Error Signal Can Drive Temporal-Difference Learning”. V: *PLoS computational biology* 7.5 (2011). DOI: 10.1371/journal.pcbi.1001133.
- [10] Wunderlich T, et al. “Demonstrating Advantages of Neuromorphic Computation: A Pilot Study”. V: *Frontiers in neuroscience* 13.260 (2019). DOI: 10.3389/fnins.2019.00260.