

FighterZero

matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek



FighterZero

matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek

FighterZero: Učenje igranja računalniških iger z globoko obdelavo slik in drevesnim preiskovanjem Monte Carlo

Matej Vitek in Peter Peer
{matej.vitek, peter.peer}@fri.uni-lj.si

Fakulteta za računalništvo in informatiko
Univerza v Ljubljani

FighterZero

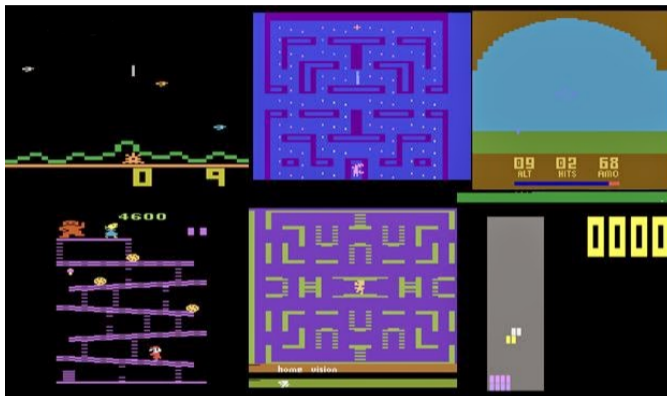
matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek



Slika: Igre za klasično konzolo Atari 2600.

Slika predelana iz: <https://www.downarea51.com/>

FighterZero

matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek

- ▶ Ekipa Google DeepMind.
- ▶ Globoka Q-mreža (Deep Q-Network, DQN).
- ▶ Spodbujevalno učenje na podlagi slikovne informacije.
- ▶ Leta 2014 nato ekipa z michiganske univerze DQN nadgradi z drevesnim preiskovanjem Monte Carlo (MCTS).

FighterZero

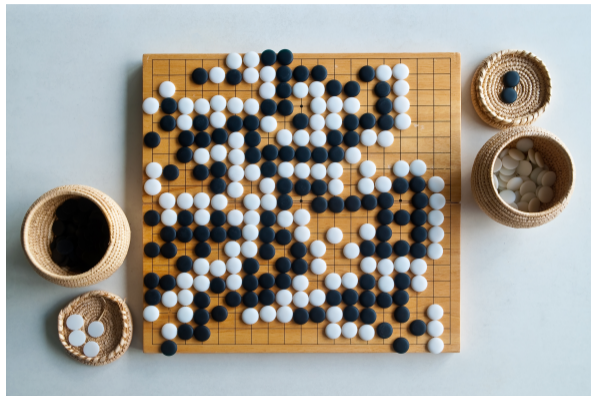
matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek



Slika: Namizna igra Go.

Slika je last Getty Images/iStockphoto – Saran Poroong

FighterZero

matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek



Slika: Namizni igri šah in šogi.

Sliki sta last Pixabaya (<https://pixabay.com/>) in Oliverja Orschiedta (CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=11038951>)

FighterZero

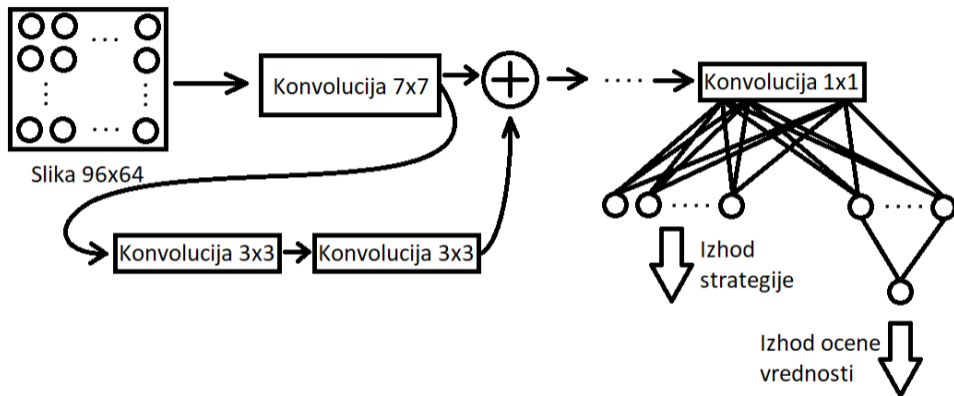
matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek



Slika: Uporabljena konvolucijska nevronska mreža preostankov.

FighterZero

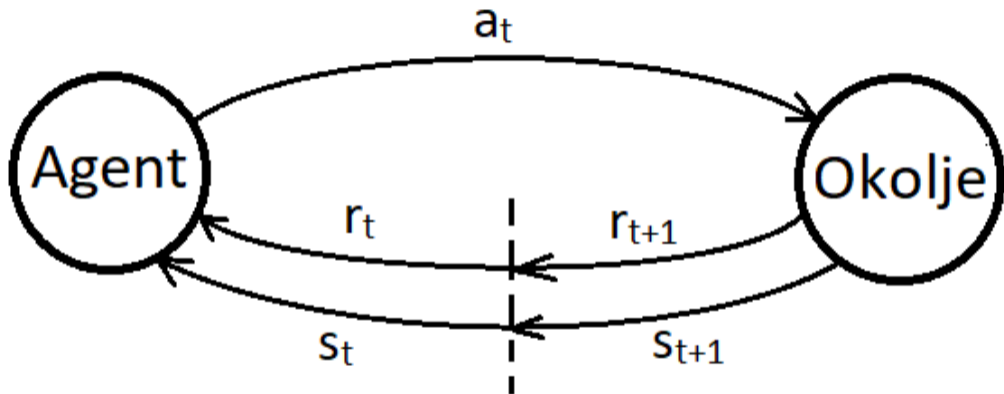
matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek



Slika: Skica poteka spodbujevalnega učenja.

FighterZero

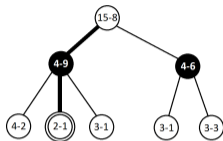
matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

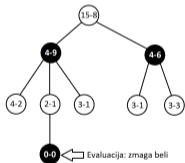
FighterZero

Rezultati

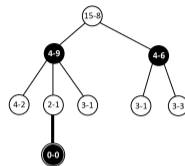
Zaključek



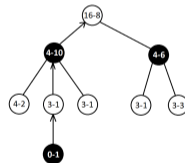
(a) Izbira vozlišča



(c) Evaluacija



(b) Razširitev



(d) Vzratno širjenje izida

Slika: Ena ponovitev drevesnega preiskovanja Monte Carlo (MCTS).

FighterZero

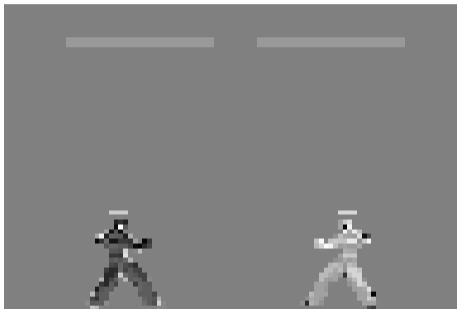
matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek



(a) Začetna slika



(b) Simulirana slika

Slika: Delovanje našega simulatorja slikovne informacije prihodnjega stanja.

FighterZero

matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek

Algoritem 1 FighterZero

- 1: Naključno inicializiramo nevronske mreže n
 - 2: **for** $i = 1$ to $BIGNUMBER$ **do**
 - 3: **while** igra ni končana **do**
 - 4: Izberemo potezo m z uporabo MCTS, v koraku evaluacije uporabimo n
 - 5: Shranimo trenutno stanje in potezo m
 - 6: Izvedemo potezo m
 - 7: **end while**
 - 8: Shranjenim primerom dodelimo nagrade glede na končni izid
 - 9: Izvedemo učenje mreže n s tako dobljenimi primeri
 - 10: **end for**
 - 11: **return** n
-

FighterZero

matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek

KickAI	MctsAi	GigaThunder
8,5–0,5	0–9	0–9

Tabela: Rezultati našega agenta proti testnim agentom.

- ▶ KickAI je preprost agent, ki stalno izbira samo akcijo brcanja.
- ▶ MctsAi je agent, ki je v tekmovanju FTGAIC uporabljen kot primerjalni agent. Uporablja klasično obliko MCTS.
- ▶ GigaThunder je zmagovalec FTGAIC iz leta 2017. Prav tako si pomaga z MCTS, poleg tega pa ima še mnogo drugih pravil oblike če-potem.

FighterZero

matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek

- ▶ Kratek čas za odločanje: AlphaZero 1 minuto, naš agent 16 ms (3600× manj).
- ▶ V kodi precej klicev metod iz ogrodja FightingICE – precej počasni.
- ▶ Tako le 1–2 ponovitvi algoritma MCTS na potezo (AlphaZero nekaj milijonov na potezo).
- ▶ Zato smo:
 - ▶ postopek izbiranja podaljšali na 5 sličic igre → čas odločanja 80 ms;
 - ▶ minimizirali število klicev metod iz ogrodja;
 - ▶ zmanjšali uporabljeno mrežo;
 - ▶ napovedi mreže poganjali na CPU namesto na GPU;
 - ▶ uporabili bolj nizkonivojsko metodo za napovedovanje;
 - ▶ poskusili upočasniti igro; neuspešno, saj ogrodje take upočasnitve ni dovoljevalo.
- ▶ Na koncu 20–30 ponovitev MCTS na potezo – žal še vedno premalo.

FighterZero

matej.vitek@
fri.uni-lj.si

Pregled
zgodovine

FighterZero

Rezultati

Zaključek

- ▶ Agent je deloval slabše kot pričakovano – premagal je namreč le najosnovnejšega nasprotnika.
- ▶ Glavni problem je bil kratek čas za odločanje.
- ▶ Celoten algoritem smo pospešili za red $10\times$, kar pa še vedno ni bilo dovolj za dobro delovanje.
- ▶ Z boljšo izbiro ogrodja in programskega jezika bi ga lahko verjetno pospešili še za red $10\times$.
- ▶ Da pa bi agent (tako kot AlphaZero) dosegel nadčloveški nivo igranja, bi najbrž potrebovali Googlov superračunalnik.