Inteligentné agenty pre stratégiu reálneho času

ŠTEFAN KRIŠTOFÍK

stefan.kristofik@savba.sk

SLOVENSKÁ AKADÉMIA VIED
ÚSTAV INFORMATIKY
ODDELENIE NÁVRHU A DIAGNOSTIKY
ČÍSLICOVÝCH SYSTÉMOV

3.6. 2020

ŠACH

- Hráči sa striedajú (turn-based)
- Ťah = posun 1 figúrky
  - Je to instantná akcia
- Deterministickosť
- Odhad stavového priestoru ≈ 10^{50}
- Hracia plocha odhalená
- Cieľ: šach mat
- Dosiahnutie cieľa
  - Predvídať čo najviac vopred
  - „Preľstiť“ súpera

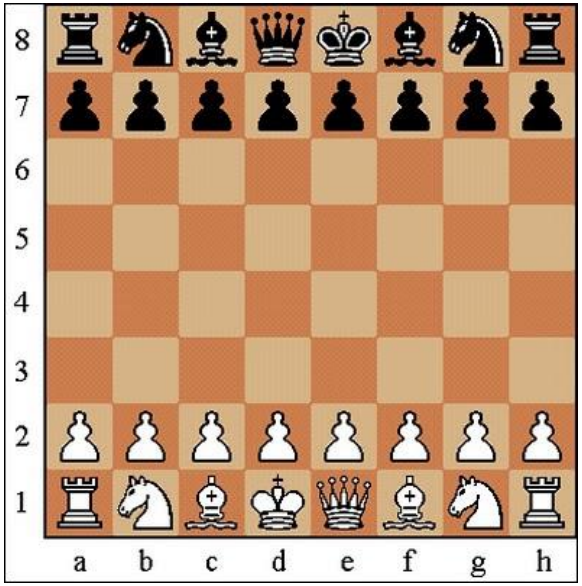
RTS

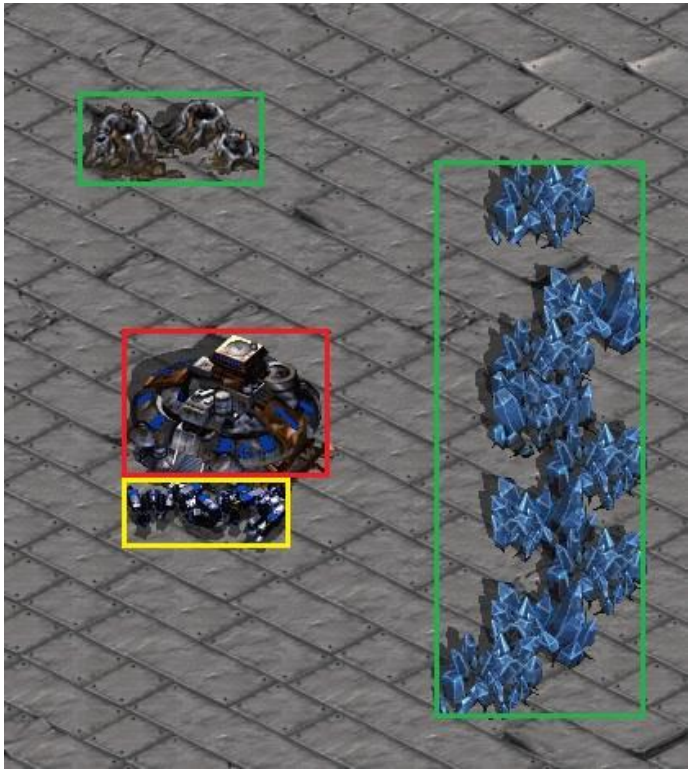
- Hráči idú naraz (real-time)
- Ťah = posun jednotky, útok na cieľ, výroba jednotky, postaviť budovu, ...
  - Nie sú to vždy instantné akcie, trvajú určitý čas
- Nedeterministickosť
- Spodný odhad ≈ 10^{1685}
- Odhalené iba malé časti hracej plochy okolo vlastných jednotiek a budov
- Cieľ: zničiť všetky budovy protivníka
- Dosiahnutie cieľa – princíp 4X
  - eXplore – prieskum
  - eXpand – rozširovanie územia a vplyvu
  - eXploit – využívanie surovín
  - eXterminate – zničenie protivníka

SPOLOČNÉ ČRTY

- 1v1
- Súťaže, turnaje, ratingy
- Rôzne úrovne hráčov – od amatérskych až po expertov
- Známe, dobre zdokumentované stratégie
- Neexistencia jednej univerzálnej, neporaziteľnej stratégie
- Možnosť učiť sa z minulosti od tých najlepších a tak sa zlepšovať
- Nutná nielen teoretická znalosť pravidiel a stratégií, treba aj veľa trénovať
- Možnosť skoršieho ukončenia hry

ZAČIATOČNÁ POZÍCIA

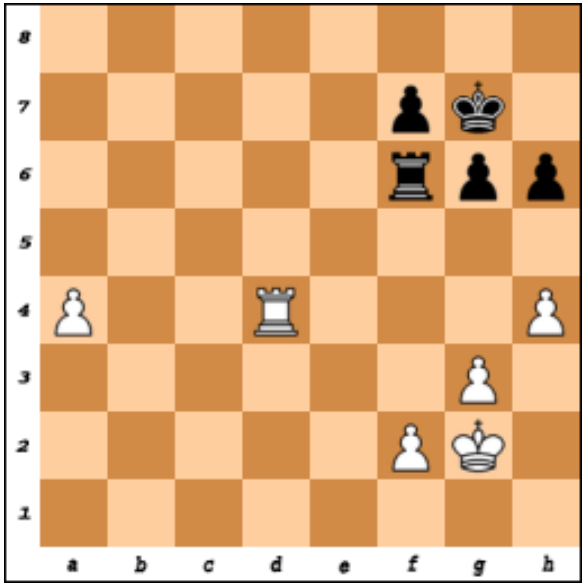




- Fixný počet a typ figúrok

- Fixný počet budov a jednotiek
- Suroviny

PRIEBEH HRY



- Počas hry nemôže pribudnúť nová figúrka
- Ale môže sa zmeniť typ figúrky

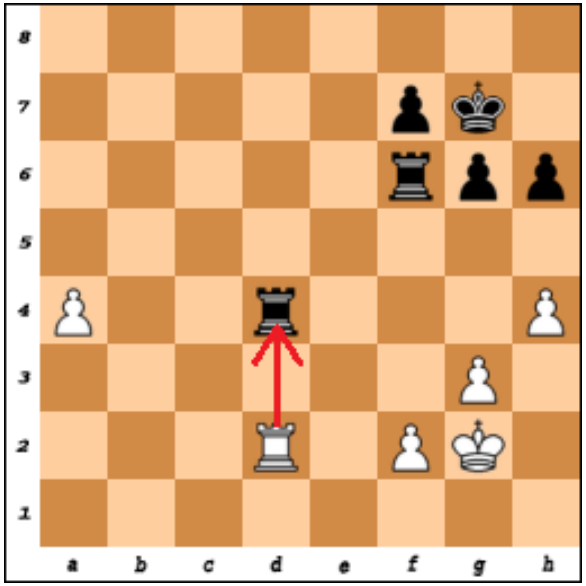
- Počas hry môžu pribúdať nové jednotky a budovy
- A tiež sa môžu meniť ich typy
- Priebeh = sekvencia snímok („frames“)

PRIESKUM

Doposiaľ nenavštívené lokality

Už navštívené lokality, ale momentálne bez našej prezencie

SÚBOJE



- Jedna figúrka zaútočí
- Jedna figúrka ubudne

- Množstvo jednotiek útočí v tom istom čase
- Množstvo objektov ubudne

ZÁZNAM HRY

1. e4 e5 2. Nf3 Nc6 3. Bc4 Bc5 4. b4 Bb4 5. c3 Ba5 6. d4 ed4 7. O-O d3 8. Qb3 Qf6 9. e5 Qg6 10. Re1 Nge7 11. Ba3 b5 12. Qb5 Rb8 13. Qa4 Bb6 14. Nbd2 Bb7 15. Ne4 Qf5 16. Bd3 Qh5 17. Nf6 gf6 18. ef6 Rg8 19. Rad1 Qf3 20. Re7 Ne7 21. Qd7 Kd7 22. Bf5 Ke8 23. Bd7 Kf8 24. Be7#

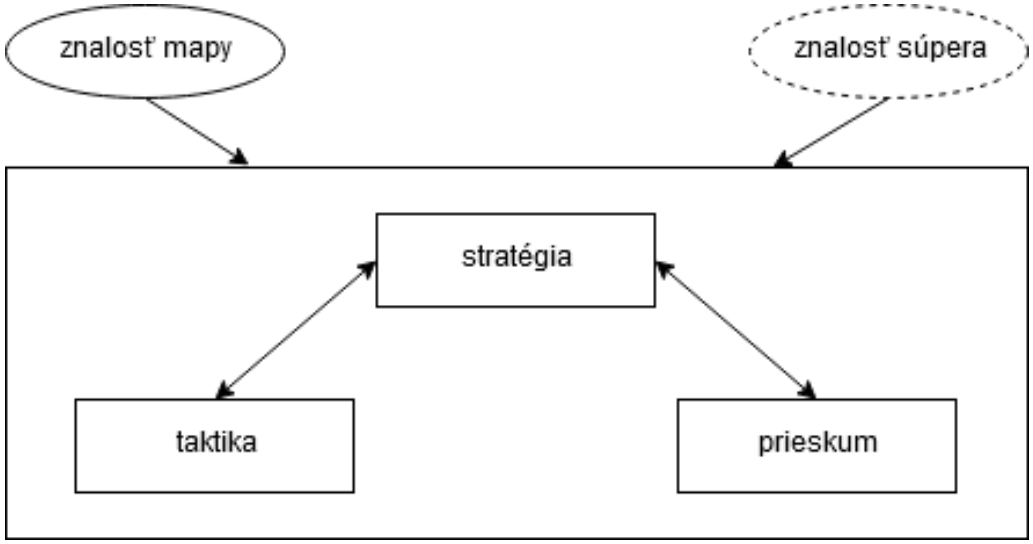
- Súbor so špecifickým formátom
- Zoznam akcií hráčov
  - Timestamp – kedy
  - Označenie – kto
  - Príkazy – čo
- Využitie záznamu
  - Spätné prehratie video záznamu
  - Spracovanie a analýza údajov

VÝZVY PRE AI V RTS

- Obrovský stavový priestor – klasické metódy nemôžu fungovať
- Vysporiadanie sa s neúplnými informáciami
  - Nutné robiť neinformované rozhodnutia
  - Kompenzácia – prieskum, skúsenosti
- Dlhodobé plánovanie (makromanažment)
- Krátkodobé plánovanie (mikromanažment)
- Učenie sa
  - Pred, počas, medzi zápasmi
  - Ako využiť rozsiahlu znalostnú bázu?
  - Ako využiť znalosť konkrétnych protivníkov?
- Spatial reasoning – „cit“ pre priestor
- Temporal reasoning – „cit“ pre čas

TYPICKÁ ARCHITEKTÚRA

- Dekompozícia na menšie úlohy – moduly
- Stratégia – globálna, vyššia úroveň, dlhodobé ciele
- Taktika – lokálna, nižšia úroveň, krátkodobé ciele, pre naplnenie dlhodobých cieľov
- Prieskum
- Analýza mapy



https://certicky.github.io/files/publications/Starcraft-AI-ToG-2018.pdf

SÚČASNÉ TECHNIKY (2018 UPDATE)

Rules – rule-based system

ML – machine learning

HS – heuristic search

IO – využitie I/O súborov pre učenie

SIM – simulácie počas hry

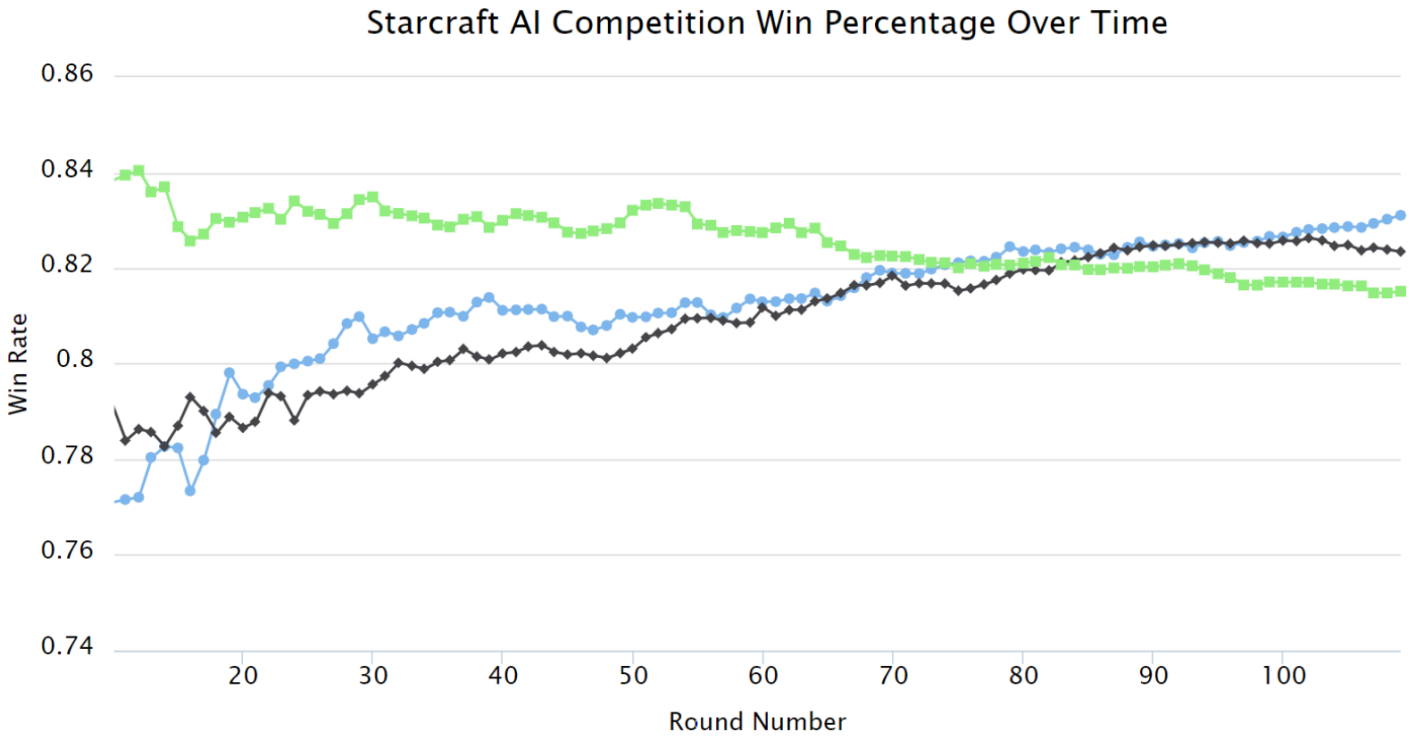
Agent	Created	Rules	ML	HS	IO	SIM
CherryPi	2017					
Cpac	2017					
ForceBot	2017					
Iron	2016					
KillAll	2017					
Krasi0bot	2010					
LetaBot	2014					
McRave	2017					
MegaBot	2016					
PurpleWave	2017					
StarcraftGP	2015					
Steamhammer	2016					
Tscmoo	2015					
UAlbertaBot	2010					
ZZZKbot	2015					

https://certicky.github.io/files/publications/Starcraft-AI-ToG-2018.pdf

POROVNANIE TECHNÍK

Turnaj pri konferencii AIIDE 2017
3 najúspešnejšie agenty

Agent	Created	Rules	ML	HS	IO	SIM
Iron	2016					
PurpleWave	2017					
ZZZKbot	2015					



NAŠE AGENTY

Rules – rule-based system

ML – machine learning

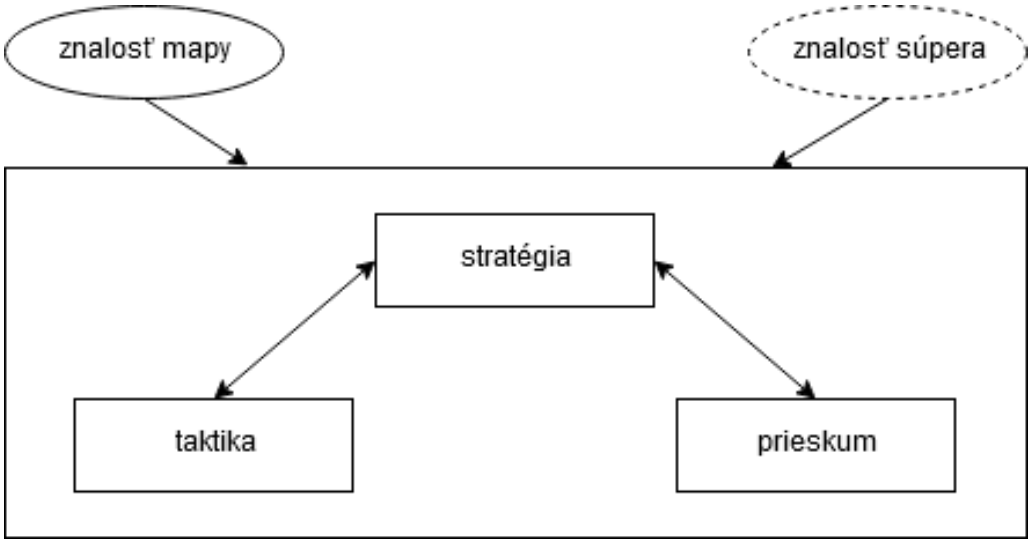
HS – heuristic search

IO – využ. súborov pre učenie

SIM – simulácie

Agent	Created	Rules	ML	HS	IO	SIM
KasoBot	2020					
DTDBot	2020					
NuiBot	2020					

- Zamerané na
  - Stratégiu
  - Prieskum
- Zaostávajú v
  - Taktike



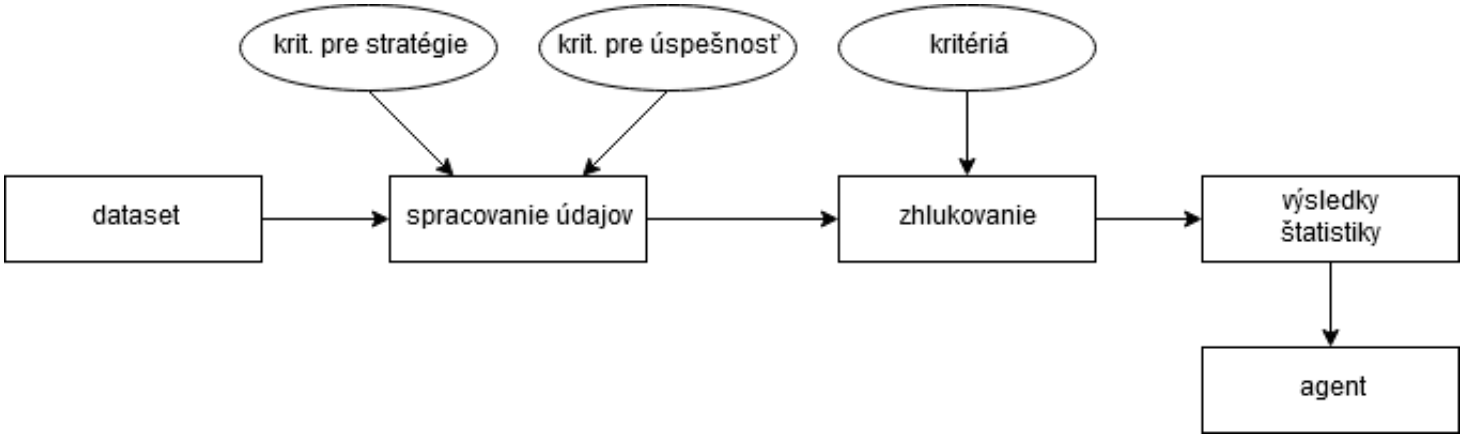
AGENT 1 – KASO

- Hlavné myšlienky
  - Skúsme využiť rozsiahly dataset záznamov hier
  - Automaticky identifikovať a „odkukať“ najčastejšie používané stratégie
  - Štatisticky vyhodnotiť vzájomnú úspešnosť stratégií
- Cieľ
  - Agent bude schopný emulovať a nasadiť štatisticky výhodné stratégie
- Téza
  - Agent s uvedenými vlastnosťami by mohol byť konkurencieschopný
- Použité metódy
  - Unsupervised learning (učenie bez učiteľa)
  - Clustering

DATASET

- STARDATA
  - Tím Facebook, 2017
  - arXiv:1708.02139v1
- Spolu 65646 záznamov
  - Train 90%, validation 5%, test 5%
- Použitých 16713

AGENT 1 – KASO



- Definovanie stratégie v kontexte dostupných dát
  - = Poradie postavených budov
  - = Početnosť postavených budov
  - = Početnosť vyrobených jednotiek
- Kritérium pre hodnotenie úspešnosti
  - = Identifikovanie víťaza
- Výber kritérií pre zhlukovanie
  - = Konkrétne objekty
- Algoritmus: K-Means
- Implementácia stratégií

VÝSLEDKY

- Stratégie

cluster	counts	Academy	Armory	Command Center	Control Tower	Engineering Bay	Factory	Machine Shop	Starport	Marine	Vulture	Goliath	Siege Tank	Wraith	Medic	Firebat
0	1126	6.50	3.73	2.98	8.48	5.25	1.38	3.14	7.77	3.30	3.05	1.75	3.63	7.93	7.48	7.92
1	2660	2.14	8.92	1.17	7.07	2.89	3.88	5.39	5.73	1.04	7.98	7.96	2.64	7.59	2.70	5.24
2	1063	1.78	9.00	3.27	9.00	4.02	9.00	9.00	9.00	1.16	8.00	8.00	8.00	8.00	2.75	4.36
3	792	8.63	8.99	7.15	9.00	8.65	8.27	8.93	9.00	1.38	7.69	8.00	7.98	8.00	8.00	8.00
4	683	1.48	8.89	7.70	6.75	3.79	2.08	5.03	5.39	1.06	7.16	7.92	4.20	7.70	2.70	4.42
5	423	5.44	8.73	7.15	5.19	6.42	1.08	7.52	2.20	1.63	3.59	7.92	7.98	2.76	4.94	7.52
6	1134	2.44	8.60	1.16	7.00	2.99	3.70	5.49	5.66	1.16	2.71	7.68	4.18	7.59	3.32	5.23
7	1339	2.20	8.87	1.25	7.66	2.85	3.85	8.00	6.37	1.03	7.91	7.98	8.00	7.55	2.21	5.15
8	583	5.16	8.30	5.46	4.77	5.79	1.08	4.49	2.38	1.27	4.64	7.83	2.79	4.85	3.36	7.13
9	428	7.90	4.73	5.39	5.12	6.70	1.07	3.50	2.61	3.62	2.96	1.58	3.87	5.78	7.76	7.98

- Úspešnosť stratégií

	0	1	2	3	4	5	6	7	8	9
0	72.22	67.50	50.00	65.43	48.14	50.00	31.71	37.04	51.04	65.00
1	63.65	79.17	100.00	67.35	70.27	55.84	75.00	28.06	71.43	100.00
2	14.29	20.00	74.07	25.00	37.74	7.81	29.72	0.00	63.04	64.14
3	50.00	0.00	47.15	16.67	50.00	0.00	10.53	50.00	0.00	28.04
4	47.45	61.11	82.35	50.00	46.88	34.02	46.03	13.33	82.76	79.31
5	46.15	54.43	81.48	45.76	31.58	28.57	25.00	0.00	62.92	78.26
6	60.81	44.44	100.00	60.00	50.00	65.00	40.00	21.71	58.33	50.00
7	64.71	75.65	75.00	32.09	62.50	41.70	61.82	44.05	75.00	81.08
8	59.76	56.25	100.00	62.35	55.36	50.63	50.00	25.93	57.97	60.00
9	52.63	38.24	100.00	48.28	48.95	60.00	26.92	25.00	58.00	100.00

VÝSLEDKY

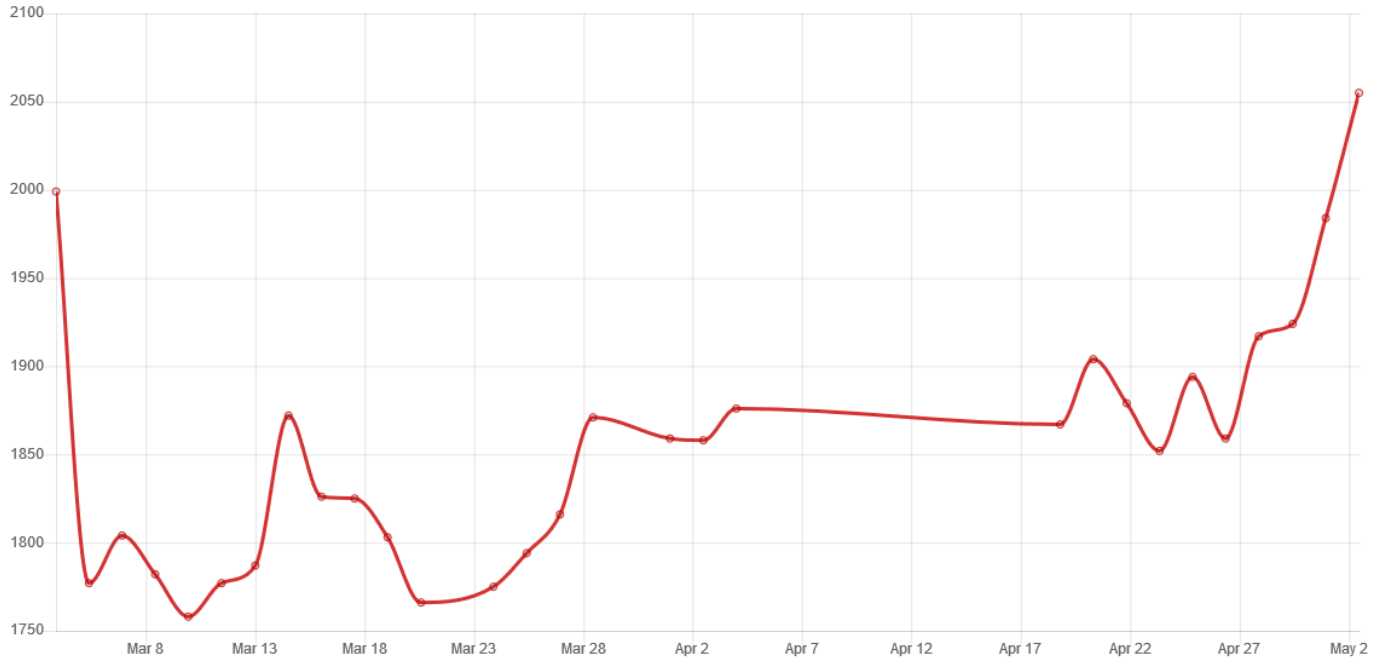
- SSCAIT
  - Náhodný súper – študent/neštudent, akýkoľvek rating
  - 89W – 293L
  - ELO rating 2004
  - Známka „D-“

- Offline
  - Iba proti 6 vybratým študentským riešeniam
  - 44W – 16L

VÝSLEDKY

- BASIL Ladder
  - Riešenie momentálne vyradené, nepostačujúca úspešnosť – nehľadí sa len na ELO, aktívnych len top 58
  - Globálne umiestnenie: pozícia 94 zo 131
  - Max. dosiahnutý ELO: 2067
  - Možné opätovné zapojenie v prípade aktualizácie agenta
  - Technické problémy

Vývoj ELO



AGENT 1 – ZHODNOTENIE

- Vlastný, nový spôsob spracovania voľne dostupného datasetu
- Získané veľmi zaujímavé štatistiky a výsledky
- Doposiaľ podobná, takáto podrobná analýza tohto datasetu nie je k dispozícii

- Agent sa udržiava približne na východiskovom ELO ratingu
- Dobre konkuruje iným študentským riešeniam
- Zaostáva za najlepšími neštudentskými
- Potrebné omnoho viac testovať, testovalo sa len 2 mesiace

- Veľmi sľubný prístup, dobrý základ pre ďalšiu prácu

Agent	Created	Rules	ML	HS	IO	SIM
KasoBot	2020					

AGENT 2 – DTD

- Hlavné myšlienky
  - Skúsme využiť rozsiahly dataset záznamov hier
  - Skúsme naučiť neurónovú sieť dávať odporúčania pre nasledujúce kroky agenta
  - Rozšírenie existujúceho agenta o modul strojového učenia
- Cieľ
  - Agent bude prijímať rozhodnutia zo siete a bude schopný podľa nich upraviť stratégiu, nie však taktiku
- Téza
  - Agent s uvedenými vlastnosťami by mohol byť konkurencieschopný
- Použité metódy
  - Supervised learning (učenie s učiteľom)
  - LSTM

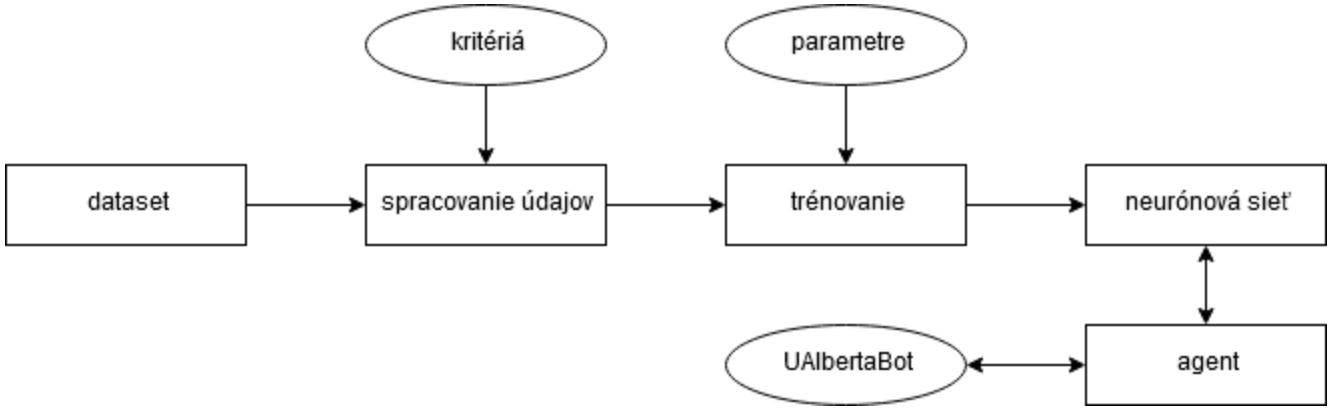
Agent	Created	Rules	ML	HS	IO	SIM
DTDBot	2020					

DATASET

- STARDATA
  - Tím Facebook, 2017
  - arXiv:1708.02139v1
- Spolu 65646 záznamov
  - Train 90%, validation 5%, test 5%
- Použitých 15920

KRITÉRIÁ PRE SPRACOVANIE

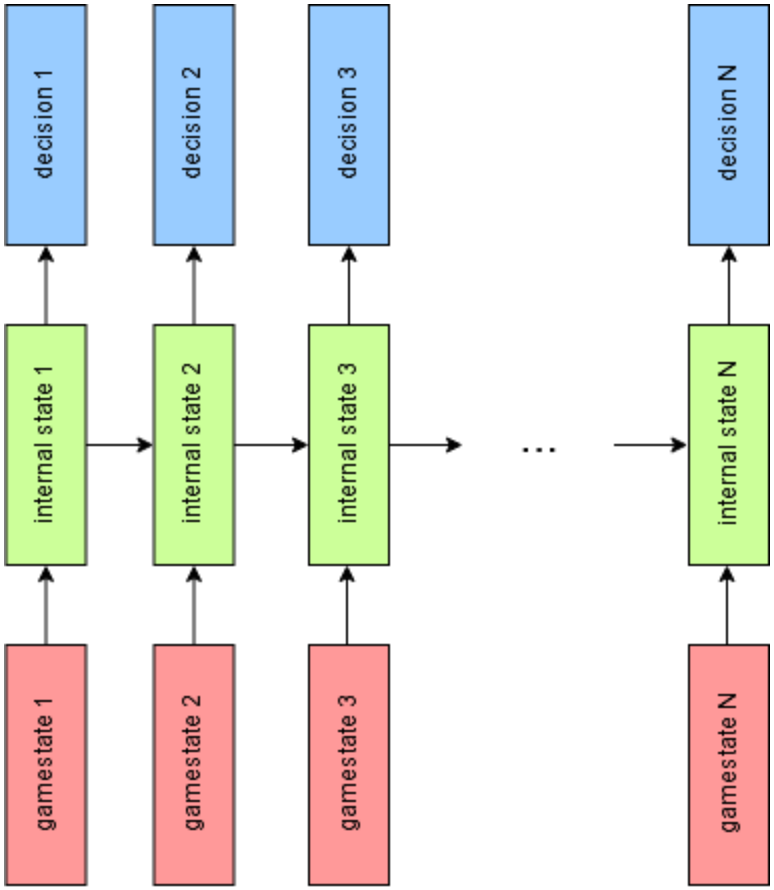
- Príprava dát do potrebného formátu pre trénovanie
- Vstup: aktuálny stav hry
  - Reprezentovaný ako 1-D vektor premenných reprezentujúcich stav hry
  - Mnoho údajov o oboch protivníkoch
  - Frekvencia vzorkovania: nepravidelne, vždy keď je vykonaná akcia „sledovaného“ hráča
- Label: samotná vykonaná akcia za daného stavu (44 možností)



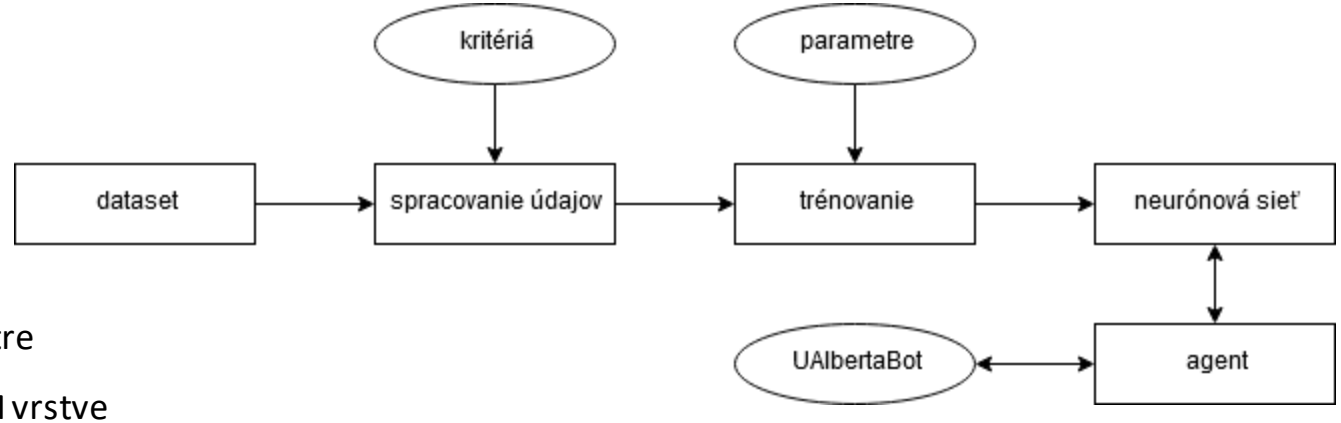
MODEL

- 1 LSTM vrstva
  - Many-to-many
- 1 plne prepojená vrstva
  - Aktivačná funkcia softmax

- Vstup
  - Vektor stavu hry v časových okamihoch
- Výstup
  - Predikcia (odporúčanie) ďalšej strategickej akcie agenta
  - Akcia s najvyššou váhou
  - Súčet váh = 1



MODEL



- Ladené hyperparametre
  - Počet „units“ v LSTM vrstve
  - Dropout
- Fixné hyperparametre
  - Stratová funkcia: categorical crossentropy
  - Optimizer: Adam
  - Metrika: categorical accuracy
- Tréning
  - Tuning na malej podmnožine datasetu (500 train, 100 test)
  - Najlepšia top-1 presnosť: 38,17 % pre 256 units a dropout 0.2
  - Tieto hodnoty následne použité pre tréning na celom datasete, top-1 presnosť: 30,75 %
- Inferencia: Na základe stavu hry dá odporúčanie pre ďalšiu akciu
- Modul pridaný do agenta UAlbertaBot

VÝSLEDKY

- Spracovaná veľká podmnožina záznamov z pôvodného datasetu
- K dispozícii je veľmi detailná analýza jednotlivých zápasov, vývoja hry
- Možné ďalšie využitie

VÝSLEDKY

- SSCAIT
  - Technické problémy s modulom pre strojové učenie

- Offline
  - Iba proti 9 vybratým študentským riešeniam
  - 13W – 32L

AGENT 2 – ZHODNOTENIE

- Vlastný, nový spôsob spracovania voľne dostupného datasetu
- Doposiaľ podobná, takáto podrobná analýza tohto datasetu nie je k dispozícii

- Získané zaujímavé poznatky
- Zrejme zvolený príliš veľký počet prvkov vektora stavu hry
- Odporúčania často nie sú „logické“
- Agent potrebuje dodatočne korigovať rozhodnutia zo siete
- Potrebné omnoho viac testovať, testovalo sa len približne 1 mesiac

- Sľubný prístup, dobrý základ pre ďalšiu prácu

Agent	Created	Rules	ML	HS	IO	SIM
DTDBot	2020					

Agent	Created	Rules	ML	HS	IO	SIM
NuiBot	2020					

AGENT 3 – NUI

- Hlavné myšlienky
  - Skúsme sa pozrieť na súčasné najúspešnejšie agenty a učiť sa od nich
  - Identifikovať ich najčastejšie stratégie
  - Pripraviť kontra-stratégie pomocou zostavenia sady pravidiel
  - Postupne si vytvárať a aktualizovať modely súperov, proti ktorým sa postavíme
- Cieľ
  - Agent bude schopný identifikovať model súpera a vybrať vhodnú kontra-stratégiu
- Téza
  - Agent s uvedenými vlastnosťami by mohol byť konkurencieschopný
- Použité metódy
  - Reinforcement learning (učenie s odmenou)
  - Manuálne štúdium video záznamov

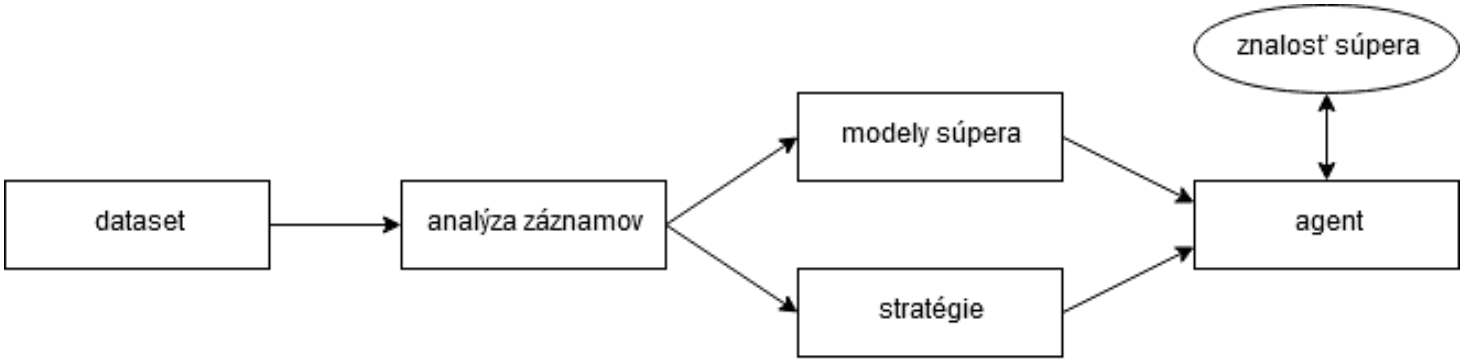
AGENT 3 – NUI

- Analýza agentov
  - Top 16 súťažiacich zo súťaže SSCAIT 2018/2019
  - Celkovo 112 vzájomných zápasov

- Dôraz kladený na prieskum

VÝSLEDKY

- Definovaných 12 rôznych modelov súpera
- História a štatistiky o každom súperovi, aktualizované po každom zápase
  - Odmena: pri výhre proti konkrétnemu súperovi je odmenená použitá stratégia
- 12 vlastných stratégií
  - Najviac odmenené stratégie sú vyberané častejšie



VÝSLEDKY

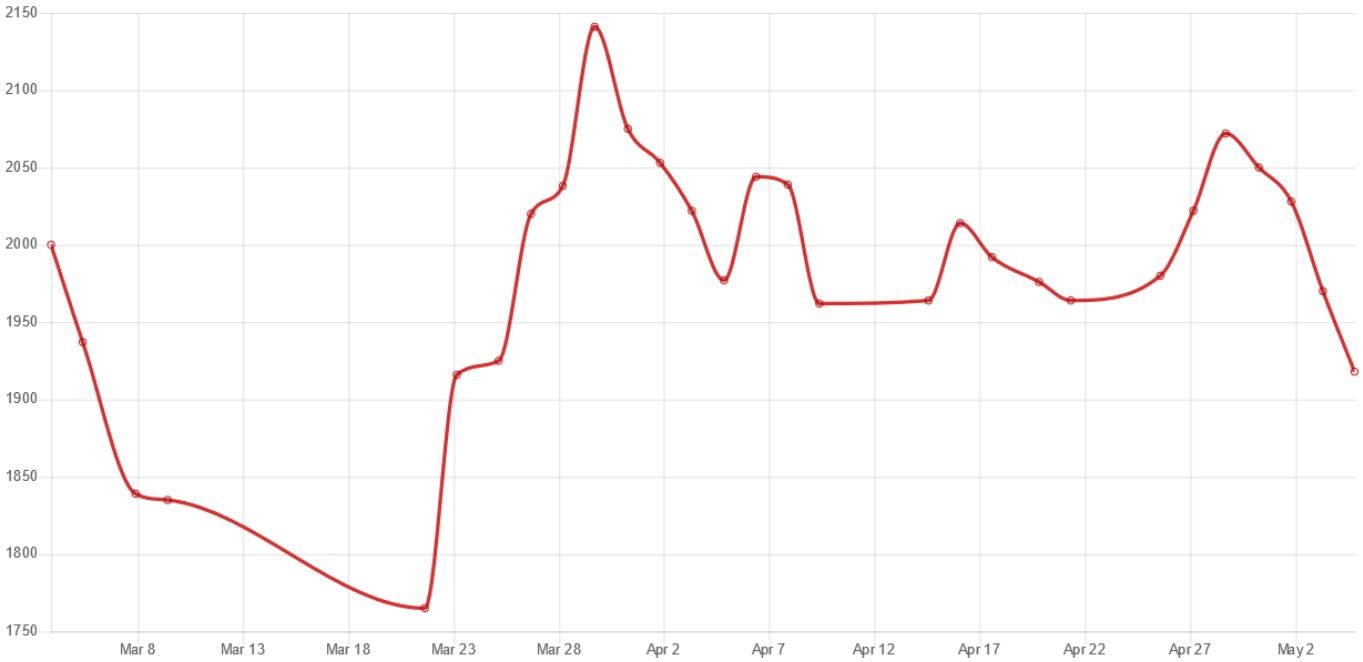
- SSCAIT
  - Náhodný súper – študent/neštudent, akýkoľvek rating
  - 101W – 250L
  - ELO rating 1915
  - Známka „E“

- Offline
  - Proti 7 rule-based agentom s podobným ELO
  - 120W – 30L

VÝSLEDKY

- BASIL Ladder
  - Riešenie momentálne vyradené, nepostačujúca úspešnosť – nehľadí sa len na ELO, aktívnych len top 58
  - Globálne umiestnenie: pozícia 103 zo 131
  - Max. dosiahnutý ELO: 2169
  - Možné opätovné zapojenie v prípade aktualizácie agenta
  - Technické problémy

Vývoj ELO



AGENT 3 – ZHODNOTENIE

- Získané zaujímavé výsledky a poznatky
- Možné vytvoriť konkurencieschopný agent starostlivou analýzou súčasných agentov
- Rule-based agenty sú stále zaujímavé a môžu predstavovať výzvu aj pre kvalitnejšie agenty

- Agent sa udržiava približne na východiskovom ELO ratingu
- Dobre konkuruje iným rule-based
- Zaostáva za najlepšími
- Potrebné omnoho viac testovať, testovalo sa len 2 mesiace

- Veľmi sľubný prístup, dobrý základ pre ďalšiu prácu

Agent	Created	Rules	ML	HS	IO	SIM
NuiBot	2020					