



Učení se hraní strategických her reálného času z demonstrací s využitím decentralizovaného MAS

JAN MALÝ

ME@JANMALY.NAME

[HTTPS://GITHUB.COM/HO
NZAMALY/KUSANAGI](https://github.com/HONZAMALY/KUSANAGI)



Cíl práce

- Vytvoření kompetentního bota pro StarCraft: Brood War
- Bot se rozhoduje na základě toho, co se naučil z demonstrací
- Pro kompetentního bota je nutné integrovat řadu technik a problém dekomponovat/abstrahovat



Motivace práce

Svět RTS je komplexitou podobný tomu našemu

- Výzkum se může uplatnit v reálném životě

Adaptace

- Většina botů není schopná se přizpůsobit situaci

Integrace doménové znalosti

- Zakomponovat doménovou znalost je těžké.
- Pro lidi je přirozené se učit z demonstrace, pro boty ne.

Dekompozice/abstrakce problému hraní RTS a integrace technik

- Současní boti nepoužívají žádnou jednotnou architekturu pro dekompozici/abstrakci problému a integraci technik

Proč nás UI ještě neporazila v RTS



Velké množství akcí

- Zjednodušená mapa 256x256 ve StarCraftu jen s 50 dělníky vede na 1 miliardu možných akcí

Velké množství stavů

- Šachy mají 10^{50} stavů, Go 10^{170} , Stracraft je odhadnut minimálně o několik řádů výše

Velice těžké i pro člověka

- Dokonce i člověk problém abstrahuje a dekomponuje

Jak je řešena dekompozice problému



Vlastní framework

- Deklarativní
- Doménově nezávislý

Základ - multiagentní systém

- Přirozené pro dekompozici Rozděl a panuj, a abstrakci
- Agenti se systémem sdílejí své touhy, záměry a přesvědčení. Systém definuje prostředí agentů

Vlastní přístup k vytváření plánů agenta

- Další úroveň dekompozice/abstrakce
- Adaptivní vytváření prostoru (sub)plánů a následné plánování na tomto prostoru

Jak se bot učí rozhodovat na základě demonstrací



Dataset pro učení

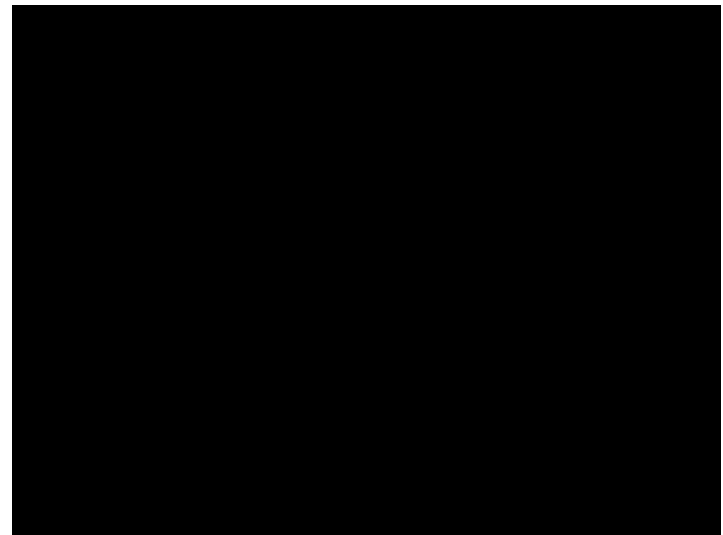
- Demonstrace her profesionálů z různých turnajů
- Sledují se převážně rozhodnutí na nejvyšší úrovni
- Zaznamenávání stavů a rozhodnutí v nich učiněných

Rozhodovací moduly

- Rozhoduje, jestli se bude daný plán vykonávat na základě pozorovaného stavu
- Trénování pomocí Markovského rozhodovacích procesů a techniky Inverse Reinforcement Learning

Jak si vede vytvořený bot

- Dokáže porazit vestavěnou UI v závislosti na použitém datasetu demonstrací a scénáři
- Dokáže stavět infrastrukturu, expandovat a (většinou) adaptuje se situaci
- Staví různé druhy jednotek podle situace a posílá je na nepřítele





Přínosy a prostor pro dotazy

- **Nová technika pro integrování doménové znalosti v této oblasti řešící celou řadu problémů**
 - Eliminace pevně zakódovaných pravidel
 - „Odpozorování“ rozhodování od profesionálů
- **Doménově nezávislý framework s vlastní adaptivní architekturou plánování**
 - Vhodný na dekompozici obdobných problémů
- **Reprezentace školy na různých turnajích botů v rámci AI konferencí (po vyladění)**