

Umělá inteligence a Starcraft 2

Jonáš Kordík

January 2020

1 Úvod

V této práci se budu zabývat tím, jak se umělá inteligence za pomoci deep neural network, supervised learningu a deep learningu naučila hrát komplexní počítačovou hru jako je Starcraft 2 a dokázala porazit jedny z těch nejlepších hráčů kteří tuto hru hrají.

1.1 Začátky

Začátky nebyly nejjednodušší, neboť i když bylo doposud AI schopné zvládat s velkými úspěchy hrát hry jako Mario nebo DOTA 2, nebylo schopné pobrat komplexitu hry jako je Starcraft 2. Nejlepší výsledky byly tedy dosaženy, když byly implementovány omezení pravidel a když byla část systému ručně naprogramována. AI byly také dány schopnosti a rychlost, která byla na nebo na vyšší úrovni než ta profesionálních hráčů. A i přes tyto omezení, úpravy a podporu nebylo AI schopné dosáhnout úrovně profesionálních hráčů. To dalo vznik Alphastar, které používalo deep neural network, která byla trénovaná z čistých herních dat za pomoci supervised a reinforced learning.

1.2 Co je Starcraft 2

Starcraft 2 je hra od vyvojářské společnosti Blizzard zasazená do fiktivního světa, která poskytuje bohatý víceúrovňový herní zážitek, který je zaměřený na to, aby otestoval lidský intelekt a vynalézavost. Tato hra patří mezi jedny z nejúspěšnějších her, které byly vytvořeny. I po 20 letech má bohatou komunitu, profesionální scénu a turnaje o peněžní ceny.

Je zde několik možností jak Starcraft hrát, ale profesionální e-sport formát je hra 1v1 (jeden hráč proti druhému) na 5 her. Hráč si na začátek vybere jednu ze tří hratelných ras- Zerg, Protoss a Terran, každá se svými charakteristickými schopnostmi a stylem hry. Je zvykem, že profesionální hráč se soustředí na jednu rasu. Každou hru začíná hráč s několika dělníky, kteří slouží ke sbírání surovin a stavění budov, ve kterých se po té trénují pokročilejší jednotky a zkoumají vylepšení pro ně. Pro vítězství musí hráč pečlivě vybalancovat kontrolu nad svojí ekonomikou v každé ojedinele hře - macro a až delikátní kontrolou jednotlivých jednotek na bojišti - micro. Potřeba pro balancování

dlouhodobé a krátkodobé cíle a adaptování se na nepředvídatelné situace je pro systémy, které jsou křehké a neflexibilní skoro nespílitelná. Pro překonání tohoto problému je potřebné obrovské pokroky v některých problémech AI jako jsou například: Herní teorie - Ve starcraftu totiž není jedna nejlepší strategie a proto je třeba, aby AI neustále zkoumalo a neustále rozšiřovalo své strategické a taktické obzory. Neperfektní informace - Narozdíl od šachů nebo hry Go, kde hráči vše vidí, ve starcraftu jsou důležité informace hráčům zkruté. Dlouhodobé plánování - jako v reálném světě jsou některé akce a následující reakce oddělené prodlevou a protože hra Starcraftu může trvat až hodinu, mohou se některé akce na začátku hry ukázat jako neoptimální nebo dokonce i špatné. Real time - narozdíl od stolních her, je hra hraná v reálném čase a vyžaduje adaptaci v každém dalším momentu hry. Velké pole akcí - Je možné kontrolovat desítky jednotek a budov najednou, které vytvoří kombinatorické pole možností. Alphastarova parametrizace hry měla přibližně 10 na 26 možných akcí v každém bodě hry.

2 Jak byl Alphastar trénovaný

Alphastarovo chování je generováno z deep neural network která získává informace přímo z herního rozhraní(ze seznamu jednotek a jejich vlastností) a výstupem je seznam instrukcí, které se skládají z následných akcí ve hře. Přesněji stavba této sítě aplikuje torso transforméru jednotkám, společně s LSTM jádrem, ukazatelovou sítí a centralizovanými základními hodnotami. Tato neuronová síť byla z počátku trénována za pomoci supervised learningu z anonymních her proti lidem. To dovolilo alphastaru učit se pomocí imitace základního micra a macra hráčů na hodnotícím žebříčku. Tento prvotní systém dokázal porazit už předtím přítomné Elite AI - které je na úrovni přirovnatelné ke zalté v žebříčku(průměrný hráč) v 95 procentech her.

2.1 Pokrok

V průběhu testování si Alphastar sám postupem času vymýšlel nové a nové strategie a proti-strategie, některé byly vylepšení těch předešlých a některé byly zbrusu nové a inovativní. Alphastar dokonce začal využívat různé "cheese" strategie, ale ty byly brzo škrtnuty v prospěch standartnějších strategií s větším procentem úspěšnosti a s menším rizikem.

Všichni agenti vytvořené alphastarem hráli s cílem porazit určitou skupinu hráčů podle žebříčkového systému a každý nově vytvořený agent si byl lépe schopný poradit s těžšími strategiemi s tím, že si v paměti zachoval, jak porazit ty předešlé. Někteří agenti byli nastaveni na to, aby porazili jednoho specifického hráče či strategii, jiní byli nastaveni na širší rozsah protivníků. Váha neuronové sítě každého agenta je aktualizována z her proti ostatním agentům pro optimalizování výsledného cíle. Toto je dosaženo za pomoci reinforcement learning algoritmu se schopností opětovného přehrání zkušeností, sebe-imitačního učení a strategického rozdělení.

Alphastar byl trénován v prostředí vytvořeném v Google's v3 TPUs, které podporovalo populaci agentů, kteří se učili z tisíců her běžících zároveň. Systém byl trénován po dobu 14 dní za použití 16ti TPU pro každého jednoho agenta. Každý agent zažil tréninkovou dobu ekvivalentní k 200 letům hraní Starcraftu 2. Konečný Agent obsahuje směsici nejefektivnějších strategií, které byly nalezeny, které běží na jednom počítačovém GPU.

Profesionální hráči starcraftu dokáží zadávat stovky příkazů za minutu. To je mnohem méně, než dokáže většina botů, kteří jsou schopní zadávat tisíce až desítky tisíc příkazů za minutu. Ve hrách proti profesionálním hráčům měl alphastar průměrných 280 akcí za minutu. To je méně než většina profesionálních hráčů, ale tyto akce byly značně přesnější a cílenější. Nižší akce za minutu, byly výsledkem toho, že se Alphastar učil ze zánamů her imitovat lidský styl hraní Starcraftu 2. K tomu ještě alphastar reagoval s prodlevou 350ms mezi pozorováním a akcí.

Alphastar hrál hru tím, že interagoval přímo s enginem hry přes jeho interface. To znamená, že mohl pozorovat vše co bylo viditelné na celé mapě bez toho aniž by musel hýbat kamerou. Na rozdíl od lidských hráčů, kteří musí kamerou ručně hýbat a vybírat si na co se zaměří. Druhá verze Alphastaru musela také ale už hýbat vlastní kamerou.

Alphastar byl zaměřen na jednu specifickou rasu - Protoss- pro redukci tréninkového času. Po té byl otestován proti profesionálnímu hráči TLO - Zerg- kterého porazilo 5-0. Sám hráč byl překvapen tím, jak silný byl tento daný agent, Který dokonce ukázal strategie, které hráč předtím ani nevídal v potaz. Po dalším týdnu tréninku hrál Alphastar proti Jednomu z nejlepších hráčů na světě MaNovi, kterého opět porazil 5-0. Hráč byl opět překvapen tím jak komplexní strategie byl schopný Alphastar demonstrovat. Tímto byl test ukončen a na systému se dále pracovalo. Postupně byly přidány a vytrénovány i zbylé dvě rasy. Nyní je Alphastar opravdu nejlepším hráčem hry Starcraft 2.