

Hraní her pomocí neuronových sítí

Algoritmus DQN a jeho vylepšení

Petr Buchal

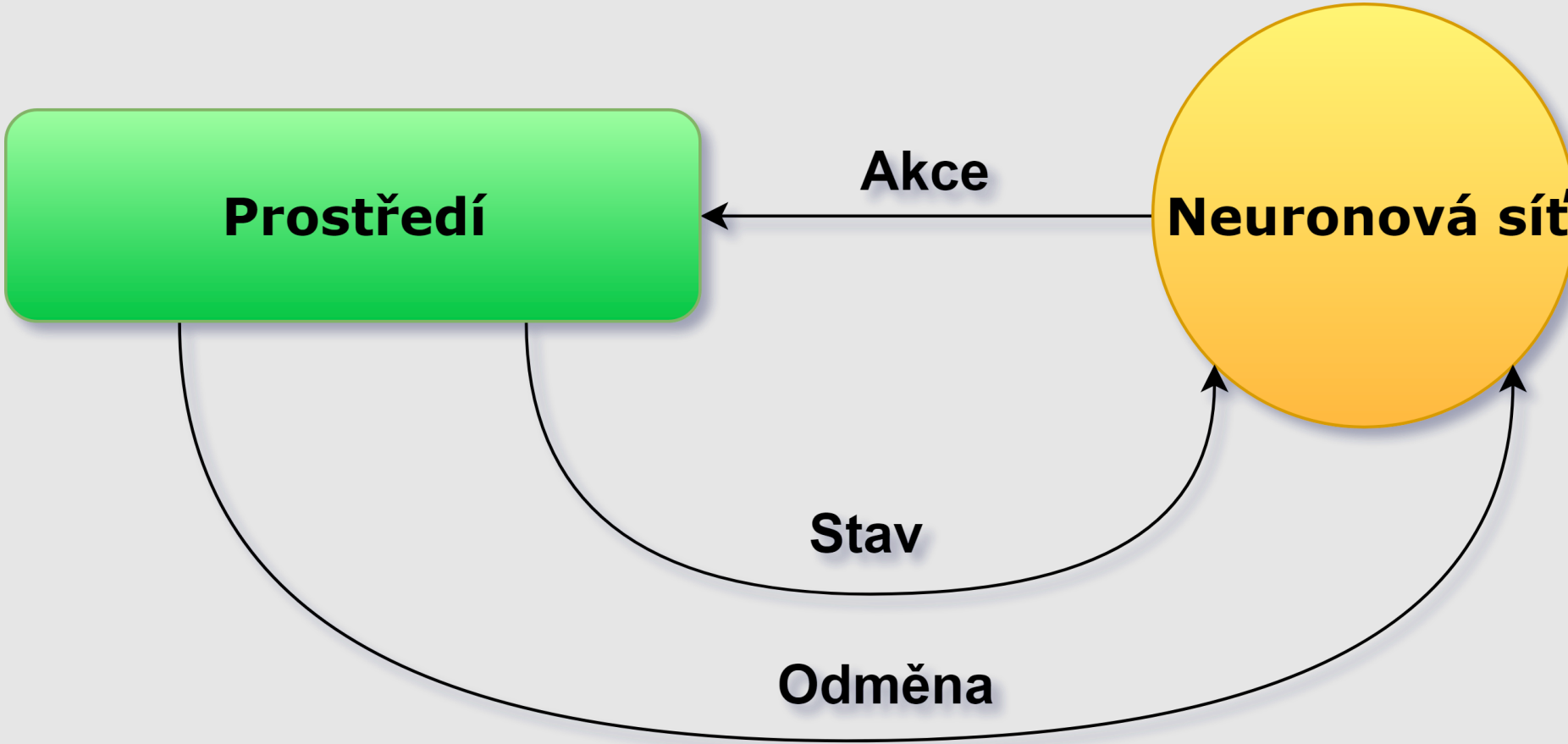


Motivace

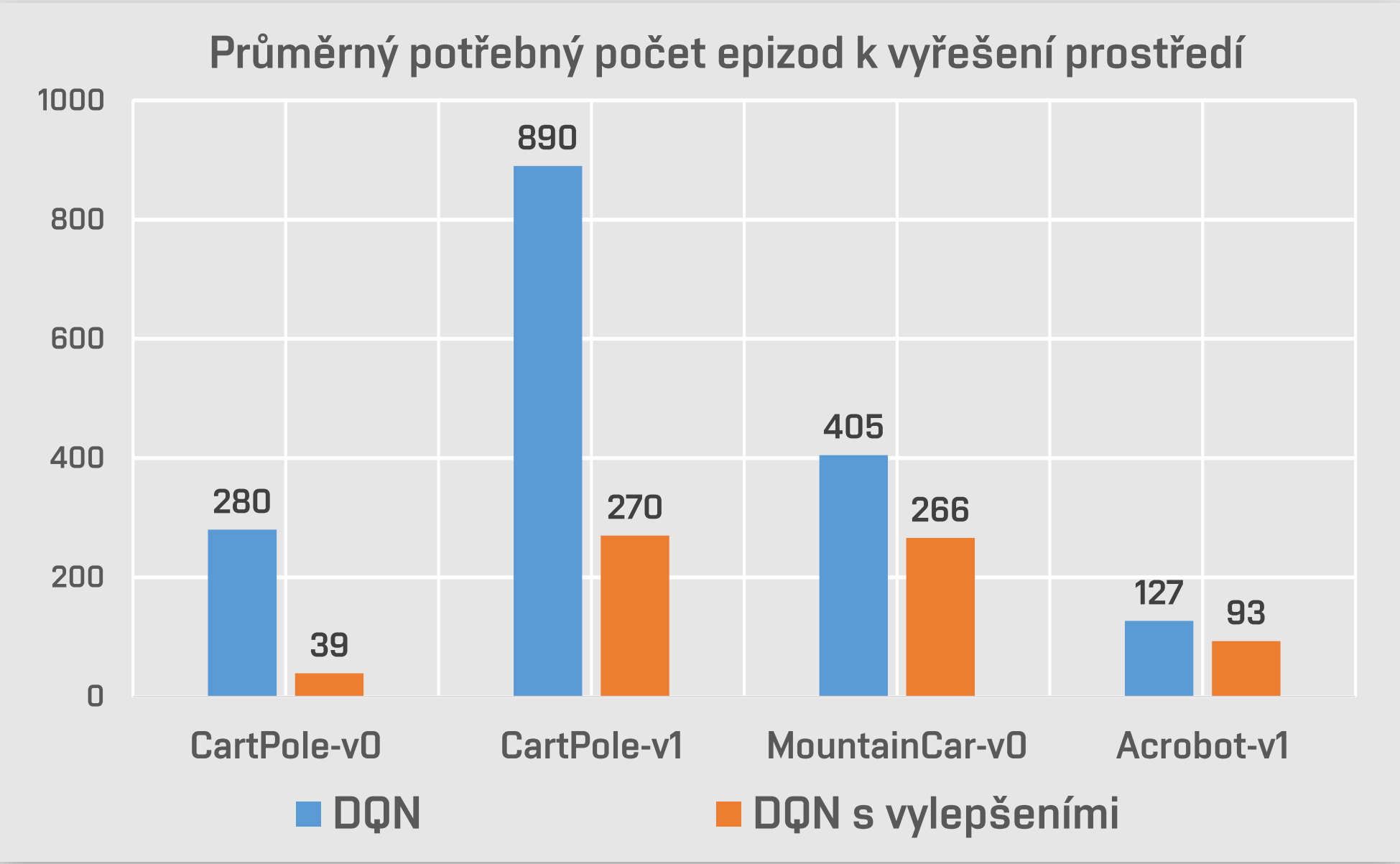
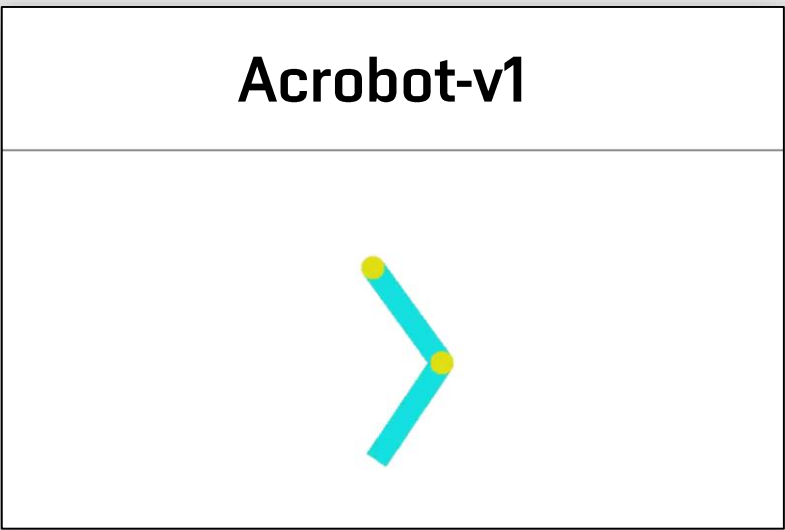
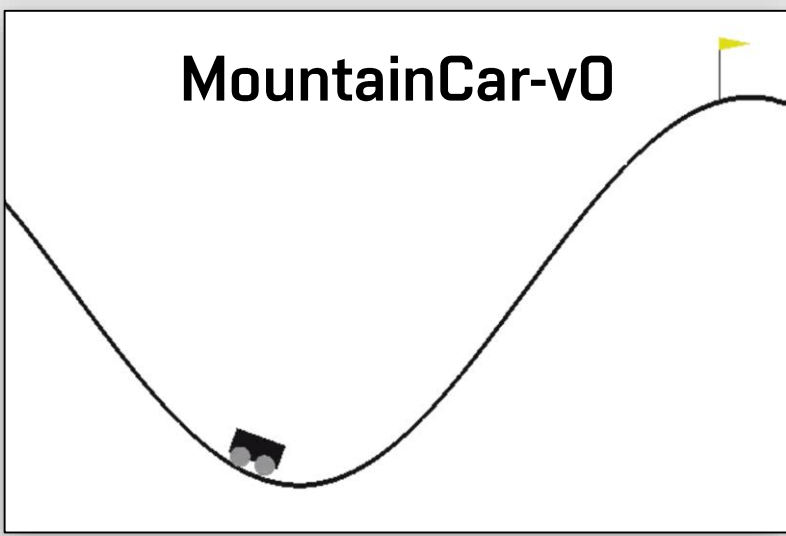
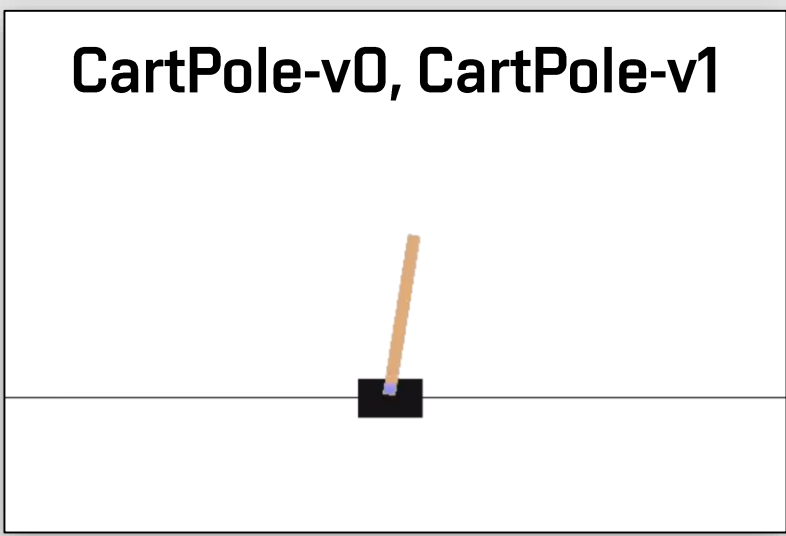
Historie má několik významných milníků, které jsou spojeny s uvedením agentů, kteří dokázali plnit konkrétní úlohu v konkrétním prostředí. Pravděpodobně nejznámějším takovým milníkem je vítězství počítače Deep Blue nad Garri Kasparovem v šachové partii v roce 1997. Tehdy se jednalo o agenta specializovaného na hraní šachů, tedy nic jiného než hrát šachy nedokázal. Navíc se neučil pomocí zpětné vazby, ale pouze prozkoumával nejlepší možné budoucí tahy. Když v roce 2013 představila firma DeepMind algoritmus DQN, byla to do jisté míry revoluce. Algoritmus se totiž dokázal úspěšně naučit hrát různé Atari hry, aniž by se specializoval na každou hru zvlášť a to vše pouze se snímky obrazovky jako vstupem.

DQN a vylepšení

DQN je algoritmus zpětnovazebního učení. Jedná se o typ Q-learningu, který Q-hodnoty místo ukládání do tabulky aproximuje pomocí neuronové sítě. Ta se učí z paměti, ve které je uložena historie stavů, kterými agent prošel. Od představení DQN bylo navrženo několik vylepšení, která až v řádu stovek procent vylepšují jeho efektivitu. Tato práce zkoumá, jak se jednotlivé varianty algoritmu chovají v různých prostředích. Mezi vylepšení algoritmu patří například přidání cílové sítě, její využití v dvojitém DQN, paměť s prioritou vzpomínek nebo duální architektura neuronové sítě.



Jednoduchá prostředí



Prostředí her

	2048	Space Invaders	Breakout	Beam Rider
Náhodný hráč	1011	153	1.3	361
Matice hracího pole (průměr)	3700	x	x	x
Matice hracího pole (nejlepší)	10 896	x	x	x
RAM (průměr)	x	398	3.8	674
RAM (nejlepší)	x	955	12	1284
Obraz (průměr)	x	280	3.1	433
Obraz (nejlepší)	x	830	8	804