

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Petr Zábaj
Název práce Efektivní paralelizace evolučních algoritmů
Rok odevzdání 2020
Studijní program Informatika **Studijní obor** Umělá inteligence

Autor posudku Roman Neruda **Role** Oponent
Pracoviště ÚI AV ČR, v.v.i.

Text posudku:

Předkládaná práce se zabývá návrhem efektivního paralelního evolučního algoritmu, který vychází z algoritmu s prokládáním generací a obohacuje jej o spekulativní vyhodnocování hodnot fitness funkce. Paralelizace evolučních algoritmů je obecně velmi důležitou oblastí výzkumu, protože umožňuje efektivní implementace v dnes běžných výpočetních prostředích obsahujících desítky až stovky výkonných CPU s relativně pomalou komunikací.

Text práce je rozdělen do šesti kapitol. Po úvodu následuje kapitola 1, která je úvodem do evolučních algoritmů a možností jejich paralelizace. Rešerše literatury je velmi dobře zpracována a obsahuje jasnou motivaci pro vlastní výzkum autora. Těžiště práce je pak v kapitolách 2 a 3, ve kterých je popsán návrh algoritmu a jeho implementace. Autor popisuje několik variant svého algoritmu IGEA-SE, které se liší přechody mezi generacemi podle schématu známého z evolučních strategií. Věnuje se i verzím ukončování spekulativního vyhodnocení. Poměrně rozsáhlá kapitola 4 obsahuje několik pečlivě provedených a zevrubně zhodnocených experimentů s různými variantami algoritmu a nastaveními jeho parametrů. Celá práce je stručně shrnuta v kapitole Závěr včetně naznačení možností dalšího výzkumu. Součástí práce je také příloha se zdrojovými texty v jazyce Python a kompletními výsledky všech provedených experimentů.

Za hlavní přínos práce považuji návrh původního algoritmu, který kombinuje stávající algoritmus s prokládáním generací s metodou spekulativního vyhodnocení. Autor vytvořil několik verzí algoritmu, které se liší v politice ukončování spekulativních vyhodnocení a jsou tak vhodné pro různé praktické úlohy. Celý algoritmus je důkladně otestován na různých syntetických datech, které simulují dva podstatné faktory reálných úloh – úspěšnost náhradního modelu použitého pro spekulativní vyhodnocení a rozložení doby vyhodnocování jedinců včetně jejich korelace s hodnotou fitness. Student tak ukázal, že v případě možnosti ukončení spekulativních vyhodnocení je jeho varianta vždy efektivnější než původní algoritmus. Další experimenty ukazují efektivitu variant algoritmu v závislosti na počtu procesorů. Velmi zajímavá je i relativní robustnost algoritmu k chybě náhradního modelu, kterou autor také experimentálně ověřoval. Součástí experimentů je i reálná úloha evoluční optimalizace perceptronové neuronové sítě v prostředí reálného výpočetního clusteru.

Dotazy k obhajobě se týkají hlavně oblasti, kterou autor zmiňuje v diskusi, a to jsou rozdíly v konečné hodnotě fitness u různých verzí algoritmu. Zajímalo by mne autorův názor, zda jde o statisticky významné rozdíly, či jen možný následek velkého rozptylu výsledků

jednotlivých běhů. Pokud jsou rozdíly statisticky podložené, co by mohlo být jejich příčinou?

Závěrem bych rád konstatoval, že jde o nadprůměrně kvalitní práci, jejíž výsledky si potenciálně zaslouží publikaci v mezinárodním prostředí.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

Pokud práci navrhuje na zvláštní ocenění (cena děkana apod.), prosím uveďte zde stručné zdůvodnění (vzniklé publikace, významnost tématu, inovativnost práce apod.).

Datum 9. září 2020

Podpis