

Oponentský posudek na bakalářskou práci

Frauknecht, J.: Koza a Prolog

Katedra logiky FF UK, únor 2015

Genetické algoritmy, inspirované biologickou evolucí, patří do třídy stochastických vyhledávacích algoritmů. Zejména v některých oblastech se ukázaly velmi účinné a prakticky aplikovatelné (např. návrh integrovaných obvodů, rozvrhování, ...). Práce se zabývá *genetickým programováním*, kdy jsou genetické algoritmy užívány k automatickému hledání počítačových programů. Cílem práce bylo zhodnotit možnosti použití Prologu v kontextu genetického programování a na konkrétní úloze provnat její řešení pomocí genetických algoritmů s řešením v Prologu. Výstupem práce měla být i praktická implementace genetického algoritmu v Prologu pro řešení zvolené úlohy. **Zadaný cíl práce z větší části splnila.**

Práce je rozdělena do sedmi kapitol. Úvodní kapitola shrnuje obsah práce a zmiňuje motivace, které k jejímu vypracování vedly. Následuje kapitola, která v krátkosti představuje obecný princip genetických algoritmů a zároveň zmiňuje dva biologicky inspirované algoritmy (PSO2 a Bee colony), které mezi genetické algoritmy nepatří. Třetí kapitola je krátkým úvodem do tzv. logického programování; v její poslední části je uvedena úloha, na kterou se autor v dalším textu soustředil. Čtvrtá kapitola obsahuje hrubý popis implementovaného algoritmu, detailnímu popisu je pak věnována kapitola 5. Kapitola 6 rozebírá výsledky algoritmu na konkrétních datech a závěrečná sedmá kapitola diskutuje příčiny vesměs neuspokojivých výsledků.

Práce je psána v angličtině. Z formálního hlediska čtenáře zarazí způsob citací — část citací odkazuje do seznamu literatury (jak je obvyklé v technických oborech), část citací je uvedena pod čarou (obvyklé spíše v humanitních oborech). Ze seznamu literatury vypadly háčky a čárky, některé citace jsou zvláštní (“Evolutionary computation, IEEE Transactions on”). Pozitivně naopak působí ilustrativní obrázky a grafy. Nicméně čtení ztěžují stylistické neobratnosti a formulační nepřesnosti (např. “many general concepts, that can be used as artificial intelligence (AI) algorithms” — koncepty mohou algoritmy inspirovat, *nemohou* však být jako algoritmy použity). Velké části vysvětlujících pasáží jsou pro nezasvěceného čtenáře velmi těžko čitelné: autor skáče z tématu na téma, mnoho věcí naznačí pouze v jedné větě, která se k okolí váže jen velmi volně; například (str. 8, první dva odstavce sekce 2.1):

The oldest and the best known BI algorithm is the evolutionary algorithm. Evolution refers to Charles Darwin’s theory of evolution by natural selection. This theory was the inspiration of 1950’s programmers like Holland, Ross, Bremermann, and Friedberg [Back et. al. 1997].

For purposes of this work it is necessary to distinguish the genome and the individual¹. For the EA, the genome is the individual representation stored in the computer memory (and it is this genome that is evolved). Fitness can be counted only for an individual (not for

the individual genetic code). This distinction was not introduced by Charles Darwin. The discovery of genotype was made in the 20th century².

Z celé práce není patrný rozdíl mezi genetickými algoritmy a genetickým programováním.

I po obsahové stránce má práce řadu nedostatků. Některá fakta jsou nepřesná: například přes nesporný význam J. Kozy je tvrzení z úvodu, že “John Koza is the author of the idea of genetic programming”, přinejmenším velmi sporné — naprosto pomíjí dřívější práce z 60. let (L. J. Fogel a další). Z mého pohledu největší slabinou textu je sedmá kapitola, která rozebírá příčiny neuspokojivých výsledků a naprosto pomíjí pravděpodobně hlavní důvod malé úspěšnosti: zásadním předpokladem efektivity genetických algoritmů pro řešení nějakého problému je rozložitelnost daného problému, t.j. předpoklad, že dvě částečná řešení problému dají dohromady úplné řešení. Zkoumané problémy (resp. jejich řešení) jsou příliš malé, než aby byl tento předpoklad splněn a genetický algoritmus v tomto případě degeneruje v ne příliš efektivní stochastický “hill-climbing” algoritmus. Vzhledem k tomu, že výsledný program nezvládá uspokojivě řešit ani takto malé instance problémů, nepříjde mi příliš rozumné pokračovat směrem naznačeným v sekci 7.4.

Navzdory výše uvedeným nedostatkům je z práce zřejmé, že autor věnoval problému velké úsilí. Samotná implementace GP v Prologu, ač z mého hlediska ne příliš rozumná (jak koneckonců naznačuje i citovaná práce Tang, L.R., Califf, M.E., Mooney, R.J.: An experimental comparison of genetic programming and inductive logic programming on learning recursive list functions), vyžaduje velkou dávku vynalézavosti. Z tohoto důvodu navrhuji práci i přes závažné nedostatky hodnotit stupněm **dobře**.

V Hejnicích a Praze, dne 6. února 2015

Mgr. Jonathan Verner, PhD.