

Posudek oponenta bakalářské práce

Autor/autorka práce: **Jakub Mikeš**

Název práce: **Desktopová aplikace pro generování školního rozvrhu malé školy**

Obsah práce

Cílem práce je tvorba aplikace pro generování školních rozvrhů, která má sloužit primárně jako benchmark pro metody automatického generování testů – jde tedy především o tvorbu aplikace, která nebude vyžadovat interakci s uživatelem a poskytne dostatečně otestované a dobře testovatelné, realistické jádro aplikace. Zároveň ale ukazuje jaké metody je možné pro generování rozvrhů použít. Pro hodnocení práce tedy není rozhodující kvalita generovaných rozvrhů, ale především kvalita aplikace jako takové a její otestovanost.

Kvalita řešení a dosažených výsledků

Aplikace je schopna generovat použitelné rozvrhy, což je ověřeno poměrně rozsáhlými testy popsanými na konci práce. Implementace genetických algoritmů ale není příliš efektivní a chybí porovnání s literaturou – není jasné jestli jsou genetické algoritmy obecně nevhodné ke generování rozvrhů nebo jde o problém s jejich implementací.

U popisu existujících reálných řešení není nikde zmíněno, jaké algoritmy generování používají (pokud jsou známy), což je v kontextu práce nejdůležitější informace. V sekci 2.3 bych čekal alespoň stručnou analýzu existující literatury, místo velmi obecného popisu s několika odkazy na konci odstavce. Dělení na heuristické a genetické algoritmy nedává smysl, genetické algoritmy jsou podmnožinou heuristických. Popis genetických algoritmů je poměrně povrchní, nijak se nevyjadřuje k jejich omezením nebo nevýhodám. V ukázce algoritmu 2.3 chybí informace o tom, jaká hodnota bude vrácena. Popis hledání hrubou silou, backtrackingu a „generuj a testuj“ je příliš stručný a po jeho přečtení není mezi metodami vidět žádný velký rozdíl.

Při tvorbě genetického algoritmu není zdůvodněna prakticky žádná z voleb, které student učinil. Popis fitness funkce je poměrně rozsáhlý a určitě by pomohlo, kdyby byla funkce zapsána i v podobě vzorce, ne jen volného textu.

Při popisu aplikace zcela chybí její návrh, text rovnou skočí k implementačním detailům (konkrétně k podobě datových souborů), místo aby byl na prvním místě uveden vysokoúrovňový popis struktury aplikace. Ani z diagramu 4.11 není patrné, jestli má aplikace nějakou strukturu (např. datový model, moduly s algoritmy moduly pro import a export dat).

Při popisu testování aplikace chybí informace o pokrytí kódu testy, vzhledem k účelu vytvářeného programu bych čekal, že tomu bude věnována větší pozornost.

Formální úroveň

Text je místy matoucí nebo obtížně srozumitelný (popis globálního a lokálního optima na s. 13, popis rulety na s. 14, ...).

Dekompozice aplikace je velmi nešťastná. Veškeré načítání a ukládání dat (bez ohledu na formát) je soustředěno v jedné třídě, která má tak za úkol všechno od parsování csv souborů po generování

výstupu do PDF. Třídy nesoucí data vypadají velmi podobně jako by reprezentovaly tabulky v ERA modelu, nejsem si jistý, jestli je to záměr nebo náhoda. Třída FitnessFunkce je obtížně čitelná a nepřehledná, především bych zde čekal metodu schopnou ohodnotit vlastnosti jedince (na vstupu jedinec, na výstupu hodnota fitness), místo toho je zde ale balík mnoha metod, často bez parametrů, zpracovávajících celou populaci. Jedinec, pro kterého je hodnota počítána je (nejspíš) nastaven jako atribut třídy, třída jako takový v sobě musí nést celý datový model, aby fungovala. Navíc je zde možné najít řadu prohřešků proti dobrým zvyklostem (jako jsou magická čísla – navíc jsou v kódu napevno konstanty, které nejspíš musí hodnotou odpovídat proměnným ze třídy GenetickyAlgoritmus, nebo nesrozumitelná jména proměnných). Samotný genetický algoritmus využívá pseudonáhodná čísla, není ale možné nastavit jeho seed a při testování tak nelze docílit opakovatelného běhu.

Práce s literaturou

Použitá literatura je relevantní, ale zejména pokud jde o genetické algoritmy, práci by prospěla důkladnější rešerše. Velká část citací je udělána velmi obecně – jako odkaz na několik zdrojů na konci odstavce – není tak možné ověřit ze kterého zdroje jaká informace pochází.

Splnění zadání

Zadání bylo po formální stránce splněno.

Dotazy k práci

Jak byly zvoleny konstanty v práci (počet prvků v populaci, počet jedinců ke křížení, počet generací při generování, způsob mutace a křížení)?

Jak je aplikace pokrytá testy, alespoň z hlediska základních metrik?

Vzhledem k uvedeným připomínkám navrhuji hodnocení známkou **dobře** a práci doporučuji k obhajobě.

V Plzni 22.5.2020

Ing. Richard Lipka, Ph.D.