

Protokol k semestrální práci z BI-ZUM

FIT ČVUT, LS 2019/2020

Jméno studenta: Tomáš Kořistka

Username: koristol

Název semestrální práce: OCR (Optical Character Recognition) snadno a rychle

1 ZADÁNÍ SEMESTRÁLNÍ PRÁCE

Mějme znaky nějaké abecedy jako obrázky, například velikosti 16x16 ve stupních šedi. Vaším úkolem je najít pozice několika pixelů tak, aby bylo možné podle hodnot těchto pixelů znaky abecedy od sebe rozeznat. Pro abecedu s 26 znaky s černo-bílými obrázky by teoreticky mělo stačit 5 pixelů, protože hodnoty v těchto pixelech mohou nabývat celkem $2^5 = 32$ různých kombinací což je více než 26.

- Pokuste se úlohu vyřešit pro latinskou abecedu.
- Najděte co nejmenší počet bodů, podle nichž jste schopni rozpoznávat znaky jedné z japonských abeced, a sice hiraganu.

2 STRUČNÝ ROZBOR, ANALÝZA PROBLÉMU ZADÁNÍ

Zadání vyžaduje nalezení několika různých pixelů, s jejich pomocí lze od sebe odlišit jednotlivé obrázky. Obrázek se skládá z $16 \times 16 = 256$ bodů. Tedy, hledá se pětice bodů, jejichž kombinace jasových úrovní je unikátní pro každý z 26, 50 znaků (obrázků) respektive.

3 METODY

Problém byl řešen evolučním algoritmem, a pro každý krok bylo naimplementováno několik metod.

3.1 Fitness

Každému jedinci je přiřazena fitness hodnota rovna počtu různých obrázků, které od sebe dokáže rozeznat. Jedná se o unikátní kombinace pixelů obrázku. V případě latinské abecedy je tedy nejvyšší hodnotou 26, v případě japonské hiragany pak 50.

V případě, že počet bodů jedince se mění (tj. v druhé části úlohy), bylo množství bodů zohledněno dělením této fitness hodnoty délkou chromosomu, tj. $\frac{\text{počet různých rozeznávaných obrázků}}{\text{délka chromosomu}}$, a to k motivaci kratšího genotypu.

3.2 Selektce

- Random selection – Jedinci jsou náhodně vybráni ke křížení
- Fitness proportionate selection – Jedinci s nejlepší fitness hodnotou mají největší šanci na výběr pro křížení
- Truncation selection – Ke křížení je vybrána poměrná část jedinců seřazených podle fitness hodnoty
- Tournament selection – Populace je rozdělena do skupin stejné velikosti, kde s každé skupiny je vybrán nejlepší jedinec.

3.3 Křížení

- K-point crossover - V obou rodičích je vybráno K pozic, které rozdělují genotypy rodičů na intervaly. Liché úseky (nebo sudé) jsou poté v potomcích vyměněny
- Uniform crossover – Potomci mají stejnou pravděpodobnost zdědit jednotlivé geny od obou rodičů

V žádné implementaci nemají rodiče šanci na přežití do další generace.

3.4 Mutace

Každý potomek má šanci mutovat a narušit tak konvergenci k neoptimální hodnotě. V rámci mutace mutuje náhodný počet genů jedince.

- Random reset – každý gen jedince má určitou šanci na resetování – náhradu náhodným novým

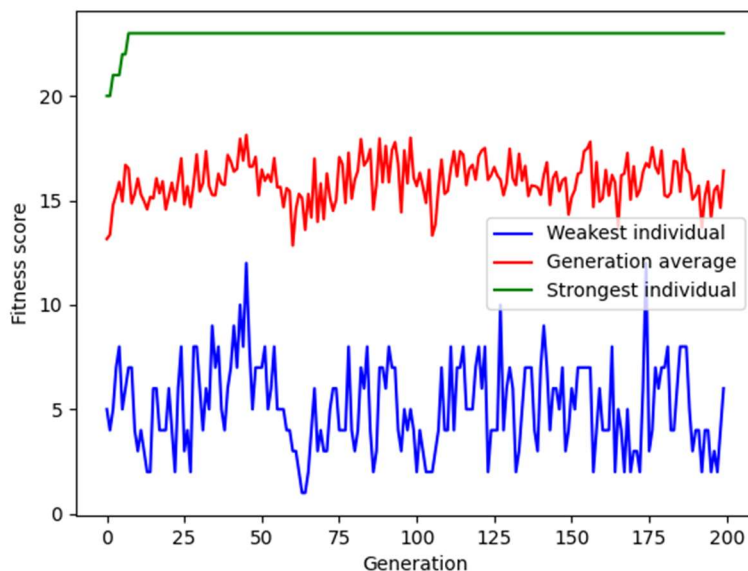
2. Single axis shift – každý gen jedince má šanci se posunout o jednu jednotku po jedné ose libovolným směrem. Nová pozice je poté podrobena operaci modulo, aby se předešlo přesunu bodu mimo rozměry obrázku.
3. Two axis shift – každý gen má šanci se posunout o jednu jednotku po obou osách libovolným směrem. Směry posunu po osách jsou na sobě nezávislé. Stejně jako u single axis shift, je na každou novou souřadnici aplikována operace modulo.

3.5 Ostrovy

Metoda ostrovů zvyšuje diverzitu populace. Úloha je rozdělena na podúlohy – ostrovy. Původní populace je rovnoměrně rozdělena mezi ostrovy a každý ostrov má svůj způsob výběru, křížení a mutace generace, pravděpodobnosti mutací genotypů a genů atp. V pravidelných intervalech dochází k migraci poměrné části jedinců mezi ostrovy, kde každá migrace z ostrova na ostrov je neidentickou permutací.

3.6 Katastrofa

Katastrofa je dalším nástrojem pro zachování diverzity. Katastrofa nastane v případě, že aktuální generace zdegenerovala na malé množství rozdílných jedinců (vyhodnocováno vzorem $\frac{\text{počet unikátních jedinců}}{\text{počet jedinců}} < k$, kde k je desetinné číslo, v aktuální implementaci 0.05). Spuštěním katastrofy na ostrově dojde k odstranění většiny populace a zachování malého zbytku populace (která je téměř identická) a nahrazení odstraněné populace novou náhodnou.



Obr. 1

Na obrázku 1 výše je záznam jednoho běhu programu úlohy b) (hledání nejmenšího množství bodů k rozeznání znaků hiragany), s 4 ostrovy a populací velikosti 200 (tedy 50 jedinců na ostrov). Výrazné poklesy v průměrné fitness hodnotě jsou způsobeny vyvolanou katastrofou. Dále lze vidět, že optimální řešení je nalezeno brzy, a vzhledem ke své fitness hodnotě se toto řešení dostává do dalších generací.

4 VÝSLEDKY

Algoritmus nalezne v i rámci krátkého běhu (10-20 generací) řešení, které rozpozná alespoň 25 různých obrázků. S větším množstvím generací algoritmus vždy najde nejlepší řešení. Příkladem pětice bodů, které dokáží odlišit od sebe všechny znaky latinské abecedy jsou body [1, 3], [5, 5], [11, 12], [13, 6], [14, 11]. V případě druhé části bylo nejlepší nalezené řešení s počtem pixelů 5, a to sice body se souřadnicemi [9,6], [11,8], [4,4], [6,7], [7,3].

Zlepšení se dalo docílit laděním parametrů, konkrétně velikost populace a počet iterací, či častější mutací.