Fabian Kuriata 238016

**Sztuczna inteligencja i inżynieria wiedzy**

Ćwiczenie 1

Algorytmy genetyczne

**Poruszany problem**

Problemem do rozwiązania jest problem mobilnego złodzieja (TTP travelling thief problem), który zawiera w sobie dwa inne podproblemy to rozwiązania, a mianowicie problem komiwojażera oraz problem plecakowy.

Jednym z możliwych podejść do rozwiązania tego problemu jest stworzenie programu ewolucyjnego, który przy ustawieniu odpowiednich parametrów powinien nam dać satysfakcjonujący wynik.

**Opis algorytmu**

Program zakłada, że złodziej kradnie jeden najbardziej wartościowy przedmiot w stosunku do wagi przy odwiedzeniu każdego miasta, o ile jest w stanie go udźwignąć, jeśli nie to próbuje z kolejnym aż do wyczerpania wszystkich możliwości, sam złodziej przed wejściem nie wie co się w mieście znajduje oraz że musi odwiedzić wszystkie podane miasta. Przy czym najlepiej, aby jak najkrócej mu to zajęło na co oprócz przebytej drogi wpływa również ciężar plecaka.

Zatem naszym zakładanym celem jest minimalizacja różnicy wartości wszystkich przedmiotów zabranych przez złodzieja i czasu poświęconego na przebycie drogi. Co można ująć w poniższy wzór.

**G ( x , y ) = g ( y ) −f ( x , y )**

Gdzie

**g(y)** – wartość wszystkich przedmiotów

**f(x, y)** – całkowity czas podróży z trasy (x) z uwzględnieniem ciężaru plecaka ( y )

Ten wzór jest wykorzystywany przy ocenie osobnika przy ewaluacji

Programowanie genetyczne bazuje na 4 aspektach:

- selekcja

- krzyżowanie

- mutacja

- ewaluacja

Podczas selekcji wykorzystana została metoda turniejowa, która zakłada jeden parametr ( tour ) określający ilu osobników bierze udział w turnieju. Polega on na wybraniu najlepszego osobnika ze względu na jego ocenę i dodaniu go do nowej populacji.

Metoda krzyżowania natomiast wykorzystuje sposób OX (ordered crossing), który zapobiega powtarzającym się genom ( miastom) przy skrzyżowaniu genów dwóch rodziców.

Mutacja zamienia miejscami dwa losowe geny ( miasta ), przy czym z założenia, że pierwszy podane miasto, zawsze jest miastem, z którego zaczynami to pomija zamianę pierwszego.

Podczas ewaluacji oceniani są wszyscy osobnicy z nowej populacji, przed dokonaniem ewentualnej kolejnej selekcji.

Sztuką znalezienia jak najlepszego wyniku przy użyciu programu ewolucyjnego, jest znalezienie odpowiednich parametrów. Problemem jest, że nikt tak naprawdę nie wie co oznacza „odpowiednie”, w sensie nie ma jednej słusznej odpowiedzi w jaki sposób je ustawiać.  
Natomiast przy badaniu różnych najlepiej skrajnych wartości parametrów możemy zauważyć pewne zależności, które nakłaniają do stawiania różnych tez.

**Testowanie algorytmu**

Do testowania działania użyto 6 zestawów danych ustawiając parametry zawsze tak samo.

Wyniki są na bazie uśrednienia z 10 uruchomień, natomiast ich stabilność określa odchylenia standardowe, które jest obliczane ze średnich wszystkich pokoleń.

**Skróty używane poniżej**

**PopSize** – rozmiar populacji

**Gen** – ilość generacji

**Px** – prawdopodobieństwo krzyżowania

**Pm** – prawdopodobieństwo mutacji

**Tour** – parametr metody turniejowej

**Dev** – odchylenie standardowe

**Best** – ostatni najlepszy osobnik

**Avg** – ostatni średnia ocen osobników

**Worst** – ostatni najgorszy osobnik

**Ustawienia parametrów**

PopSize = 100 Gen= 100 Px = 70% Pm = 10% Tour = 5

**Wyniki**

**Dev** = 3080,10 **Best** = -48676,5 **Avg** = -67695,8 **Worst** = -116930,03

**Dev** = 6331,29 **Best** = -94005,3 **Avg** = -131870,14 **Worst** = -183264,39

**Dev** = 908,21 **Best** = 9066,8 **Avg** = 5070,23 **Worst** = -1212,73

**Dev** = 536,26 **Best** = 10393,81 **Avg**: 6819,32 **Worst** = 1828,89

**Dev** = 21744,46 **Best** = -1306409,29 **Avg** = -1349902,83 **Worst** = -1403518,35

**Dev** = 18717,06 **Best** = -1314291,7 **Avg** = -1357002,86 **Worst** = -1403141,42

**Eksperymentowanie z wartościami parametrów**

**1. Brak krzyżowania**

**PopSize** = 100 **Gen** = 100 **Px** = 0% **Pm**: 70% **Tour**: 5

**Dev** = 3332,58 **Best** = -74025,84 **Avg** = -101069,35 **Worst** = -163163,55

Bez krzyżowania rozwój jest wolniejszy i jest podatny na wahania zależne od tego jak wiele genów mutujemy.   
W przypadku algorytmu, który wygenerował powyższe wyniki mutowany był jedyny jeden gen stąd też widoczne są niewielkie wahania w rozwoju.

**2. Brak mutacji**

**PopSize** = 100 **Gen** = 100 **Px** = 70% **Pm** = 0% **Tour**: 5

**Dev** = 6931,16 **Best** = -88840,6 **Avg** = -88840,6 **Worst** = -88840,6

Bez stosowania mutacji w zależności od presji selekcyjnej cała populacja zmierza do posiadania tych samych genów ( składa się z tych samych miast ), czyli ujednolicenia wyników i braku dalszego rozwoju.

**3. Wpływ parametru metody turniejowej**

**PopSize** = 100 **Gen** = 100 **Px** = 0,7 **Pm** = 0,1 **Tour** = 100

**Dev** = 651,95 **Best**: =17389,3 **Avg** = 15738,54 **Worst** = 12123,26

**PopSize** = 100 **Gen** = 100 Px = 0,7 Pm = 0,1 **Tour**: 50

**Dev** = 659,34 **Best** = 16758,16 **Avg** = 14570,34 **Worst** = 9755,04

Przy dużej presji selekcyjnej występuje bardzo szybko wzrost rozwoju, lecz równie szybko on wyhamowuje. Wydaje się to być dobrym sposobem na uzyskanie szybciej „dobrego” rozwiązania, ale trzeba się liczyć z tym, że przy takiej presji zmienia się zróżnicowanie populacji, co na dłuższą metę prowadzi do stagnacji.

**4. Wpływ rozmiaru populacji i ilości generacji**

**PopSize** = 1000 **Gen** = 100 **Px** = 70% **Pm** = 10% **Tour** = 50

**Dev** = 64,07 **Best** = 18304,84 **Avg** = 16003,04 **Worst** = 10028,39

**PopSize** = 1000 **Gen** = 100 **Px** = 70% **Pm** = 10% **Tour** = 50

**Dev** = 1726,74 **Best** = -875914,4062 **Avg** = -905169,52 **Worst** = -949397,73

Zwiększenie populacji przyczynia się do szybszego rozwoju, polepszenia wyników, gdyż rośnie zróżnicowanie osobników.

**PopSize** = 10 **Gen** = 1000 **Px** = 70% **Pm** = 10% **Tour**: 5

**Dev** = 33948,40 **Best** = 5760,88 **Avg** = 3996,97 **Worst** = 1645,25

Gdy liczba populacji jest zbyt mała i to przy sporej presji selekcyjnej rozwój szybko wyhamowuje oraz występują wahania przy wynikach przez zbyt małe zróżnicowanie osobników.

**PopSize** = 1000 **Gen** = 1000 **Px** = 70% **Pm** = 10% **Tour** = 50

**Dev** = 4726,14 **Best** = -38236,7 **Average** = -56883,2 **Worst** = -114980,81

**PopSize** = 1000 **Gen** = 1000 **Px** = 70% **Pm** = 10% **Tour** = 50

**Dev** = 13515,07 **Best** = -458609,38 **Average** = -481581,61 **Worst** = -522637,9872

Przy zwiększonej populacji i większej ilości generacji oraz zastosowaniu krzyżowania, mutacji oraz niewielkie presji selekcyjnej widać ciągły wzrost ewolucyjny, lecz nie jest on linią prostą, a z generacji na generację robi coraz to mniejsze postępy**.**

**5. Zmiana parametrów krzyżowania i mutacji**

**PopSize** = 100 **Gen** = 100 **Px** = 50% **Pm** = 50% **Tour** = 5

**Dev** = 9599,87 **Best** = -96931,54 **Avg** = -139873,30 **Worst** = -219404,23

**PopSize** = 100 Gen = 100 Px = 50% Pm = 50% **Tour** = 5

**Dev** = 422,67 **Best** = 8329,25 **Avg** = 4423,61 **Worst** = -1421,28

Gdy mutacja oraz krzyżowanie występuje z takim samym prawdopodobieństwem to otrzymane wyniki są gorsze niżeli w przypadku Px=70%, Pm=10%, spowodowane prawdopodobnie zmniejszeniem różnorodności, gdyż sama mutacja zmienia jeden gen, a szansa na krzyżowanie, które najbardziej wpływa na zmianę jest mniejsze.

**6. Zmiana parametru turniejowego, rozmiaru populacji oraz ilości generacji**

**PopSize** = 10 **Gen** = 1000 **Px** = 70% **Pm** = 10% **Tour** = 2

**Dev** = 27849,84 **Best** = 356,35 **Avg** = -2067,49 **Worst** = -4669,4

Zbyt mała liczba populacji sprawia, że wyniki w dużej mierze zależą od pierwszej wygenerowanej populacji co wydać po wartości odchylenia.

**PopSize** = 1000 **Gen** = 100 **Px** = 70% **Pm** = 10% **Tour**: 2

**Dev** = 32,93 **Best** = 5225,40 **Avg** = -2320,4 **Worst** = -11592,8

**PopSize** = 1000 **Gen** = 100 **Px** = 70% **Pm** = 10% **Tour** = 2

**Deviation** = 600,24 **Best** = -1578830,7 **Avg** = -1704581,8 **Worst** = -1839873,33

Niska presja selekcyjna sprawia, że więcej gorszych osobników przechodzi do kolejnej populacji. Przez to jest ona bardziej zróżnicowana, ale zbyt duże zróżnicowanie prowadzi do wolniejszego rozwoju.

**PopSize** = 10 **Gen** = 1000 **Px** = 70% **Pm** = 10% **Tour**: 2

**Dev** = 1580231,18 **Best** = -1539402,13 Avg = -1560107,32 **Worst** = -1578744,22

Zmniejszona liczba populacji przyczynia się do mniejszej różnorodności co wiąże się z gorszym rozwojem.