

Part 2

Τρίμας Χρήστος , Παντελής Κωσταντίνος
2016030054 , 2015030070

Τεχνητή Νοημοσύνη
LAB41744549
ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

April 30, 2020

Υλοποίηση ενός Γενετικού Αλγορίθμου για το Πρόβλημα του Χρονοπρογραμματισμού Προσωπικού WHPP

A) Εισαγωγή.

Στο δεύτερο μέρος της 1ης προγραμματιστικής εργασίας του μαθήματος, κληθήκαμε να υλοποιήσουμε έναν γενετικό αλγόριθμο για το πρόβλημα του χρονοπρογραμματισμού του προσωπικού ενός νοσοκομείου, για ένα διάστημα 14 ημερών και 30 εργαζομένων.

Αρχικά περιγράφονται όλα τα βασικά κομμάτια του γενετικού μας αλγορίθμου και οι σχεδιαστικές μας επιλογές. Στη συνέχεια ακολουθεί σχολιασμός των αποτελεσμάτων και ιδέες για βελτίωση του προγράμματος.

B) Γενετικός Αλγόριθμος.

Όλοι οι γενετικοί αλγόριθμοι ακολουθούν έναν βασικό σκελετό υλοποίησης(Figure 1). Επί γραμματικά, τα βήματα είναι τα εξής:

- 1) Δημιουργία ενός "ψευδό"-τυχαίου πληθυσμού.
- 2) Έλεγχος συνέπειας.
- 3) Αξιολόγηση αντικειμενικής συνάρτησης.
- 4) Ικανοποίηση των κριτηρίων βελτιστοποίησης ?

Σε αυτό το σημείο, αν τα κριτήρια δεν ικανοποιούνται, συνεχίζουμε με την δημιουργία ενός νέου πληθυσμού, στον οποίο εφαρμόζονται οι εξής γενετικοί παράγοντες:

α) Selection

β) Crossover

γ) Mutation

δ) Check

Επιστροφή πίσω στο βήμα 3.

Όταν και εφόσον τα κριτήρια στο βήμα 4 ικανοποιούνται, συνεχίζουμε στην εξαγωγή του αποτελέσματος.

Σε γενικές γραμμές, αυτά είναι τα βήματα για τον σχεδιασμό ενός γενετικού αλγορίθμου. Στη συνέχεια της αναφοράς, ακολουθεί αναλυτική περιγραφή για κάθε βήμα.

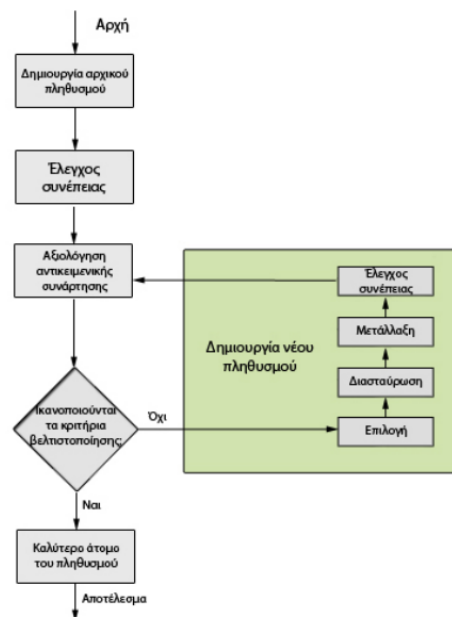


Figure 1: Genetic Algorithm

Γ) Αρχικοποίηση πληθυσμού.

Με τον όρο πληθυσμός, χαρακτηρίζεται το πλήθος των λύσεων του προβλήματος μας ή αλλιώς το πλήθος των προγραμμάτων που περνάνε τους αυστηρούς ελέγχους (Hard Constraints). Το πλήθος των χρωμοσωμάτων (δηλαδή ο πληθυσμός) έχουν δυο κύριους τρόπους αρχικοποίησης.

1) Τυχαία αρχικοποίηση (η επιλογή μας).

2) Ευριστική αρχικοποίηση.

Μέσω της βιβλιοθήκης numpy και της συνάρτησης randint που εγκαταστήθηκε για την python, και για ένα άνω "φράγμα" αριθμού προγραμμάτων(population size=3000/2000), δημιουργήθηκε ένα τυχαίος αριθμός χρωμοσωμάτων, τα οποία περνάνε τους αυστηρούς περιορισμούς(όχι πάντα).

Στην αμέσως επόμενη γενιά, ο αριθμός των χρωμοσωμάτων είναι πιο κοντά σε αυτόν της αρχικοποίησης, και αυτό οφείλεται στις διαδικασίες επιλογής-διασταύρωσης-μετάλλαξης.

Σε ότι αφορά το πληθυσμιακό μας μοντέλο, επιλέχτηκε ένα δυναμικό μοντέλο το οποίο σε κάθε γενιά αλλάζει τον αριθμό των χρωμοσωμάτων που περνάνε στην επόμενη. Για να αποσαφηνιστεί αυτό, στην συνέχεια της αναφοράς, ακολουθεί το κριτήριο τερματισμού του αλγορίθμου μας. Να τονιστεί σε αυτό το σημείο, ότι για την συνέπεια παραγόμενων χρωμοσωμάτων, η συνάρτηση feasibility() ακολουθεί αυστηρά τους περιορισμούς της εκφώνησης που βρίσκονται στην σελίδα 15.

Δ) Κριτήριο Τερματισμού.

Ως κριτήριο τερματισμού του αλγορίθμου εκ πρώτης όψης είναι ο αριθμός των επαναλήψεων που ορίσαμε με βάση την εκτέλεση του αλγορίθμου ως 400 γενιές. Ωστόσο, το πραγματικό κριτήριο είναι ο μηδενισμός των χρωμοσωμάτων που διέρχονται στην επόμενη γενιά. Πιο συγκεκριμένα, καθώς επιλέχτηκε ανακατανομή του πληθυσμού από γενιά σε γενιά, με τον μηδενισμό των χρωμοσωμάτων, καταλήγει το πρόγραμμα σε έναν μικρό αριθμό από βέλτιστες λύσεις(με την καλύτερη ποινή βάση των hard και soft constraints).

Ε) Επιλογή.

Η επιλογή, είναι η διαδικασία επιλογής γονέων που "ταιριάζουν" και συνδυάζονται για την αναπαραγωγή της επόμενης γενιάς χρωμοσωμάτων. Η επιλογή του αλγορίθμου υλοποίησης, είναι ζωτικής σημασίας για την σύγκλιση του αλγορίθμου σε ένα βέλτιστο αποτέλεσμα. Η συγκεκριμένη υλοποίηση, έγινε ακολουθώντας τον αλγόριθμο του τουρνουά. Η επιλογή αυτή οφείλεται στο γεγονός, ότι είναι αρκετά διαδεδομένη τεχνική στην βιβλιογραφία που υπάρχει, ενώ παράλληλα είναι αρκετά εύκολη η υλοποίηση του αλγορίθμου.

Βασική ιδέα, είναι η τυχαία επιλογή χρωμοσωμάτων, και η σύγκριση αυτών για την εύρεση του "καλύτερου" γονέα(Figure 2).

ΣΤ) Διασταύρωση.

Εννοιολογικά, είναι η ίδια διαδικασία με την φυσική διασταύρωση δυο οργανισμών. Πρόκειται για την αναπαραγωγή ενός παιδιού από το γενετικό υλικό δυο γονέων. Ζητούμενο της εργασίας είναι η δημιουργία δυο τελεστών διασταύρωσης. Επιλέχτηκε η διασταύρωση ενός σημείου (one point) και η διασταύρωση πολλών σημείων(five points).

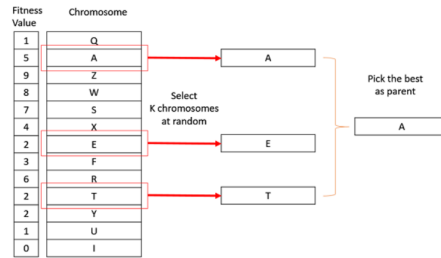


Figure 2: Tournament selection

Για τον 1ο τελεστή(Figure 3-4), τυχαία επιλέγεται το σημείο διασταύρωσης των γονέων, και στη συνέχεια γίνεται η "ένωση" του "γενετικού υλικού".

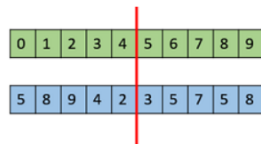


Figure 3: Parents

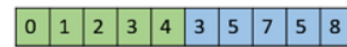


Figure 4: 1-point crossover Child

Αντίστοιχη διαδικασία, ακολουθεί και ο five-point crossover operator(Figure 5-6) αλλά για 5 τυχαία σημεία.

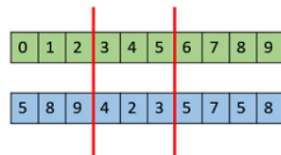


Figure 5: Parents

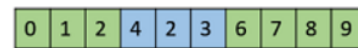


Figure 6: Multi-point crossover Child

Z) Μετάλλαξη.

Ο δεύτερος από τους δύο γενετικούς παράγοντες, και ο τρόπος για να ξεφεύγουμε από τοπικά ακρότατα, η μετάλλαξη αποτελεί μια τυχαία επέμβαση σε κάποιο χρωμόσωμα για την εύρεση κάποιας άλλης λύσης. Η μετάλλαξη αφορά το κομμάτι ανακάλυψης του χώρου καταστάσεων και αποτελεί και αυτή ζωτικής σημασίας παράγοντα για την σύγκλιση σε κάποιο αποτέλεσμα. Υλοποιήθηκαν δύο αλγόριθμοι μετάλλαξης. Ο Random resetting και ο swap mutation.

Ο πρώτος αποτελεί μια διεύρυνση του bit flip mutation algorithm, για αναπαράσταση ακεραίων. Συγκεκριμένα, επιλέγει τυχαία κάποιο στοιχείο του χρωμοσώματος και το αλλάζει τιμή σύμφωνα με τον εξής μετασχηματισμό:

$$0 \rightarrow 3, 1 \rightarrow 2, 2 \rightarrow 1, 3 \rightarrow 0$$

Ο Swap mutation(Figure 7), με την σειρά του επιλέγει τυχαία δύο στοιχεία του χρωμοσώματος και τα αλλάζει θέση.

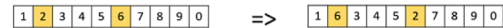


Figure 7: Swap mutation

Η) Πειραματικά Αποτελέσματα.

Η εκτέλεση της επιλογής νέου πληθυσμού, καθώς και των γενετικών τελεστών, πραγματοποιείται με κάποια πιθανοτικά κριτήρια. Συγκεκριμένα, η πιθανότητα επιλογής και διασταύρωσης πρέπει να είναι αρκετά μεγάλη έτσι ώστε να παρθεί το καλύτερο δυνατό αποτέλεσμα, ενώ η πιθανότητα μετάλλαξης πρέπει να είναι αρκετά μικρή, όχι όμως μηδενική, καθώς αυτό θα σημαίνει πρόορη σύγκλιση του αλγορίθμου.

Μετά από συνεχείς εκτελέσεις του προγράμματος για τυχαία φράγματα πληθυσμού, πιθανότητας επιλογής-διασταύρωσης-μετάλλαξης το ιδανικό που επιλέχτηκε ήταν:

- 1) population size = 3000 και 2000
- 2) Probability selection = 0.85
- 3) Probability crossover = 0.85
- 4) Probability mutation = 0.05

Δεν δίνεται κάποιο άλλο σεντ τιμών, καθώς οποιοδήποτε σεντ πιθανοτήτων με μικρή απόκλιση από αυτές τις τιμές έχει σχεδόν πανομοιότυπα αποτελέσματα.

Ακολουθούν τα αποτελέσματα σύμφωνα με το ζητούμενο της εκφώνησης.

- 1) tournament selection - one point crossover - random reseting - population = 2000.

- a) Best = 57435, Avg = 54456
- b) Best = 68323, Avg = 65695
- c) Best = 62042, Avg = 59380
- d) Best = 58931, Avg = 58094
- e) Best = 64931, Avg = 61367
- AvgBest : = 62332.4**

- 2) tournament selection - one point crossover - swap mutation - population = 2000.

- a) Best = 62834, Avg = 62834
- b) Best = 55742, Avg = 55742

- c) Best = 63240, Avg = 62881
- d) Best = 60740, Avg = 60740
- e) Best = 67243, Avg = 65397
- AvgBest* : = 61959.8**

3) tournament selection - one point crossover - random reseting - population = 3000

- a) Best = 62438, Avg = 62438
- b) Best = 58733, Avg = 58733
- c) Best = 71227, Avg = 68850
- d) Best = 53837, Avg = 53435
- e) Best = 61043, Avg = 56464
- AvgBest* : = 61455.6**

4) tournament selection - one point crossover - swap mutation - population = 3000

- a) Best = 60731, Avg = 58051
- b) Best = 70432, Avg = 61082
- c) Best = 69237, Avg = 65095
- d) Best = 56731, Avg = 56731
- e) Best = 65237, Avg = 65237
- AvgBest* : = 64473.6**

5) tournament selection - five point crossover - random reseting - population = 3000

- a) Best = 64832, Avg = 61523
- b) Best = 69834, Avg = 67623
- c) Best = 66943, Avg = 62482
- d) Best = 62524, Avg = 60849
- e) Best = 75630, Avg = 68815
- AvgBest* : = 67952.6**

6) tournament selection - five point crossover - swap mutation - population = 3000

- a) Best = 51644, Avg = 51644
- b) Best = 66725, Avg = 66725
- c) Best = 49244, Avg = 49244
- d) Best = 61741, Avg = 60774
- e) Best = 59134, Avg = 58909
- AvgBest* : = 57697.6**

7) tournament selection - five point crossover - swap mutation - population = 2000

- a) Best = 62039, Avg = 59261
- b) Best = 61238, Avg = 59711

- c) Best = 62151, Avg = 58309
- d) Best = 65230, Avg = 60194
- e) Best = 61644, Avg = 57431
- AvgBest** : = 62460.4

8) tournament selection - five point crossover - random reseting - population = 2000

- a) Best = 63373, Avg = 62155
- b) Best = 66831, Avg = 63107
- c) Best = 57126, Avg = 57126
- d) Best = 59035, Avg = 58868
- e) Best = 74734, Avg = 70099
- AvgBest** : = 64219.8

Γραφήματα καλύτερων περιπτώσεων.

Ακολουθούν οι τίτλοι και τα γραφήματα

- 1) tournament - one point crossover - random reseting - population = 3000
- 2) tournament - one point crossover - swap mutation - population = 3000
- 3) tournament - one point crossover - random reseting - population = 2000
- 4) tournament - one point crossover - swap mutation - population = 2000
- 5) tournament - multi point crossover - random reseting - population = 3000
- 6) tournament - multi point crossover - swap mutation - population = 3000
- 7) tournament - multi point crossover - random reseting - population = 2000
- 8) tournament - multi point crossover - swap mutation - population = 2000

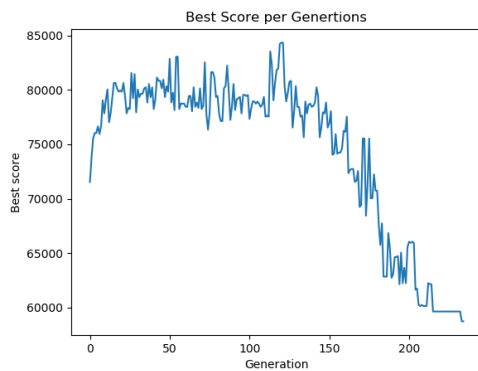


Figure 8: Best score(1)

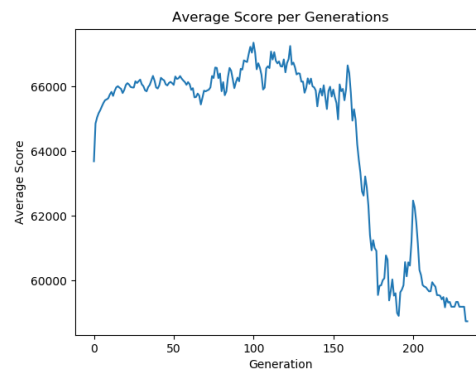


Figure 9: Average Score(1)

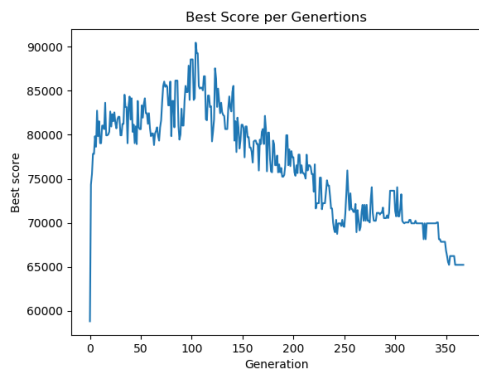


Figure 10: Best score(2)

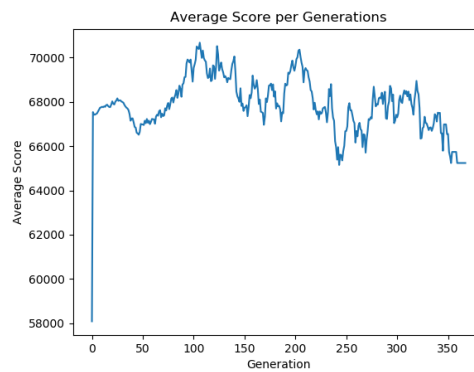


Figure 11: Average Score(2)

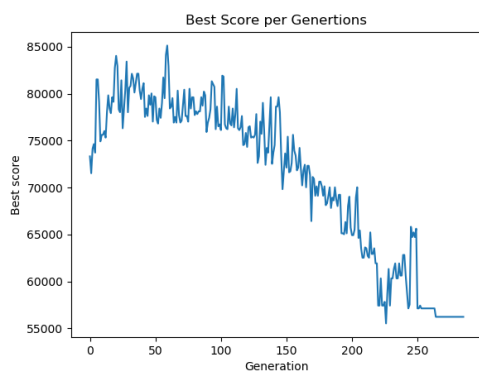


Figure 12: Best score(3)

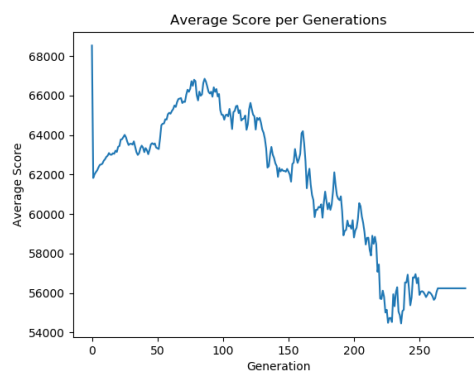


Figure 13: Average Score(3)

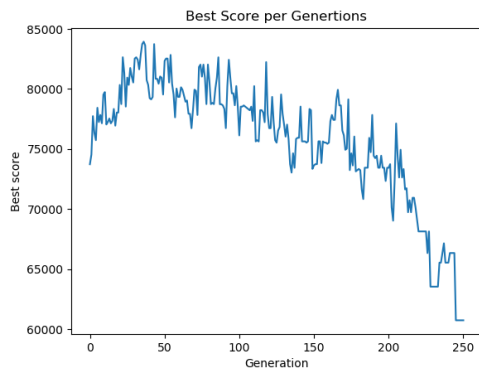


Figure 14: Best score(4)

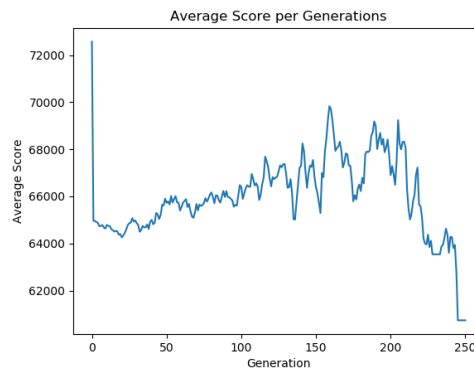


Figure 15: Average Score(4)

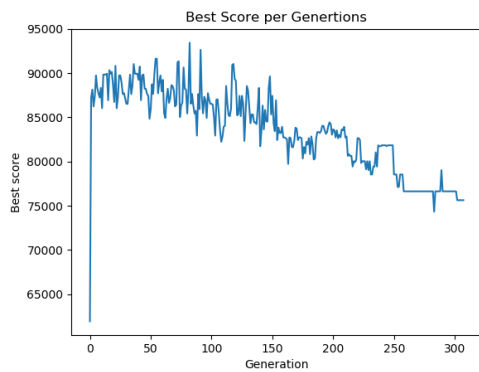


Figure 16: Best score(5)

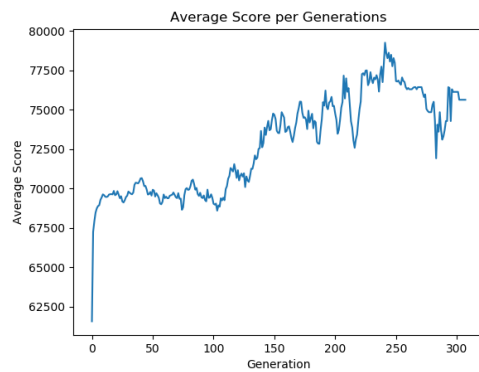


Figure 17: Average Score(5)

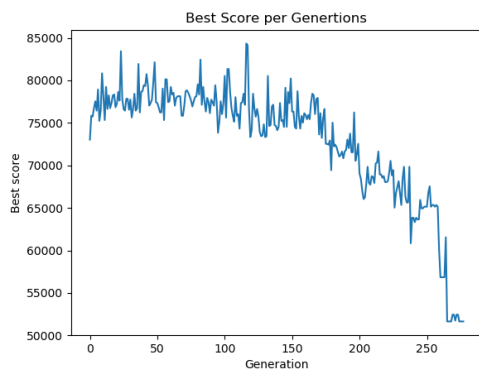


Figure 18: Best score(6)

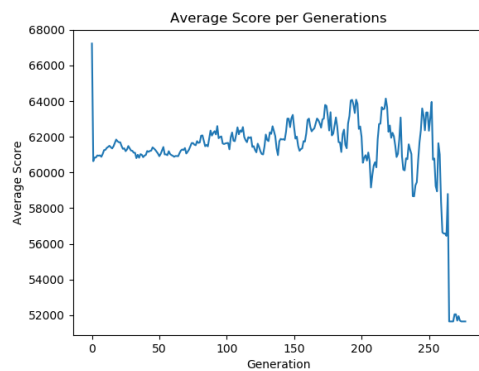


Figure 19: Average Score(6)

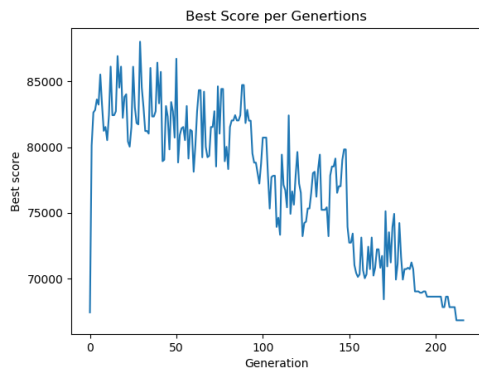


Figure 20: Best score(7)

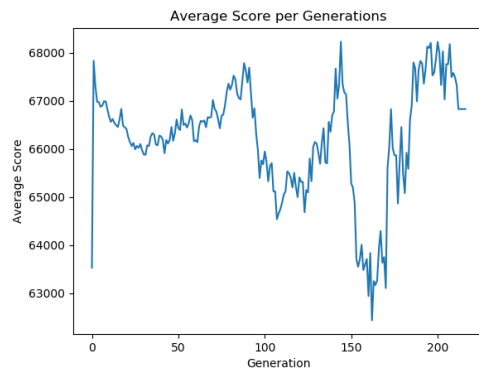


Figure 21: Average Score(7)

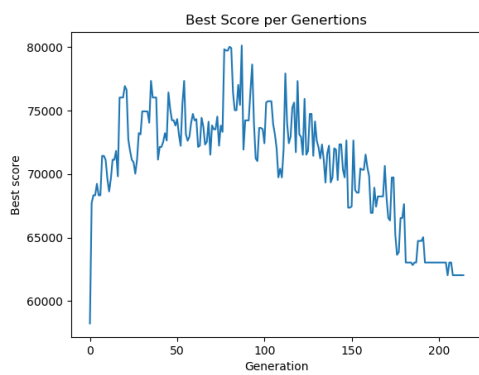


Figure 22: Best score(8)

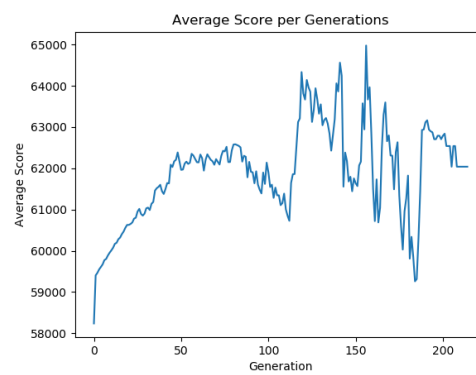


Figure 23: Average Score(8)

Θ) Σχολιασμός.

Αρχικά οι ταλαντώσεις στα γραφήματα είναι αναμενόμενες, καθώς ο πληθυσμός αλλάζει διαρκώς, με αποτέλεσμα να αλλάζει και η καλύτερη τιμή διαρκώς. Η ταλαντώσεις πάλι της μέσης τιμής είναι απολύτως δικαιολογημένες, όπως επίσης και η αύξηση κάποιες φορές των τιμών, καθώς όταν το πλήθος αλλάζει, αλλάζει και ο παρονομαστής που ζυγίζει τον μέσο όρο.

Παρατηρείται γενικά ότι η swap μετάλλαξη φέρνει καλύτερα αποτελέσματα, όταν συνδυαστεί με five point crossover σε άνω φράγμα πληθυσμού 3000. Από την άλλη, η random reseting μετάλλαξη, συνδυάζεται καλύτερα με one point crossover και πάλι άνω φράγμα πληθυσμού 3000.

Είναι αναμενόμενο, για όσο μεγαλύτερο αριθμό χρωμοσωμάτων έχουμε, τόσο καλύτερα αποτελέσματα να παίρνουμε, καθώς ο αλγόριθμος συγκλίνει πιο αργά και ομαλά στην βέλτιστη λύση.

Σε ότι αφορά τους γενετικούς τελεστές, η σύγκριση τους δεν είναι δίκαιη, αν και γενικά όπου υπάρχει swap mutation τα αποτελέσματα είναι σχετικά καλύτερα.

Για πιθανή βελτίωση του προγράμματος μας, θα μπορούσε να εξεταστεί και κάποιο άλλο είδος επιλογής ή από γενιά σε γενιά να διατηρηθεί σταθερό το πλήθος των χρωμοσωμάτων.

Η προσέγγιση μας βασίστηκε σε papers σημειώσεις και forums τα οποία βρίσκονται στην ακόλουθη ενότητα.

I) Βιβλιογραφία.

https://www.researchgate.net/post/How_to_overcome_strong_local_minima_in_Genetic_Algorithm

https://www.researchgate.net/publication/325011628_Comparative_Study_of_Different_Selection_Techniques_in_Genetic_Algorithm

<https://tik-old.ee.ethz.ch/file//6c0e384dceb283cd4301339a895b72b8/TIK-Report11.pdf>

https://www.researchgate.net/post/Which_selection_technique_is_best_in_genetic_algorithm

https://www.tutorialspoint.com/genetic_algorithms/genetic_algorithms_population.html

<https://numpy.org/>

<https://matplotlib.org/>