Si è condotto una raccolta dati sulle prime 10000 generazioni di 3 diversi run del GA con i seguenti parametri:

Fitness: energy

Crossover: singlePoint

Mutations:

0) addRandom 0.15

1) removeRandom 0.15

2) moveRandom 0.15

3) changeRandom 0.15

4) filter 0.00

Strategy size: 3 – 100

Questa raccolta dati è stata effettuata poi per ogni tipologia di metodo di selezione:

* Fitness proportional
* Linear rank con selection pressure a 1.3
* Tournament

Obiettivo dell’esperimento era osservare come variava la similitudine della popolazione durante l’evoluzione. La misura di similarità è basata sulla similarità del coseno e calcolata sul profilo di mappe generate dai punti di cambio di ogni individuo. Per ogni generazione ogni individuo è stato confrontato con il migliore della propria generazione e per ogni generazione si è tenuto traccia della media di questi valori per avere una misura delle similarità media della popolazione con il miglior individuo di essa.

Per quanto riguarda invece il Tournament selection si ha che i valori di similarità sono tutti compresi in [0.9,1.0]. La cosa è abbastanza sorprendente perché in teoria dovrebbe essere il metodo di selezione, tra i 3 scelti, che riesca a mantenere una diversità maggiore dall’individuo migliore. Questo sarà oggetto di studio per eventuali errori di implementazioni o per motivazioni che rendono inefficace questo metodo di selezione.

Per i primi due metodi di selezione si sono ottenuti risultati simili. In entrambi i casi si ha che la similitudine oscilla continuamente tra valori compresi nel range [0.6,0.95]. In entrambi i casi poi non si osservano variazioni del valore migliore di fitness in corrispondenza di cambi di trend della similitudine. Un prima conclusione che si può trarre da ciò è che il GA sia bloccato in un minimo locale e non riesca ad allontanarsi più di tanto da esso. Analizzando anche la variazione della fitness media si ha che quest’ultima rimane molto distaccata dalla minima e quindi c’è segno di una convergenza della popolazione verso una soluzione. Questo è solitamente segno di un tasso di mutazione troppo elevato. Per come i parametri del GA sono stati settati ogni gene di un individuo ha una probabilità di subire una mutazione che è 15% / N dove N è il numero di punti di cambio dell’individuo. In entrambi i casi la lunghezza media degli individui è circa 60 punti di cambio quindi la probabilità di mutazione media per gene è circa 0.25% per generazione.

Per investigare ulteriormente la diversità della popolazione si introducono nuove misurazioni di similarità tra ogni individuo e il migliore della generazione, basate su:

* Genotipo -> essendo il genoma di dimensione variabile una prima misura di similarità proposta si basa semplicemente sulla differenza del numero di punti di cambio tra l’individuo e il migliore. Si avrà così un valore positivo se vi sono più punti di cambio del migliore e un valore negativo se vi sono meno punti di cambio rispetto al migliore.
* Fenotipo -> si considera sempre la similarità del coseno, ma di essa se ne calcola l’angolo in gradi per rimuovere la non linearità del coseno che non fornisce una misura diretta di quanto due soluzioni distino tra loro nello spazio di ricerca.
* Fitness -> si considera la differenza tra la fitness di un individuo e quella migliore. Questo valore servirà per osservare eventuali correlazioni tra la variazione della popolazione e la variazione della fitness.