EC HW02 - 313551047 陳以瑄

1. **Discuss why ES and steady-state GAs form two extremes regarding the population size and the number of offspring created.**

   ES typically use a small population size (μ) and generate a relatively large number of offspring (λ), often following a relationship such as λ = 7μ. On the other hand, steady-state GAs maintain a larger population size but replace only one or a few individuals with new offspring in each iteration.

2. **Given a population of μ individuals, which are bit-strings of length L. Let the frequency of allele 1 be 0.25 at position i, that is, 25% of all individuals contains a 1, and 75% a 0 at the ith position on the chromosome. How does this allele frequency change after performing k crossover operations with one-point crossover? How does it change if uniform crossover is performed?**

   The probability that one parent has allele 1 and the other has allele 0 at position i:

   $P(one\ is\ 1\ and\ one\ is\ 0) = 2 * 0.25 * 0.75 = 0.375$

   (1) one-point crossover:

   the probability that the ith position is changed, means the selected crossover point is before i:

   $$P(crossover\ point\ before\ i) = \frac{i-1}{L-1}$$

   The probability that the ith allele is changed from 1 to 0 or from 0 to 1:

   $$P(ith\ allele\ is\ changed) = 0.375 * \frac{i-1}{L-1}$$

   The expected total change in allele frequency after k iterations is:

   $$allele\ frequency\ change = k * 0.375 * \frac{i-1}{L-1}$$

   (2) uniform crossover:

   the probability that the ith position is changed:

   $P(crossover) = 0.5$

   The probability that the ith allele is changed from 1 to 0 or from 0 to 1:

   $P(ith\ allele\ is\ changed) = 0.375 * 0.5 = 0.1875$

   The expected total change in allele frequency after k iterations is:

   $allele\ frequency\ change = k * 0.1875$

3. **Minimize the n-dimensional sphere mode**

Results for (1,1)-ES:

	σ = 0.01	σ = 0.1	σ = 1.0
Run #1	10000000	10000000	10000000
Run #2	10000000	10000000	10000000
Run #3	10000000	10000000	10000000
Run #4	10000000	10000000	10000000
Run #5	10000000	10000000	10000000
Run #6	10000000	10000000	10000000
Run #7	10000000	10000000	10000000
Run #8	10000000	10000000	10000000
Run #9	10000000	10000000	10000000
Run #10	10000000	10000000	10000000

Results for (1+1)-ES:

	σ = 0.01	σ = 0.1	σ = 1.0
Run #1	791	304378	10000000
Run #2	825	29374	10000000
Run #3	869	336347	10000000
Run #4	913	161457	10000000
Run #5	825	182346	10000000
Run #6	798	169468	10000000
Run #7	901	240310	10000000
Run #8	828	24397	10000000
Run #9	849	42442	10000000
Run #10	869	191510	10000000

4. **Compare and contrast the results you obtained in problem 3 and discuss what you think about the difference between (1,1)-ES and (1 + 1)-ES.**

   In (1,1)-ES, the offspring simply replaces the parent without comparing fitness, so there's no guarantee of improvement in each generation. This lack of selection pressure leads to random walks in the search space, preventing convergence for any value of σ.

   In (1+1)-ES, a selection step ensures that only an offspring with better fitness replaces the parent, making each generation at least as good as the previous one. The choice of σ had a noticeable impact on performance. With σ=0.01, small, steady mutations allowed effective convergence. For σ=0.1, the larger step size sped up initial progress but caused frequent overshooting, often stalling around values like 0.03. With σ=1, the step size was too large, resulting in low chances of improvement due to large jumps, leading to stagnation around 1 without reaching the target.

5. **Repeat problem 3 with uncorrelated Gaussian mutation with n step sizes.**

   I use $\tau = \frac{1}{\sqrt{2\sqrt{10}}}$, $\tau' = \frac{1}{\sqrt{20}}$, and $\varepsilon_0 = 10^{-6}$

(1,1)-ES with Uncorrelated Mutation:

	σ = 0.01	σ = 0.1	σ = 1.0
Run #1	10000000	10000000	10000000
Run #2	10000000	10000000	10000000
Run #3	10000000	10000000	10000000
Run #4	10000000	10000000	10000000
Run #5	10000000	10000000	10000000
Run #6	10000000	10000000	10000000
Run #7	10000000	10000000	10000000
Run #8	10000000	10000000	10000000
Run #9	10000000	10000000	10000000
Run #10	10000000	10000000	10000000

(1+1)-ES with Uncorrelated Mutation:

	σ = 0.01	σ = 0.1	σ = 1.0
Run #1	1928	1136	12084
Run #2	25953	1097	7027
Run #3	4924	3122	1958
Run #4	7037	4473	1024
Run #5	7497	9449	1741
Run #6	2832	2006	2179
Run #7	6861	10055	3501
Run #8	10204	3673	2184
Run #9	9480	2988	40831
Run #10	4783	2807	4741

6. **Compare and contrast the results you obtained in problems 3 and 5. Discuss what you think about the self-adaptation.**

   In (1,1)-ES with uncorrelated mutation, even with a change in mutation strategy, the lack of selective pressure still resulted in a complete failure to converge, regardless of σ.

   In (1+1)-ES with uncorrelated mutation, my choices for τ and τ ′ were not that small, allowing new step sizes to differ significantly from the old ones. In this experiment, the initial σ became less critical due to the self-adaptation process. As observed in problem 3, larger σ values could accelerate initial progress, making offspring with higher σ more likely to beat the parent in the early iterations. However, as time passed, the larger step sizes led to overshooting, giving lower σ offspring a better chance of succeeding against the parent. This created a trend of self-adaptation: increasing σ initially and then decreasing it later on. Overall, this adaptive approach performed better than using σ=0.1 or σ=1 but was not as effective as consistently using σ=0.01.

7. **Repeat problem 3 with the 1/5-rule.**

I use G = 7, a = 0.817

Results for (1,1)-ES with 1/5-Rule:

	σ = 0.01	σ = 0.1	σ = 1.0
Run #1	10000000	10000000	10000000
Run #2	10000000	10000000	10000000
Run #3	10000000	10000000	10000000
Run #4	10000000	10000000	10000000
Run #5	10000000	10000000	10000000
Run #6	10000000	10000000	10000000
Run #7	10000000	10000000	10000000
Run #8	10000000	10000000	10000000
Run #9	10000000	10000000	10000000
Run #10	10000000	10000000	10000000

Results for (1+1)-ES with 1/5-Rule:

	σ = 0.01	σ = 0.1	σ = 1.0
Run #1	333	213	274
Run #2	379	218	278
Run #3	318	244	267
Run #4	289	242	260
Run #5	308	197	246
Run #6	343	272	267
Run #7	274	252	230
Run #8	317	271	234
Run #9	251	332	306
Run #10	372	239	297

8. **Compare and contrast the results you obtained in problems 3, 5, and 7. Discuss what you think about the 1/5-rule for the self-adaptation of strategic parameters.**

In (1,1)-ES, the lack of selective pressure still prevents convergence to the target, regardless of σ.

In (1+1)-ES using the 1/5-rule, performance across all σ values improved compared to previous experiments. This enhancement is due to the self-adaptation of strategic parameters, where σ is initially increased for exploration and then decreased to avoid overshooting. This process is similar to the approach in problem 5. However, with the 1/5-rule, σ is adjusted every seven iterations in a specified direction, making it more efficient than the uncorrelated mutation, which relies solely on mutation and only changes σ when a better offspring is found. When comparing σ values, if the initial σ is too small, more iterations are needed to increase σ for better exploration. Conversely, if the initial σ is too large, it starts to decrease immediately to prevent overshooting. As a result, the iterations for σ=0.01 were higher than for σ=0.1 or σ=1.