GA2015 Project2 Report

2015-21242 박준호

**1. Abstraction**

유전 알고리즘은 선택, 교차, 변이, 세대교체를 수없이 반복하여 문제공간을 탐색하는 방법이다.이번 프로젝트에서는 순수 GA가 아닌 지역 최적화 알고리즘을 포함한 Hybrid GA를 이용하여 문제에 접근한다. 최대 600개의 도시를 가지는 Traveling Salesman Problem의 문제공간을 탐색하여 최적 해를 구하는 것이 목표이다.

**2. Introduction**

유전 알고리즘은 집단 유전학의 개체 진화 원리를 이용하는 문제공간 탐색 방법이다. 기존에 확정적으로 구할 수 없었던 NP-hard 문제들이 대표적인 유전 알고리즘이 사용되는 문제군이며, 유전학에서 사용되는 여러 현상들이 operation으로 등장한다. Selection - Crossover - Mutation - Replace 의 단순한 반복으로 좋은 품질을 가지는 해를 얻는 것이 가장 기본적인 목적이다. Traveling Salesman Problem(이하 TSP)은 위치가 명시된 도시들이 있을 때, 그 도시들을 모두 연결하는 Cycle 중에 가장 짧은 경로를 찾는 문제이다. 지난 프로젝트에서는 순수 GA의 요소만 가지고 했던 것과 달리, 이번 과제에서는 지역 최적화를 2-OPT 방식으로 제한하여 사용한 hybrid GA를 이용하여 문제공간을 탐색한다.

**3. Methodology**

먼저, 유전 알고리즘 반복구간에 진입하기 전에 초기화 과정이 필요하다. 초기화 과정에서는 정해진 개수만큼 임의로 Chromosome을 만들며, 정해진 개수를 Population이라 한다. Chromosome들의 초기값은 임의로 정해진다. 유전 알고리즘 반복구간이 종료되면, 이제까지의 해들 중 가장 품질이 좋은 해를 출력하고 프로그램을 종료한다. Representation중 Encoding의 경우, Order-based Encoding을 사용하였다.

**3.1. Selection**

품질에 비례한 룰렛 휠 선택 알고리즘은 가장 기본적이고 대표적인 선택 방법으로서, 각 해의 품질을 선택 한 후, 가장 좋은 해의 적합도가 가장 나쁜 해의 적합도의 K배(K>1)가 되도록 하여 가상의 룰렛에서 선택될 확률이 적합도에 비례하도록 선택하는 방식이다. 이전 프로젝트에서 잘못된 부분이 있어 수정하였다.

**3.2. Crossover**

TSP처럼 염색체가 순열(Permutation)로 표현 가능한 경우, 사이클 교차, 순서 교차, PMX등의 방법이 있다. 이 중 지난 프로젝트에서 사용하였던 PMX(Partially Matched Crossover)는 부모가 동일한 경우, 동일한 자식이 생성되어 수렴하는 상황에서 벗어날 수 없었다. 따라서 이전 프로젝트의 단점을 보완하고자 동일한 부모에서 다른 자식이 나오는 Order Crossover를 사용했다- 라고 생각했지만, 결국 Order Crossover 또한 동일한 부모에서 동일한 자식이 생성되어 결국 차이는 없었다.

**3.3. Mutation**

변이 연산은 부모에게 없는 속성을 자식에게 부여하여 염색체의 다양성을 야기하여 탐색 공간을 넓히는 용도로 사용된다. TSP에서는 순서가 인접한 도시가 연결된 순서만이 중요하므로, 일정 구간의 도시 이동 경로를 떼어 반대쪽에 붙여주는 의미를 가지게 되는 inversion mutation을 구현하였다. 이번 프로젝트에서 2-opt를 이 mutation을 유사하게 참고하여 쉽게 구현하였다.

**3.4. Local Optimization**

Mutation으로 생성된 자식을 문제공간 내부에서 지역 최적화 하는 단계이다. 이번 프로젝트에서 구현한 2-OPT 지역 최적화 방식은 inversion mutation과 동일하게 순열의 특정 구간을 fitness가 더 나아지지 않을 때까지 반복한다. 이렇게 되면 2-OPT가 종료된 시점의 chromosome은 기하학적으로 self-crossed되지 않은 상태가 된다. 더 이상 좋아지지 않는다는 것을 보장하기 위해 brute-force하게 모든 조합을 다 뒤집어보기 때문에 총 세대 수가 줄어들지만, 단번에 해의 품질이 좋아진다.

**3.5. Replacement**

교체는 GA가 더 나아지기 위한 마지막 관문이다. 만약 이 단계에서 교체가 이루어지지 않으면 GA는 Steady-state가 될 것이다. 이번 프로젝트에서는 Selection으로 선정된 부모로부터 Crossover, Mutation 이후 Local Optimization을 거쳐 자식을 만든 후, 이전 프로젝트와 동일하게 가장 나쁜 품질을 가진 것을 교체 대상으로 삼았다.

**4. Problem Representation**

도시의 개수는 N으로 문제의 입력으로 주어진다. 각 개체는 Chromosome과 Fitness를 멤버로 가지며, Chromosome은 order-based encoding으로, 방문하는 도시의 inex를 값으로 가지는 길이 N의 배열로 정하였다. 이번 과제에서는 N ≤ 600 이며, 각 세대는 PSIZE = 10 의 개수로 이루어진다. PSIZE가 적으면 수렴할 가능성이 높아지나, mutation으로 이를 어느 정도 극복할 수 있으며, 세대 교체 step의 시간 간격이 적어지는 이점이 있어 기존 값인 100에서 낮추었다.

**5. Result**

**5.1. Environment Specification**

사용된 CPU의 사양은 다음과 같다.

Intel® Core™ i7-3930K CPU @ 3.20GHz 3.20 GHz

**5.2. Hybrid GA Statistics**

N = 50, 100, 200, 318, 600인 TSP에 대해 20번씩 알고리즘을 돌린 결과 가장 좋은 품질, 평균 품질, 품질간의 표준편차는 다음과 같다.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N= | 50 | 100 | 200 | 318 | 600 |
| 최고 | 313.951234 | 1111.04 | 1071.71298 | 420.425351 | 1087.710253 |
| 평균 | 313.951234 | 1111.04 | 1074.319258 | 421.5984129 | 1093.167684 |
| 표준편차 | 0 | 0 | 2.252063295 | 1.060075158 | 2.809897122 |

Table 1. GA를 이용하여 TSP Solving을 20회 진행한 결과. (\* : 1E-10보다 작은 경우 0으로 처리)

**5.3. Single 2-OPT Statistics**

|  |  |
| --- | --- |
| N= | Single |
| 최고 | 1103.983 |
| 평균 | 1151.567 |
| 표준편차 | 19.89953 |

Table 2. Sigle 2-OPT를 100번 진행한 결과

**6. Discussion**

Selection, Crossover, Mutation, Replacement 네 가지 GA 연산자만을 사용하는 것이 순수 GA였다면, 이번에는 Mutation과 Replacement 사이에 Local Optimization을 끼워넣은 Hybrid GA를 하였다. 각각 연산자들 간에 시너지를 내는 경우도 있고, 서로 도움이 되지 않는 경우도 있는데, 이번에 선택한 네 가시 연산자 Roulette-Wheel Selection / Order Crossover / Inversion Mutation / 2-opt local optimization / Worst-case Replacement 중 가장 큰 영향을 미친 것은 지역 최적화였다. 사용된 TestCase의 문제공간이 이전 프로젝트보다 시간 대비 문제공간의 크기가 컸지만, 오히려 좋은 결과가 나왔다. 지역 최적화 중 제한이 걸린 경우가 이정도라면 나중에는 더 좋은 결과를 얻을 수 있지 않을까 생각한다.