

# 터치 스크린 유형 영문자판을 위한 유전자 알고리즘

강태원 김미숙<sup>u</sup> 조용만 한수경  
강릉대학교 컴퓨터공학과  
{twkang, mskim, ymcho, skhan}@kangnung.ac.kr

## Genetic Algorithm for English Keyboard of Touch Screen Style

Tae-Won Kang Mi-Suk Kim Yong-Man Cho Soo-Kyung Han  
Dept. of Computer Science, Kangnung University

### 요 약

터치스크린 상에 소프트웨어적으로 제공되는 자판은 대부분 크기가 작기 때문에 한 두 손가락만을 사용하여 글자를 입력한다. 그런데 이 자판은 양손을 모두 사용하는 경우를 전제로 하여 자소들을 배열한 것이다. 이 논문에서는 보다 빠르게 영문을 입력할 수 있는 영문 자판 배열을 유전자 알고리즘을 이용하여 찾은 것에 대하여 연구한다. 실험결과 유전자 알고리즘을 이용하여 생성한 자판이 기존의 자판에 비하여 평균적으로 약 40% 정도 빠르게 글자를 입력할 수 있다.

### 1. 연구배경

키보드를 사용하여 문자를 입력 할 때는 여러 개의 손가락을 사용해서 입력을 하지만, 휴대용 정보 기기와 같이 터치스크린 등의 입력장치를 사용할 때는 한번에 한 두 손가락을 사용할 수밖에 없다. 현재 가장 보편적으로 사용되는 영문 자판 배열은 QWERTY 자판이다. 이 자판은 다양한 기존의 타자기 자판과 비교하여 볼 때 받침을 치기 위해 윗글쇠(Shift)를 쳐야하는 등 많은 비효율적인 요소를 가지고 있지만, 컴퓨터가 타자기를 대체하면서 거의 대부분의 사람들이 이 자판에 익숙해져 있다. 또한 오늘날 쓰이고 있는 자판 배열의 성능을 평가하는 표준 방법이 없다. 그러한 이유에서 대개의 휴대용 정보 기기에서 영문 자판 배열은 QWERTY 자판 배열을 따르고 있다.

그러나, 이러한 자판은 모두 양손을 사용하는 경우를 전제로 한 것인 반면에, 휴대용 정보 기기는 크기가 작아서 자판의 크기가 작기 때문에 한번에 한 키를 입력하게 된다. 이 논문에서는 이러한 경우 가장 빠르게 글자

※이 논문은 2001년도 두뇌한국21사업에 의하여 지원되었음.

를 입력할 수 있는 영문 자판 배열을 유전자 알고리즘(Genetic Algorithm : GA)을 이용하여 찾아본다. 물론 현재의 자판에 익숙한 사용자 입장에서는 자판을 새로이 외워야 하는 문제점을 생각할 수 있으나, 휴대용 정보 기기의 자판의 작은 크기 때문에 자판이 한 눈에 들어와서 조금만 익숙해지면 외우지 않고도 효율적으로 사용할 수 있을 것이다.

이 논문의 구성은 다음과 같다. 2장에서는 터치 스크린 유형 영문 자판 배열을 위한 유전자 알고리즘에 대하여 설명하고, 3장에서는 실험을 통하여 결과를 분석하고 마지막으로 4장에서 결론을 맺는다.

### 2. 영문 자판 배열을 위한 유전자 알고리즘

영문 자판의 경우 하나의 자판 배열은 영문 자소 52자(대·소문자 구분)를 26개의 글쇠에 대응시킨 것이다. 그렇지만 대문자와 소문자를 다른 글쇠에 배정한다는 것이 좋지 않기 때문에, 대문자와 소문자를 동일한 글쇠에 배정한다면 영문 자판 배열 문제는 가능한 26!의 배열 중에서 하나를 선택하는 문제이다. 결론적으로, 영문 자판

배열문제는 경우의 수가 대단히 많은 최적화 문제에 해당 하는 것이며 이러한 문제에 유전자 알고리즘이 매우 효과적으로 응용될 수 있다.

어떤 문제에 GA를 적용하기 위해서는, 문제를 어떻게 유전형으로 표현하고, 연산자는 어떻게 구현하며, 적합도를 어떻게 계산하는가를 정해야 한다[5].

### 2.1 영문 자판 배열 개체의 표현

이 논문에서는 서술의 편의를 위하여 현재의 자판에 나타나는 영문 자소 26개와 대소문자 구분을 위한 윗글쇠(Shift), 공백문자(Space Bar), 그리고 구두점 “.” 및 “,” 30개와 현재 자판의 글쇠를 중심으로 설명한다. 즉, 영문 자소를 30개의 글쇠에 매핑하기 경우를 가지고 설명한다(<그림 3.1>).

자소 집합={ A, B, C, ..., Z, [Shift], “,”, “.”, [SpaceBar]}

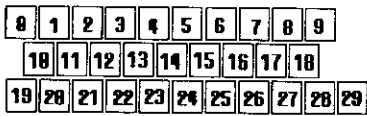


그림 3.1 자소 집합과 글쇠

이러한 전체하에 모집단 내의 각 개체는 하나의 자판 배열을 나타내는 것으로 각각 30개의 유전자를 가지며, 각 유전자는 글쇠 번호를 값으로 갖는데, 그 값은 현재 개체가 나타내는 자판에서 해당 자소의 위치를 나타낸다.

### 2.2 유전연산자 및 적합도

유전자 알고리즘을 적용할 때 사용하는 진화 연산자 즉, 교차와 돌연변이로 대표되는 유전 연산자 및 선택연산자의 종류는 매우 다양하며, 각각의 방법은 나름대로의 특징을 갖는다[3]. 이 문제에서는 순회 외판원 문제에서와 같이[2], 한 개체의 유전자 값은 유전형 내에 두 번 이상 나올 수 없다. 따라서 일반적인 교차 연산을 수행하면 안되고 부분사상교차(PMX : Partially Mapped crossover)[4]등과 같이 교차 연산 후에 새롭게 생겨난 자식 개체 역시 정당한 연산을 수행해야 한다. 돌연변이 연산을 적용하는 경우도, 어떤 방법을 사용해도 무관하지만 돌연변이 후의 결과가 항상 정당한 자판이 되는 연산자를 사용해야 한다. 임의의 두 유전자를 교환하는 치환이나 연속된 유전자 토막에서 유전자 값의 순서를 반대로 하는 전위 등의 돌연변이 연산자를 사용하면 된다. 선택연산을 위해서는 먼저 각 개체의 적합도를 평가해

야 하는데[1], 영문 자판 배열 문제에서 개체의 적합도 즉, 한 자판의 적합도는 임의의 문서 입력을 위하여 글쇠에서 글쇠로 이동한 총 이동 거리로 한다. 글쇠들 사이의 이동 거리는 임의로 정한 자판에서의 실제 거리를 측정하여 사용한다.

### 3. 실험 및 분석

실험에서는 일반적인 QWERTY 자판과, DVORAK 자판, 그리고 GA를 이용하여 만든 자판에 대해서 문서 전체의 자소를 입력하는데 필요한 전체 자간 이동 거리를 문서에 포함된 자소 개수로 나눈 평균 자간 이동 거리로 비교 분석한다. 자판을 생성하고 평가하기 위한 문서는 시, 생활문, 소설, 노래가사, 그리고 신문기사 등과 같이 5가지 성격이 다른 것을 사용하였고, 자소의 수에 따라 각각의 자판을 비교하여 보았다.

또한, GA를 이용하여 생성된 5개의 자판들에 대하여 다른 문서들에 대해서도 같은 성능을 나타내는지 알아보기 위해 50개의 문서에 대해 성능을 평가하였다.

#### 3.1 문서의 성격에 따른 성능평가

문서의 성격에 따라서, GA를 이용하여 생성된 자판의 성능이 서로 다른가를 알아내기 위하여 문서의 성격을 시(Poem), 생활문(Talk), 소설(Novel), 가사(Lyric), 그리고 신문기사(News) 등과 같이 5가지로 분류를 하였으며, 각각에 속한 문서 10개에 대하여 각각 자판 10개씩을 생성하여 문서의 성격별로 평균 자간 이동 거리의 전체 평균값을 구해서 비교하였다. 문서의 성격에 따른 실험 결과는 다음<그림 3.2>과 같다.

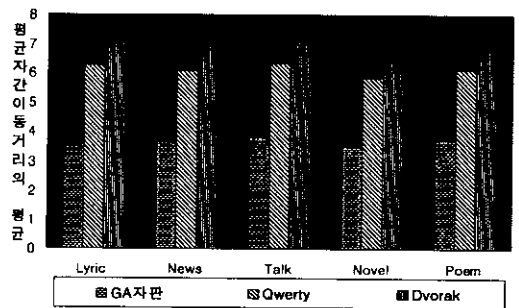


그림 3.2 문서의 성격에 따른 QWERTY, DVORAK 자판, 그리고 GA를 이용하여 생성한 자판에 대한 성능

GA를 이용하여 생성한 자판들은 그것을 생성할 때 사용한 문서의 성격과 관계없이 매우 같은 평균값을 가지며, GA를 이용하여 생성한 자판들이 평균적으로 QWERTY 및 DVORAK 자판에 비하여 40%정도의 성능

향상을 보이고 있다.

### 3.2 영문 자소의 수에 따른 성능평가

GA를 이용하여 생성한 자판들이 자소의 수에 따라 성능이 어떻게 달라지는가를 실험하여 보았다. 50개의 문서를 이용하여 생성된 50개의 자판들의 자소의 수에 따른 평균 자간 이동 거리는 <그림 3.3>과 같다.

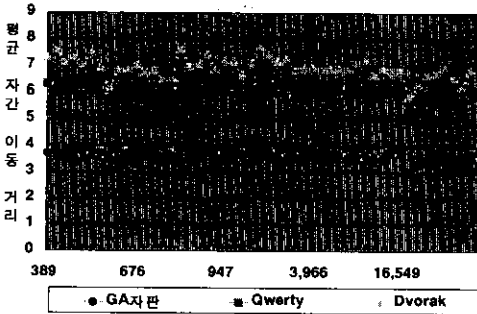


그림 3.3 자소의 수에 따른 자판별 평균 자간 이동 거리

전체적으로 자소의 수에 따른 평균 자간 이동 거리값의 변화는 거의 같은 패턴을 보여주고 있으며, 자소 수에 관계없이 평균 자간 이동 거리 역시 거의 같은 값을 갖는다는 것을 알 수 있다. 여기에서도 GA를 이용하여 생성한 자판이 QWERTY, DVORAK 자판에 비해서 평균적으로 40% 정도 문서를 빠르게 입력할 수 있다.

### 3.3 GA를 이용하여 생성한 자판에 대한 성능평가

각각의 문서를 이용하여 생성한 자판들에 대하여, 다른 문서를 입력 할 때의 성능을 평가한다. GA를 이용하여 생성한 50개의 자판 중에 문서 성격별로 가장 우수한 5개의 자판(자판1, 자판2, 자판3, 자판4, 자판5)을 선택하여 50개의 문서 전체에 대한 평균 자간 이동 거리의 평균을 구하여 비교한다. 예를 들어 GA를 이용하여 생성한 자판1의 자판 배열을 보면 <그림 3.4>와 같다.

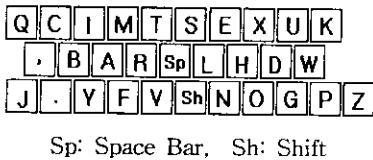


그림 3.4 GA를 이용하여 생성한 자판 배열

실험 결과는 다음<그림 3.5>과 같다.

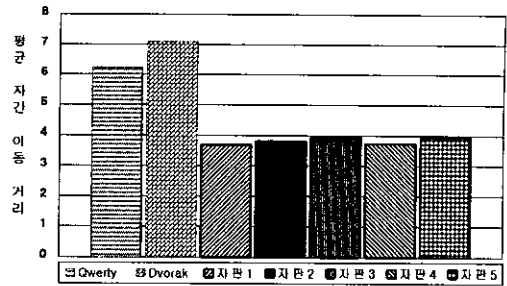


그림 3.5 50개의 문서에 대한 평균 자간 이동 거리의 평균

결론적으로 GA를 이용하여 생성한 자판들이 QWERTY, DVORAK 자판에 비해서 40%정도의 성능 향상을 보이고 있다.

### 4. 결론

터치스크린 상에 소프트웨어적으로 제공되는 자판은 대부분 크기가 작기 때문에 한 두 손가락만을 사용하여 글자를 입력한다. 이 논문에서는 보다 빠르게 영문을 입력할 수 있는 영문 자판 배열을 유전자 알고리즘을 이용하여 찾는 것에 대하여 연구하였다.

실험결과, 우리가 일반적으로 사용하고 있는 QWERTY, DVORAK 자판에 비해서 문서의 성격이나 글자의 수에 관계없이 GA를 이용하여 생성한 자판들이 평균적으로 40%정도의 성능 향상을 보이고 있다. 또한, 특정한 문서를 이용하여 생성된 자판들로 다른 문서를 입력하는 경우도 QWERTY, DVORAK 자판보다 높은 성능을 보이고 있음을 알 수 있다.

### 5. 참고문헌

- [1]. David B.Fogel, Lawrence J.Fogel, "An Introduction to Evolutionary Programming", Natural Selection Inc.
- [2]. Goldberg.D.E. and Lige, R.Jr., "Alleles, Loci and the Traveling Salesman Problem", Proc.1st. ICGA, 1985.
- [3]. Hancock, Selection Methods for Evolutionary Algorithms, in Practical Handbook of Genetic Algorithms Vol.2, pp.67-92, 1995.
- [4]. Jose L.Ribeiro Filho, Philip C.Treleaven Cesare Alippi, "Genetic - Algorithm Programming Environments", IEEE Computer, pp.28-43, June1994.
- [5]. Melanie Mitchell, Stephanie Forrest, "Genetic Algorithms and Artificial Life", The MIT Press Cambridge, Massachusetts London, England, pp.267-289, 1995.