

유전자 알고리즘

김 중 훈 (고려대학교 토목환경공학과 부교수)

Q 유전자 알고리즘이란 무엇입니까?

A 유전자 알고리즘 (Genetic Algorithm, GA)은 진화론의 적자생존(survival of the fittest)과 자연도태(natural selection)의 유전학에 근거한 적응탐색(adaptive search)기법으로서 발견적 해법이다. 유전자 알고리즘은 1973년 Holland에 의해 많은 발전을 이루었으며, 여러가지 어려운 조합문제(combinatorial optimization)에 대한 효율적인 탐색을 수행하고, 최적해에 근사한 해를 쉽게 구할 수 있는 방법이다.

해는 자연 유전체계에서 염색체(chromosome)와 유사한 형태로 암호화되어 나타내지며 목적함수의 염색체를 평가함으로써 각 염색체의 적합도를 알 수 있고, 이 값에 따라 자손생성과 다음 세대에 살아남을 가능성이 결정된다. 각 염색체는 유전인자(gene)의 이름(string)으로 이루어져 있으며 유전인자는 전체 염색체상의 특정한 위치(locus)에 특정한 유전형질(allele)을 가지고 존재한다. 그러므로 GA에서 염색체는 0과 1의 이름으로 구성된 이진암호(binary code)로 표현되며, 정수나 실수도 쉽

게 사용할 수 있다.

유전자 알고리즘을 적용하기 위한 초기해집단은 염색체집단을 임의적으로 발생시켜 형성하고, 각 염색체의 적합성을 표현할 수 있는 평가함수를 설정하여 유전해 집단을 평가한 후 적합도 함수로 변환한다. 다음 세대를 위해 적자생존의 원리에 따라 부모해로부터 염색체가 확률적으로 선택된다. 다음 세대를 생성하기 위해 부모해의 선택(selection), 상호교배(crossover)와 돌연변이(mutation) 유전연산, 복제(reproduction), 교체(replacement)의 유전자 재조합을

통해 새로운 유전자 집단을 형성하는 세대교체를 반복적으로 수행한다. 부모집단의 선택비율에 적합도 함수를 사용하며 부모는 자손발생에 돌연변이와 상호교배 유전연산자를 사용한다. 이 과정은 다음 세대에 우수한 유전자와 우성의 염색체 조합과정으로서 다음 세대에서 새로운 유전해집합을 생성한다. 상호교배는 두 부모해의 암호를 임의의 위치에서 부분적으로 상호 교환하여 재조합함으로써 새로운 자손해를 생성하는 주된 연산자이다. 돌연변이는 염색체내에 새로운 유전자를 도입하여 집단의 다양성을 보존한다. 교체유전자는 해집단의 가장 나쁜 염색체와 나쁜 적합도를 갖는 염색체를 더 나은 값을 갖는 염색체로 교체하는 것이다. 교체과정은 미리 설정한 종료조건(세대교체횟수)에 도달하거나 해의 개선이 관찰되지 않을때(converge)까지로 설정하고 있다. 이상과 같은 절차는 그림 1과 같다.

유전자 알고리즘의 기본개념은 간

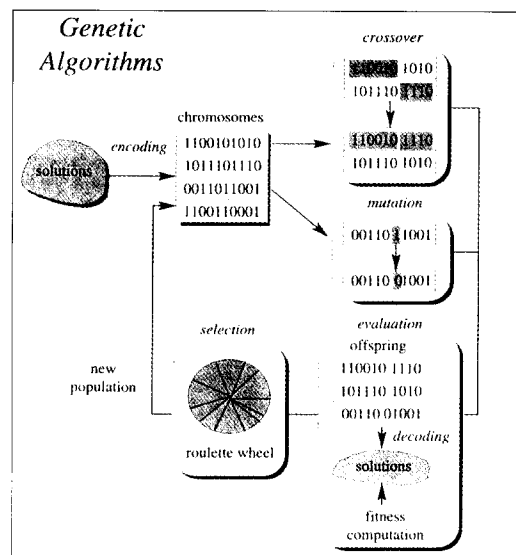


그림 1. 유전자 알고리즘의 일반적인 구조

단하고 매우 쉬우며 여러가지 방법으로 수행할 수 있다. 상호교배를 위한 염색체 결정에는 여러가지 기법이 있으며, 염색체 암호특성과 문제형태에 기초를 두고 최근 상호교배에 대한 연구가 많이 발표되었다. 한 염색체의 약간의 임의적인 변화를 소개하는 돌연변이에 대한 기법도 많이 소개되었다. 특정한 문제를 해결하기 위하여 유전자 알고리즘을 개발하는 과정은 위에서 살펴본 각종 기법을 결합하여 종합화하는 작업으로 해의 표현 방법, 초기해 발생방법, 각 해의 적합도 및 평가기법, 부모해의 선택규칙, 유전자 연산, 교배정책, 종료조건 등이 포함된다.

Q 유전자 알고리즘은 수자원 공학에 어떻게 적용됩니까?

A 유전자 알고리즘은 최적화 문제를 푸는 알고리즘의 일종으로 볼 수도 있으며 그러기에 최적화기법이 고려되는 수자원공학의 다양한 분야에 적용될 수 있다. 관망, 댐, 모형변수 등의 분야를 들 수 있는데 특히 관망에서 많은 적용예를 볼 수 있다.

관망분야에서는 1) 관로설계·계획, 2) 운영, 3) 노선, 4) 요소의 규모 등의 문제로 나뉜다.

관로설계 및 계획 분야에는 관로 신설, 기존관로 확장, 노후관로 개량 등으로 다시 나뉘는데, 관로 신설의 경우 이미 상당수의 관로가 존재하고 있는 선진국보다는 새로이 많은 개발이 필요한 개발도상국에서 각광을 받는 분야다. 실제로 미국의 코비지역의 관개용 수로에 적용되어 좋은 결과를 제시했으며 여러가지 다양한 대

안도 함께 제시 할 수 있었다. 기존관로의 확장에도 적용되었는데 뉴욕시의 경우 그간 많은 최적화 해법의 결과보다도 많게는 2배까지 유전자 알고리즘이 최적의 비용을 도출하는 결과를 보였다. 노후관로의 개량은 매설년수가 경과함에 따라 수질악화와 누수문제를 해결하기 위하여 가장 저렴한 비용으로 최대의 효과를 거두어야 하는 문제인데 이를 위해 관갱생, 관교체, 관삽입, 관청소, 관라이닝 등의 방법중 가장 좋은 방법을 가장 적절한 시기에 사용할 수 있도록 최적화 기법으로 방법과 시기를 결정해야 하며 유전자 알고리즘중에서도 여러개의 목적함수를 고려해 주는 다목적해법 SMGA(Structured Messy Genetic Algorithm)가 사용되어 실제로 모로코의 한 도시에 적용한 결과 좋은 결과가 도출되었다.

운영 분야는 펌프의 스케줄, 운영시점 결정, 비상시 계획, 소화전의 위치와 규모 등의 문제로 나뉜다. 펌프의 스케줄은 최소의 비용으로 펌프를 운영하기 위하여 어떤 펌프를 몇번이나 가동하는 지에 대한 문제이다. 예를 들어 펌프의 가동여부에 따라 컷을 때를 1, 깎을 때를 0으로 하고 하루를 24시간으로 나누면 시간당의 펌프 운영조건은 24개의 비트를 가진 염색체로 구성할 수 있게 된다. 운영점 결정은 수조, 펌프, 압력밸브 등의 운영시점과 종료시점을 결정하는 문제이다. 비상시 계획은 송배수관로의 간선에서 사고가 발생했다든지 수리부속물의 파괴, 단전으로 인한 펌프의 정지 등의 문제이다.

노선분야는 새로이 관로를 신설하는 문제나 기존관에서 덧붙이는 문

제, 수조·펌프장·감압밸브 등의 위치선정문제이다. 특히 노선을 찾아가는 문제에 있어서 기존의 DP모형과 비교해볼때 유전자 알고리즘은 빠른 결과를 도출하고 작업메모리도 줄일 수 있다.

요소의 규모는 신관의 관경과 재질 결정, 펌프장·저류조의 크기 등을 결정하는 문제이다.관망분야에 유전자 알고리즘을 도입하여 얻는 장점을 요약한다면 첫째로 기존방법보다 15~30%정도 경제적이며, 기존의 방법이 하나의 해를 제시하는 대신 유전자 알고리즘은 최적의 비용 근처에 있는 많은 대안의 해를 제시해준다. 또한 수리학적인 제약조건에 대해서는 수리시뮬레이터(KYPIPE, EPANET, CYBERNET 등)와 연계하여 가능해영역을 수시로 검사해준다. 그리고 유전자 알고리즘의 계산 결과로 제시되는 수치들이 연속적인 수치가 아닌 이산적인 규격의 값으로 제시되는데 예를들어 관의 규격관경이나 펌프의 규격용량, 수조의 규격용량을 들 수 있다.

유전자 알고리즘은 댐이나 모형변수의 검정에도 적용된다. 댐 분야는 한 수계에 여러 댐이 존재할 경우 연계운영문제나 여러수원에서 경제성이나 수질조건을 만족시키며 적절히 분배하여 취수하는 문제이다. 모형변수의 검정은 저류함수모형 매개변수나 관 조도계수 등의 최적값을 결정하는 문제이다. 특히 국내에서도 대청댐 유역의 과거 홍수자료를 이용하여 저류함수 모형의 변수인 침투홍수량과 도달시간의 최적치를 검정하는 연구가 발표되기도 하였다.