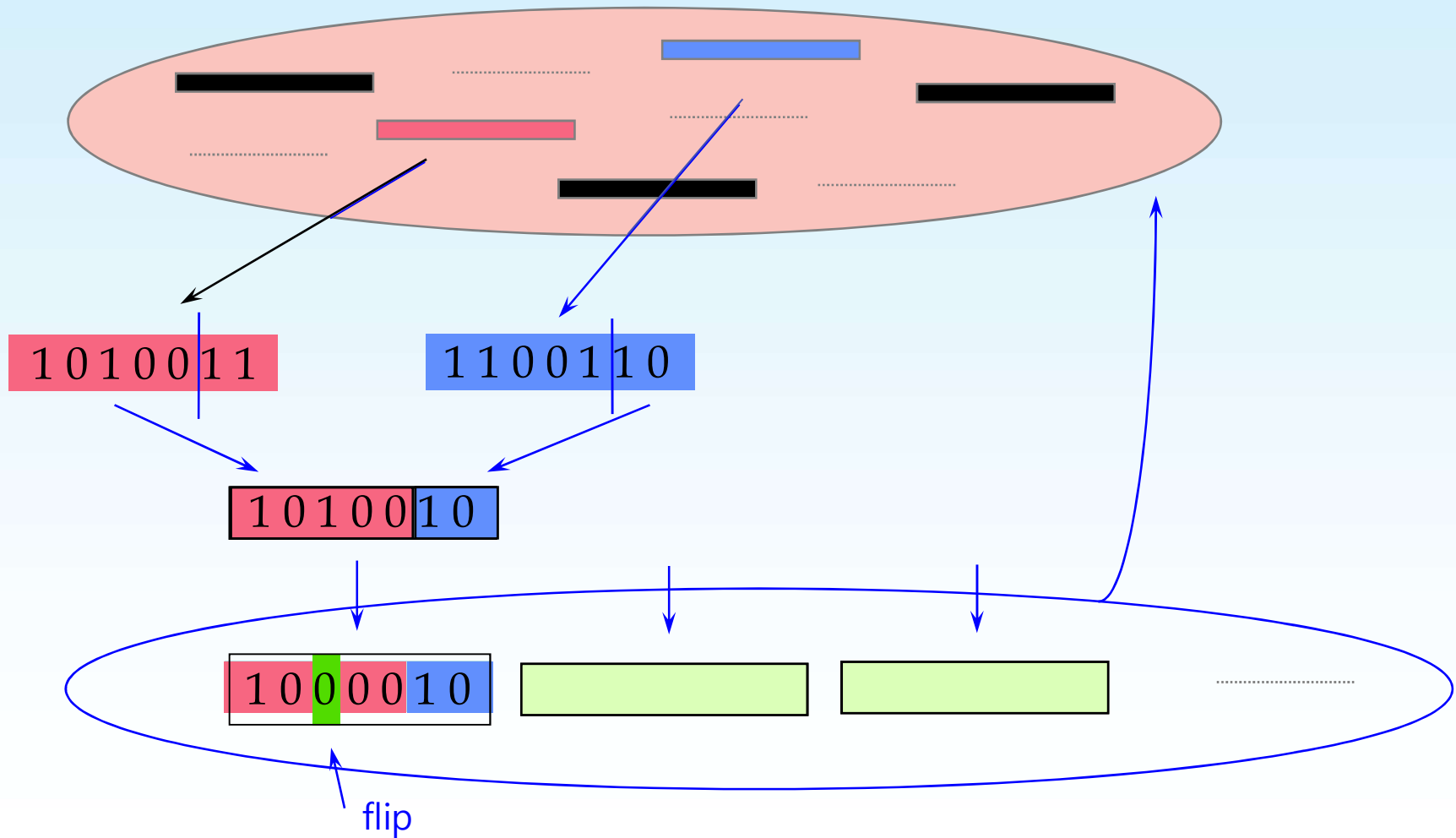


해의 표현

A Typical GA Flow



GA Template

Generate initial solutions (say, p solutions);

do {

for ($i=0$; $i < k$; $i++$){

 Select two parent solutions;

 Crossover;

 Mutation;

 }

 Replace;

} **while** (not stopping condition);

Report the best solution;

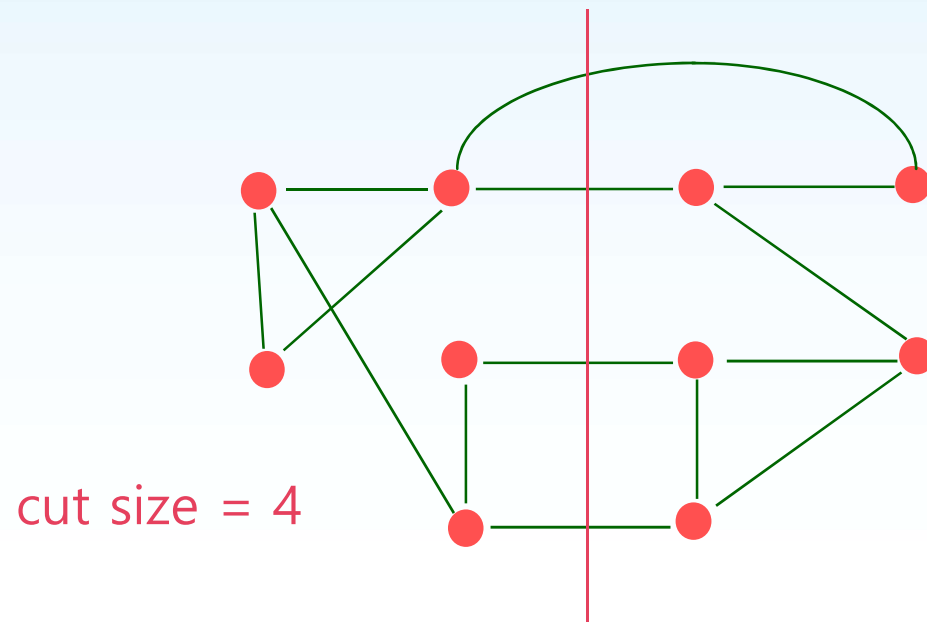
k/p : generation gap

해의 표현

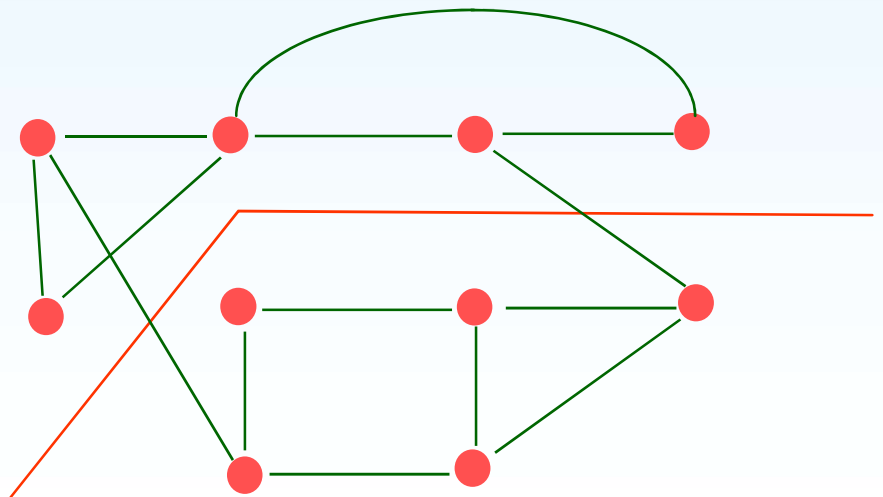
- 우선 GA가 이해할 수 있는 형태로 문제의 해를 표현해야 한다
 - 1차원 : 2차원
 - 위치기반 : 순서기반
 - Binary : Decimal
 - Integer : Real
 - ...

표현의 예

- Bisection
 - graph를 같은 사이즈의 두 subgraph로 분할한 것
- Cut size
 - 해당 bisection의 cross edge 갯수

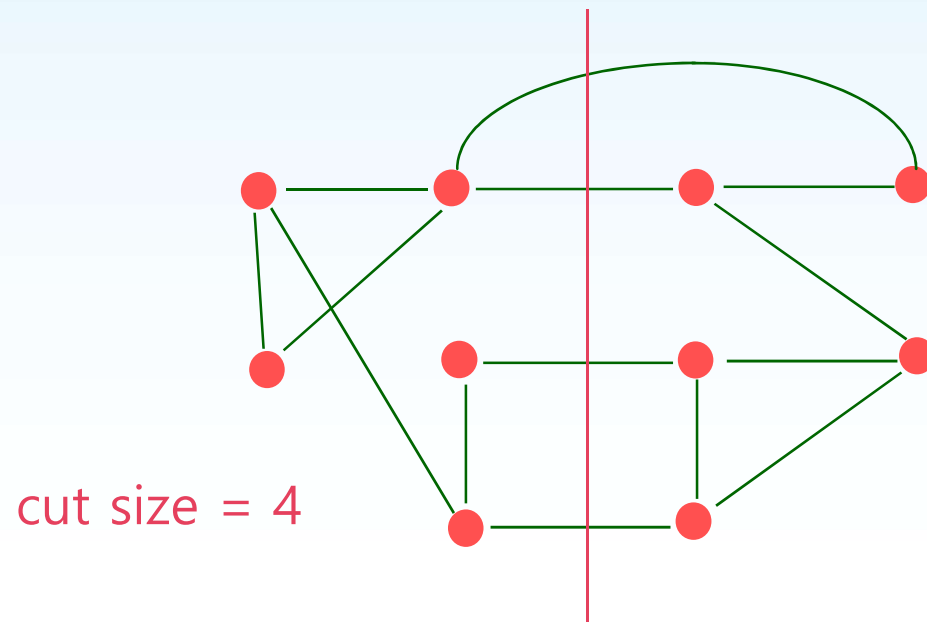


cut size = 2

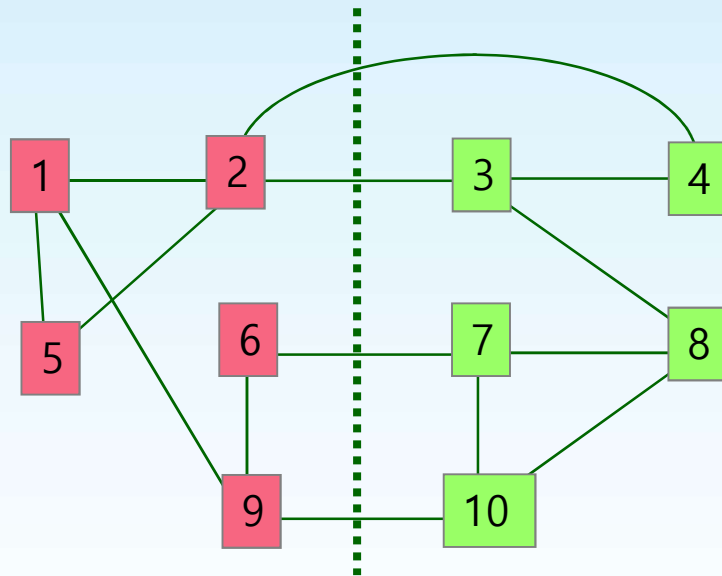


Binary : k-ary

- Bisection
 - graph를 같은 사이즈의 두 subgraph로 분할한 것
- Cut size
 - 해당 bisection의 cross edge 갯수



Binary Encoding for Graph Bisection

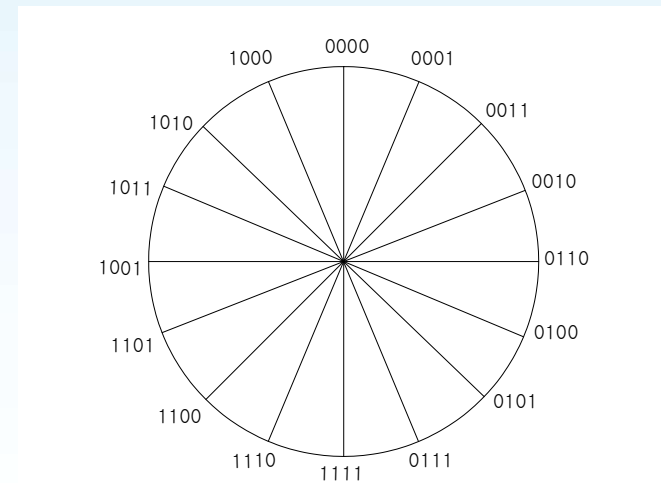


0 0 1 1 0 0 1 1 0 1

일반 이진수 : Gray Coding

- Gray coding
 - 인접한 수는 한 한 비트만 차이나도록 만든 이진수 코딩 체계

Binary	Gray
0000	0000
0001	0001
0010	0011
0011	0010
0100	0110
0101	0100
0110	0101
0111	0111
1000	1111
1001	1110
1010	1100
1011	1101
1100	1001
1101	1011
1110	1010
1111	1000

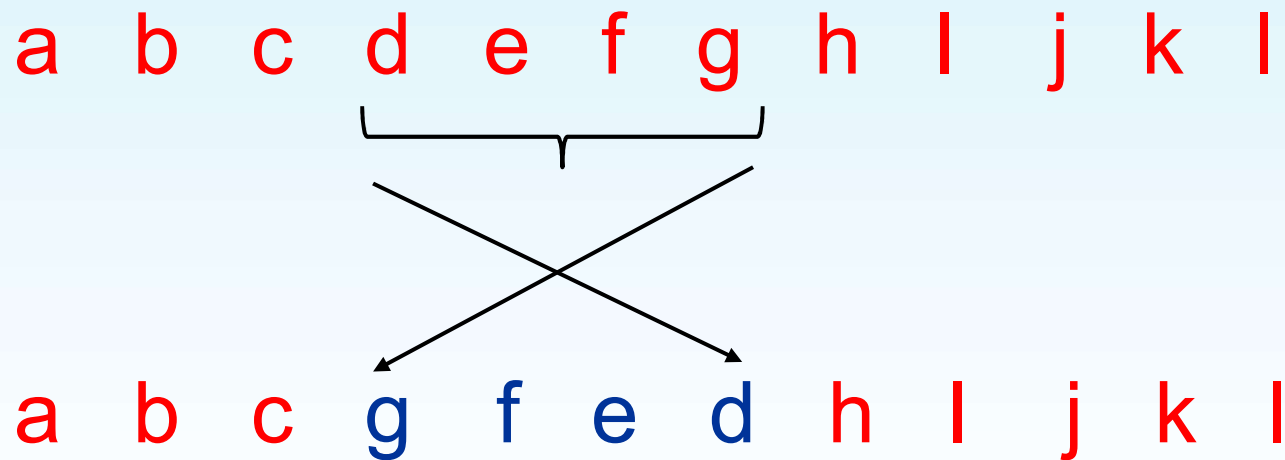


실수 표현

- 독일의 Evolutionary Strategy 그룹에서 처음 사용
 - Single thread 진화 방법이므로 태생적으로 crossover가 없었음
 - 굳이 이진 표현을 쓸 필요가 없었음
- 최초 사용
 - Bremermann, 1962
- 인자의 속성이 실수이면 인자와 유전자가 1:1 대응되어 자연스러움
- 유용한 예
 - 산술 교차(arithmetic crossover)

가변 표현

역치|inversion



schema의 모양에 영향을 미친다

Messy GA

(1, 1)(2, 0)(2, 1)(4, 0)(1, 0)

의미:

위치 1은 1, 위치 2는 0, 위치 2는 1, 위치 4는 0, 위치 1은 0

10?0

Overspecification

Underspecification



- **Overspecification**
 - 별 문제 안됨
 - 뒤에 나타나는 것 무시해 버린다
- **Underspecification**
 - 어떻게든 품질 평가를 해야 한다
 - 결손된 유전자 위치에 임의로 값을 채워 평가
 - 대표성을 갖기 힘들다
 - 결손된 유전자 자리들에 나름의 최적화해서 평가

Genetic Programming (GP)

- 트리 형태의 chromosome
- 원래 LISP 프로그램을 자동으로 만들기 위해 고안
- John Koza...

위치 기반 : 순서 기반

- 위치 기반 표현
 - 각 유전자의 위치가 의미를 갖는다
- 순서 기반 표현
 - 유전자들의 순서만이 의미를 갖는다
 - 유전자의 절대 위치는 의미가 없다

순서 기반

예: TSP의 순서 기반 표현

0 1 3 4 8 5 7 9 6 2

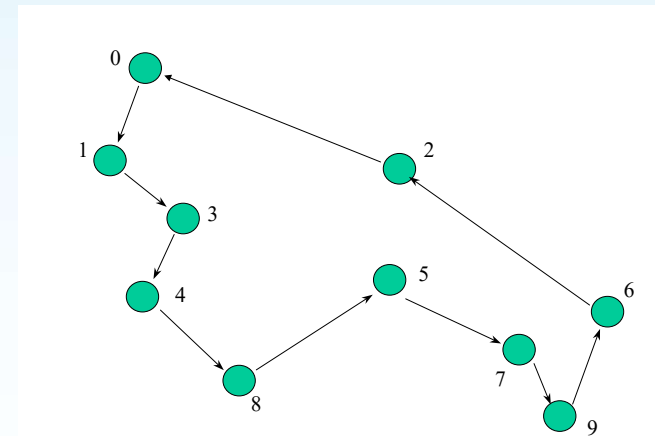
같은 의미가 다른 chromosome으로 표현될 수도 있다

1 3 4 8 5 7 9 6 2 0

3 4 8 5 7 9 6 2 0 1

4 8 5 7 9 6 2 0 1 3

...



이런 특성은 GA의 crossover에 혼란을 줄 수 있다

위치 기반

예: TSP의 위치 기반 표현

0 1 2 3 4 5 6 7 8 9

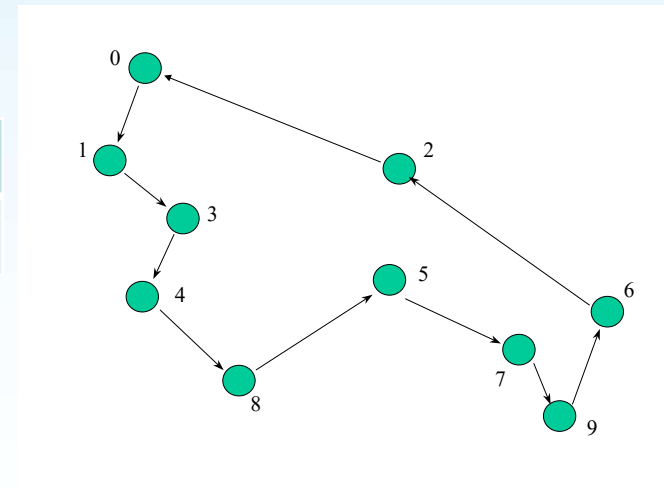
0 1 3 4 8 5 7 9 6 2

같은 의미는 하나의 chromosome으로

위치	0	1	2	3	4	5	6	7	8	9
값	1	3	0	4	8	7	2	9	5	6

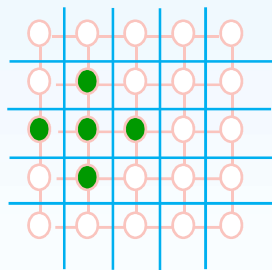
의미

Vertex 4 다음에 오는 vertex는 8이다

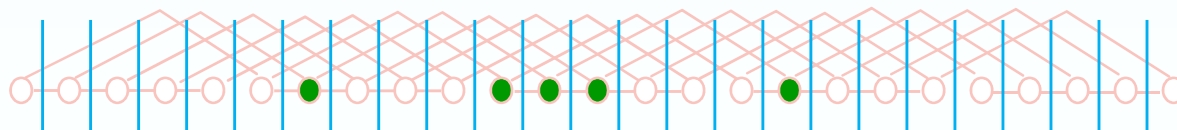


Encoding Dimensions

- A fairly high dimension is needed to embed a general graph w/ reasonable distortion (Linial-London-Ravinovich, 1994)
- Linear encoding : N -dimensional encoding

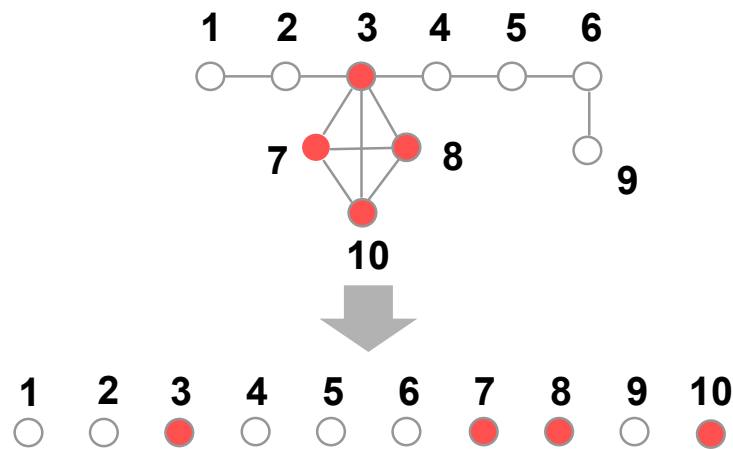


$2 (\sqrt{n} - 1)$ cuts



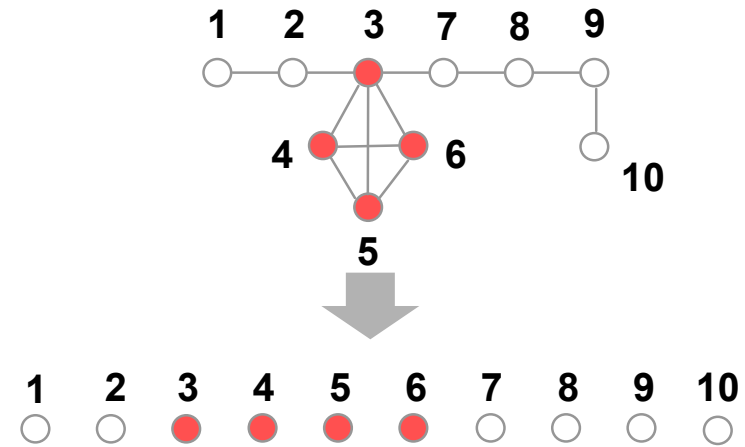
$n-1$ cuts

Reordering Genes



schema의 defining length 7

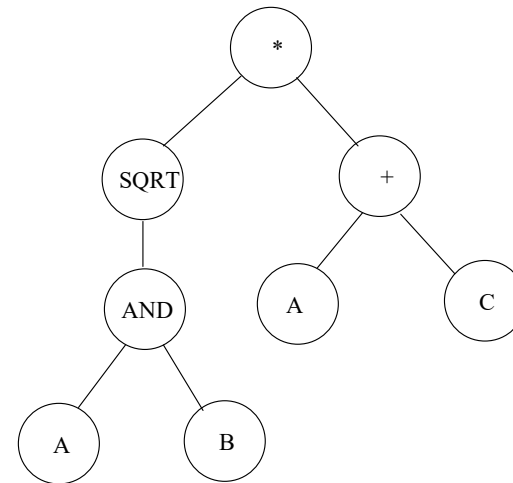
BFS



schema의 defining length 3

Tree 표현

$(* (\text{SQRT} (\text{AND } A \ B)) (+ \ A \ C))$



Linear chromosome의 관행을 깨 한 갈래

John Koza가 제안

LISP 프로그램이 트리 형태의 S-expression으로 표현

“프로그램이 프로그램을 짤다”