

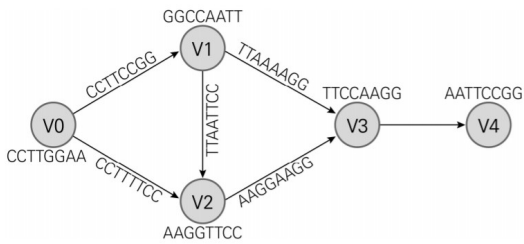
한 가닥의 DNA는 아데닌(A), 구아닌(G), 시토신(C), 티민(T)의 네 종류의 염기를 가지고 있는 뉴클레오티드가 선형적으로 이어진 사슬로 볼 수 있다. 보통의 경우 <그림 1>과 같이 두 가닥의 DNA가 염기들 간 수소 결합으로 서로 붙어 있는 상태로 존재하는데, 이를 ‘이중나선 구조’라 부른다. 이때 A는 T와, G는 C와 상보적으로 결합한다. 온도를 높이면 두 가닥 사이의 결합이 끊어져서 각각 한 가닥으로 된다.



<그림 1> 염기들 간 상보적 결합의 예

정보과학의 관점에서는 DNA도 정보를 표현하는 수단으로 볼 수 있다. 한 가닥의 DNA 염기서열을 4진 코드로 이루어진 특정 정보로 해석할 수 있기 때문이다. 즉, ‘A’, ‘G’, ‘C’, ‘T’만을 써서 순서가 정해진 연속된 n개의 빈칸을 채울 때, 총 4ⁿ개의 정보를 표현할 수 있고 이중 특정 연속체를 한 가지 정보로 해석할 수 있다.

DNA로 정보를 표현한 후, DNA 분자들 간 화학 반응을 이용하면 연산도 가능하다. 1994년 미군의 정보과학자 에이들먼은 『사이언스』에 DNA를 이용한 연산에 대한 논문을 발표했다. 이로써 ‘DNA 컴퓨팅’이라는 분야가 열리게 되었다. 이 논문에서 에이들먼이 해결한 것은 정점(예: 도시)과 간선(예: 도시 간 도로)으로 이루어진 그래프에서 시작 정점과 도착 정점이 주어졌을 때 모든 정점을 한 번씩만 지나는 경로를 찾는 문제, 즉 ‘해밀턴 경로 문제(HPP)’였다. HPP는 정점의 수가 많아질수록 가능한 경로의 수가 급격하게 증가하기 때문에 소위 ‘어려운 문제’에 속한다.



<그림 2> 정점 5개로 구성된 그래프

DNA 컴퓨팅의 기본 전략은, 주어진 문제를 DNA를 써서 나타내고 이를 이용한 화학 반응을 수행하여 답의 가능성이 있는 모든 후보를 생성한 후, 생화학적인 실험 기법을 사용하여 문제 조건을 만족하는 답을 찾아내는 것이다. 에이들먼이 HPP를 해결한 방법을 <그림 2>의 그래프를 통해 단순화하여 설명하면 다음과 같다. <그림 2>는 V0가 시작의 정점, V4가 도착 정점이고 화살표로 간선의 방향을 표시한 그래프를 보여 준다. 즉, V0에서 V1로 갈 수 있으나 역방향으로는 갈 수 없다. 먼저 그래프의 각 정점을 8개의 염기로 이루어진 한 가닥 DNA 염기서열로 표현한다. 그리고 각 간선을 그 간선이 연결하는 정점의 염기서열로부터 취하여 표현한다. 즉, V0(<CCTTGGAA>)에서 출발하여 V1(<GGCCAATT>)

에 도달하는 간선의 경우는 V0의 뒤쪽 절반과 V1의 앞쪽 절반을 이어 붙인 염기서열 <GGAAGGCC>의 상보적 코드 <CCTTCCGG>로 나타낸다. 이렇게 6개의 간선 각각을 DNA 코드로 표현한다.

이제 DNA 합성 기술을 사용하여 이들 코드를 종류별로 다량 합성한다. 이들을 하나의 시험관에 넣고 서로 반응을 시키면 DNA 가닥의 상보적 결합에 의한 이중나선이 형성되는데, 이것을 ‘혼성화 반응(hybridization)’이라 한다. 혼성화 반응의 결과로 경로, 즉 정점들의 연속체가 생성된다. 시험관 안에는 코드별로 막대한 수의 DNA 분자들이 있기 때문에, 이들 사이의 이러한 상호 작용은 대규모로 일어난다. ① 이상적인 실험을 가정한다면, 혼성화 반응을 통해 <그림 2> 그래프의 가능한 모든 경로에 대응하는 DNA 분자들이 생성된다. 경로의 예로 (V0, V1), (V1, V2), (V0, V1, V2) 등이 있다. 이와 같이 생성된 경로들로부터 해밀턴 경로를 찾아 나간다.

- [1단계] V0에서 시작하고 V4에서 끝나는지 검사한 후, 그렇지 않은 경로는 제거한다.
- [2단계] 경로에 포함된 정점의 개수가 5인지 검사한 후, 그렇지 않은 경로는 제거한다.
- [3단계] 경로에 모든 정점이 포함되었는지 검사한다.
- [4단계] 지금까지의 과정을 통해 취한 경로들이 문제에 대한 답이라고 결정한다.

에이들먼은 각 단계를 적절한 분자생물학 기법으로 구현했다. 그런데 DNA 분자들 간 화학 반응은 시험관 내에서 한꺼번에 순차적으로 일어난다는 특성을 갖고 있다. 요컨대 에이들먼은 기존 컴퓨터의 순차적 연산 방식과는 달리, 대규모 병렬 처리 방식을 통해 HPP의 해결 방법을 제시한 것이다. 이로써 DNA 컴퓨팅은 기존의 소프트웨어 알고리즘이나 하드웨어 기술로는 불가능했던 문제들의 해결에 대한 잠재적인 가능성을 보여주었다.

01

DNA 컴퓨팅에 대한 설명으로 적절하지 않은 것은?

- ① 창시자는 미국의 정보과학자 에이들먼이다.
- ② DNA로 정보를 표현하고 이를 이용하여 연산을 하는 것이다.
- ③ 기본적인 해법은 가능한 모든 경우의 수를 생성한 후, 여기서 답이 되는 것만을 찾아내는 것이다.
- ④ 기존 컴퓨터 기술의 발상을 전환하여 분자생물학적인 방법으로 접근함으로써 정보 처리 방식의 개선을 모색했다.
- ⑤ DNA 컴퓨팅을 이용하여 HPP를 풀 때, 간선을 나타내는 DNA의 염기 개수는 정점을 나타내는 DNA의 염기 개수의 두 배다.

