ACM-ICPC 문제의 출제 및 채점 과정

전남대학교 ▌임 형 석*

1. 개 요

매년 열리는 대학생 프로그래밍 경시대회인 ACM-ICPC가 점차 대학생들의 프로그래밍 능력을 시험하는 대표적인 무대로서 자리매김을 하고 있다. 국내에서는 매년 서울에서 ACM-ICPC 아시아 지역 예선[1]이 열리고 있으며 대회에 참여하는 대학교 및 학생 수가 해마다 증가하는 추세이다.

본고에서는 국내에서 열리는 ACM-ICPC 아시아 지역 예선 문제들이 어떠한 과정을 거쳐 출제되고 채점되는지를 소개한다. 그리고 출제자가 프로그래밍 학습에 대한 동기 부여와 완벽한 프로그래밍을 유도하기 위해 고려하는 사항들을 실제로 출제되었던 문제와 함께 살펴본다.

최근 인터넷 예선이 도입됨에 따라 경시대회 문제들은 크게 인터넷 예선 문제와 본선 문제로 구분된다. 본고에서는 본선 문제에 초점을 두고 세부 과정들을 설명한다. 경시대회 문제를 만들고 경시 당일 실시간으로 채점하기 위해서는 그림 1과 같은 여러 단계들을 거치게 된다.

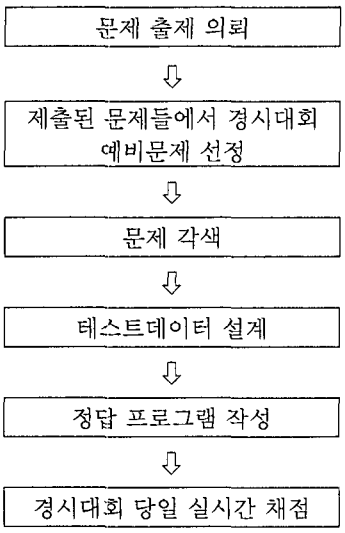

그림 1 출제 및 채점 절차

* 종신회원

본고에서는 그림 1에 나타난 절차를 따라 내용을 전개한다. 제 2절에서는 문제 출제 과정의 단계들을 살펴보고 제 3절에서는 채점 과정의 단계들을 살펴본다. 제 4절에서는 요약과 함께 교육적 효과에 대해 기술한다.

2. 문제 출제 과정

여기에서는 문제를 모집하고 대회용 문제를 선정한 후 최종적인 영문 문제를 만들기까지의 과정을 기존에 출제된 문제와 함께 설명한다.

2.1 문제 모집

경시대회에 사용할 문제는 국내 대학에 있는 알고리즘 전공 교수들에게 출제를 의뢰한다. 출제에는 문제내용과 함께 해결방법과 난이도를 포함하도록 한다. 이때 교과서에 들어 있는 문제나 기존의 여러 경시대회들에서 사용된 문제들은 사용할 수 없다.

이 단계에서 수합된 문제들은 보통 서술이 완벽하지는 않고 문제의 요점과 개략적인 출제 의도를 담고 있다. 그림 2에 문제 모집 단계에서 수합된 예비문제의 예가 있다.

> 순열매칭 문제
>
> 1부터 k까지의 숫자로 구성된 어떤 순열 $P = (p_1, p_2, \ldots, p_k)$와 1부터 n까지의 숫자로 구성된 어떤 순열 $T = (t_1, t_2, \ldots, t_n)$가 있다($k \leq n$). T에서 P의 순열 매칭이란 P와 같은 상대적 크기 관계를 가지는 T의 서브시퀀스를 말한다. 예를 들어, $T = (3, 1, 4, 2, 5)$이고 $P = (2, 1, 3)$일 때, T에서 P의 순열 매칭을 모두 나열하면 (3, 1, 4), (3, 1, 5), (3, 2, 5), (4, 2, 5)이다. 즉, 위의 순열 매칭들은 가운데 원소가 가장 작고, 첫 번째 원소가 두 번째로 작고, 마지막 원소가 가장 크므로 P와 같은 크기 관계를 가지는 T의 서브시퀀스들이다. 문제는 순열 T가 주어졌을 때 주어진 T에서 $P_1 = (2, 4, 1, 3)$ 또는 $P_2 = (3, 1, 4, 2)$의 순열매칭이 존재하는 지를 판별하는 것이다.

그림 2 예비문제의 예

그림 2의 순열매칭 문제는 문제의 요점만을 수학적으로 기술하고 있다. 이 문제는 단순한 방법으로 해결을 시도하면 $O(n^4)$ 이상의 시간복잡도를 갖게 되고, 출제자는 $O(n)$ 또는 $O(n\log n)$의 시간복잡도를 갖는 프로그램을 정답으로 하기를 의도하고 있다. 출제자가 원하는 시간제한을 고려하면 이 문제는 난이도가 높은 문제이다.

2.2 선제 과정

문제 수합이 완료되면 그 동안 여러 경시대회에서 활동한 전문가 교수들 위주로 경시대회에 사용할 문제를 선택하기 위한 선제 위원회가 열린다. 선제 위원회에서는 먼저 책에 나와 있거나 기존 경시대회 문제와 유사한 문제들은 배제한다. 다음으로 난이도 및 풀이 방법이 다양하게 배분되도록 8개 내지 10개의 문제를 선정한다.

이 과정에서 선정된 문제들에 대해서는 선제 과정에서 보완이 필요하다고 토의된 사항들을 출제자에게 전달한다. 또 다수의 문제가 인터넷 예선 문제들로 배정된다.

2.3 문제 각색

선제단계에서 선택된 문제들을 경시대회에 사용하기 위해서는 상당 부분 가공이 필요하다. 먼저 고려할 점은 문제의 내용이 되도록 현실 세계와 연관되고 학생들의 흥미를 유도할 수 있도록 만드는 것이다. 초기 문제 내용이 수학적이거나 이론 위주인 경우 출제자에게 현실적이고 흥미로운 문제로 각색해 줄 것을 요구한다.

그림 3에 앞의 순열매칭 문제를 경시대회용으로 다듬은 문제의 예가 있다. 그림 3에서 보듯이 초기 문제의 수학적인 서술은 되도록 최소화 되고 가능한 한 흥미를 유발시키도록 순열 매칭 문제가 암호화와 관련되어 특정한 키를 찾는 문제로 각색되었음을 알 수 있다.

문제각색이 완료되면 문제 내용, 입출력 방법 등을 영문으로 작성한다. 지면 관계상 그림 3의 Weak Key 문제에 대한 입출력의 예와 샘플 입출력은 2004년에 서울에서 열린 ACM-ICPC 아시아 지역대회[2]의 Problem H를 참조하기 바란다.

다듬어진 영문 문제들은 출제자 전체회의를 통해 다시 검토되며 문제 서술이 정확하고 전체 문제들이 난이도 배분 등 경시대회의 요구사항을 만족한다고 판단되면 채점을 위한 과정으로 넘어간다.

Weak Key

Cheolsoo is a cryptographer in ICPC (International Cryptographic Program Company). Recently, Cheolsoo developed a cryptographic algorithm called ACM(Advanced Cryptographic Method). ACM uses a *key* to encrypt a message. The encrypted message is called a *ciphertext*.

In ACM, to decrypt a cipher text, the same key used in the encryption should be applied. That is, the encryption key and the decryption key are the same. So, the sender and receiver should agree on a key before they communicate securely using ACM. Soon after Cheolsoo finished the design of ACM, he asked its analysis on security to Younghee who is a cryptanalyst in ICPC.

Younghee has an interest in breaking cryptosystems. Actually, she developed many attacking methods for well-known cryptographic algorithms. Some cryptographic algorithms have *weak keys*. When a message is encrypted with a weak key, the message can be recovered easily without the key from the cipher text. So, weak key should not be used when encrypting a message. After many trials, she found the characteristic of weak keys in ACM. ACM uses a sequence of mutually distinct positive integers $(N_1, N_2, \ldots, N_k)$ as a key. Younghee found that weak keys in ACM have the following two special patterns:

There are four integers N_p, N_q, N_r, N_s
$(1 \le p < q < r < s \le k)$ in the key such that
(1) $N_q > N_s > N_r > N_p$ or (2) $N_q < N_s < N_r < N_p$

Now, Younghee wants to find an efficient method to determine, for a given key, whether it is a weak key or not. Write a program that can help Younghee.

그림 3 순열 매칭 문제를 각색한 문제

3. 채점 과정

ACM-ICPC의 특징 중 하나는 각 문제에 있어 테스트 데이터의 일부에 대해서만 정답을 출력하는 경우에 부분 점수를 주지 않는다는 점이다. 즉, 어떤 문제에 대한 모든 테스트 데이터를 제한시간 내에 해결한 경우에만 그 문제를 해결한 것으로 채점한다. 이 절에서는 정확한 채점을 하기 위해 사전에 준비해 두어야 하는 세부 단계들을 살펴본다.

3.1 테스트데이터의 설계

정확하고도 효율적인 프로그램만이 통과할 수 있도

데이터크기	데이터 설명
4	20, 40, 10, 30
4	30, 10, 40, 20
30,000	증가수열
30,000	감소수열
30,000	증가하다 감소하는 수열
30,000	감소하다 증가하는 수열
30,000	증가수열에서 인접한 둘씩 교환
30,000	감소수열에서 인접한 둘씩 교환
30,000	증가수열에서 연속한 세개 반전
30,000	감소수열에서 연속한 세개 반전
30,000X5	위 5개 데이터에 답 패턴 넣음
30,000X2	스택 높이를 고려한 데이터
30,000X3	3개의 random 데이터

그림 4 Weak Key 문제의 테스트데이터

록 테스트데이터를 설계한다. 먼저, 제한적인 경우에만 문제를 해결하는 프로그램이 통과하지 못하도록 다양한 경우를 고려한 충분한 개수의 테스트 데이터를 설계한다. 또, 비효율적인 방법으로 작성된 프로그램들이 제한시간 내에 통과하지 못하도록 테스트데이터의 크기를 충분히 크게 만든다.

그림 3의 Weak Key 문제에 대한 테스트데이터의 설계가 그림 4에 나타나 있다. 이 경우 20개의 테스트데이터가 사용되었으며 18개의 데이터는 각각 30,000개의 정수로 구성되어 있다. 데이터 설명에서 보듯이 가능한 여러 가지 방법으로 정확하고 효율적인 프로그램만이 통과하도록 설계한다.

3.2 정답 프로그램 작성

출제자가 직접 자신이 고안한 방법으로 프로그램을 작성하여 다양한 형태의 데이터로 실험해보면서 변수의 범위, 테스트 데이터의 크기 및 제한시간을 설정한다. 문제를 해결하는 방법이 여러 개인 경우에는 가능한 여러 가지 방법으로 프로그램을 작성해서 최종 설계된 테스트데이터에 대한 정답을 여러모로 검증한다.

그림 3의 Weak Key 문제에 대한 정답 프로그램 중 main 함수가 그림 5에 나타나 있다.

3.3 경시대회에서의 채점

경시대회 당일에는 대회 중 각 팀에서 온라인으로 제출된 답을 정답 프로그램의 수행결과와 비교하여 그 결과를 곧바로 통보한다. 각 문제의 출제자가 직접 채점하며 제한 시간 이내에 모든 테스트 데이터에 대해 정답을 출력하는 프로그램만을 통과시킨다. 통과되지 못한 경우 그 이유에 해당하는 불합격 유형을

```
int main(void)
{
    int data_size, k;
    int data[SIZE];
    int temp[SIZE];
    ifstream fin("key.in");
//  read the number of test cases
    fin >> data_size;
//  read the length of a sequence k
    for(int i=0; i<data_size; i++) {
        fin >> k;
//  read k distinct positive integers
    for(int j=0; j<k; j++) {
        fin >> data[j]; }
//  heap sort
    for(int j=0; j<k; j++)
        temp[j] = data[j];

        make_heap(&temp[0], &temp[k]);
        sort_heap(&temp[0], &temp[k]);

    for(int j=0; j<k; j++)
data[j]=bin_search(temp, k, data[j]) + 1;
    if(check_key(data, k))
        cout << "NO" << endl;
    else cout << "YES" << endl;
    }
    fin.close();
    return 0;
}
```

그림 5 Weak Key 정답 프로그램의 main 함수

온라인으로 전달한다. 이때 전달되는 불합격 유형은 다음과 같다.

- 정답과 다른 출력
- 컴파일 에러
- 런타임 에러
- 제한시간 초과
- 출력형식 에러

각 문제에 대한 시간제한은 공개하지 않으며 보통 정답 프로그램의 수행시간에 약간의 시간을 추가하여 설정한다. 학생들과의 질문 및 답변도 온라인상에서 이루어진다. 질문에 대한 답변이 일반적인 경우에 해당하면 모든 팀이 볼 수 있도록 한다. 질문 내용이 문제에 대한 힌트를 요구하는 것이거나 문제 내용에 이미 포함되어 있는 경우는 "No Comment"로 답한다.

4. 맺음말

본고에서는 ACM-ICPC 문제의 출제과정과 채점과정

을 예와 함께 살펴보았다. 그리고 출제와 채점의 각 세부 단계에서 프로그래밍 교육 효과와 정확한 프로그래밍을 유도하기 위해 고려된 사항들을 살펴보았다. 다년간 프로그래밍 경시대회 문제들을 출제하고 채점하는 과정에서 느낀 점들 중 하나는 ACM-ICPC와 같이 테스트데이터를 잘 설계하고 모든 테스트데이터를 정확히 처리하는 경우에만 프로그램을 통과시키는 절차의 중요성이다. 실제 학부의 프로그래밍 교육에 있어서도 되도록 이러한 절차를 통해 학생들을 교육하는 것이 학생들에게 프로그래밍에 임하는 올바른 자세를 심어주기 위해 매우 중요하다고 생각된다. 그동안 대전과 서울에서 열렸던 ACM-ICPC 아시아 지역 예선에 출제된 문제들이 이와 같은 교육적 효과를 염두에 두고 고안된 문제들이므로 프로그래밍 교육 현장에서 유용하게 쓰이기를 바란다.

참고문헌

[1] http://acm.kaist.ac.kr/

[2] http://acm.kaist.ac.kr/2004/

임 형 석

1983 서울대학교 컴퓨터공학과(학사)
1985 한국과학기술원 전산학과(석사)
1993 한국과학기술원 전산학과(박사)
1996~1997 미국 Purdue 대학 (방문교수)
1987~현재 전남대학교 전자컴퓨터공학부(교수)
관심분야 : 알고리즘, 그래프 이론, 생물정보처리
E-mail : hslim@chonnam.ac.kr

SWCC 2007(하계 컴퓨터통신 워크샵)

- 일 자 : 2007년 8월 23~25일
- 장 소 : 제주 라마다프라자 호텔
- 내 용 : 논문발표 등
- 주 최 : 정보통신연구회
- 상세안내 : http://swcc.or.kr