

[10~13] 다음 글을 읽고 물음에 답하시오.

문자 입력 창에 한 글자만을 입력했는데 완성된 문구가 ㉠ 제시되는 자동 완성을 경험해 보았을 것이다. ‘코’라는 문자를 입력했다면 ‘코피’, ‘코로나’ 등이 후보로 제시되어 휴대 전화와 같이 문자 입력이 불편한 경우 문자 입력을 편리하게 할 수 있다. 이는 사용했던 단어들 중에서 입력되는 문자와 첫 글자부터 일치하는 것을 찾고 그중 사용 빈도가 높은 단어들을 후보로 제시하는 것이라고 할 수 있다. 한편 워드 프로세서에서 단어 찾기와 같은 검색은 저장되어 있는 문자열을 대상으로 검색어가 ㉡ 포함된 문자열을 찾는 것이다. 검색은 자동 완성과 달리 대상 문자열의 어느 위치에서도 검색어를 찾을 수 있어야 하며 사용 빈도를 고려하지 않아도 된다.

검색이 가능하기 위해서는 검색어를 저장되어 있는 문자열의 부분 문자열과 비교하는 알고리즘이 필요하다. 예를 들어 ‘우리글’이라는 검색어를 ‘한글: ㉢ 우리나라에서 ㉣ 창제된 ㉤ 우리글’이라는 띄어쓰기(㉢)가 포함된 18글자의 대상 문자열에서 검색한다고 ㉥ 가정해 보자. ㉦ 가장 간단히 떠올릴 수 있는 방법은 ‘우리글’이 3글자이므로 대상 문자열을 3글자씩 잘라 1글자씩 비교하는 것이다. ‘한글’, ‘글: ㉢’, ‘: ㉣우’ 등과 같이 16개의 비교 대상을 만들고 이를 검색어와 각각 비교하여 모두 같은지 확인한다. 하나의 비교 대상을 확인하기 위해서는 3글자를 각각 비교해야 하므로 총 16×3번 비교를 하게 될 것이다. 검색어 길이에 비해 대상 문자열이 짧거나 같은 경우는 없으므로 이 방법은 검색어와 비교해야 하는 대상 문자열의 길이가 길어지거나 개수가 많아지면 비교 횟수가 늘어나 검색 시간이 늘어난다.

검색 시간을 줄이기 위한 다른 방법은 없을까? 검색어와 비교 대상을 1글자씩 비교하지 않고 3글자씩 한 번에 비교할 수 있다면 그만큼 비교 횟수가 줄어들게 되어 검색 시간이 줄어들 것이다. 이를 위해 각각의 문자열에 특정 값을 ㉧ 생성하는 함수를 설정할 수 있다. 이런 함수를 해시 함수라고 하고, 어떤 문자열에 대해 해시 함수가 생성한 값을 해시값이라고 한다. 만일 해시 함수가 입력 가능한 문자열에 대해 모두 다른 해시값을 생성한다면 검색어의 해시값과 비교 대상의 해시값을 비교하여 두 문자열이

[A] 일치함을 단번에 ㉨ 판단할 수 있다.

앞의 예와 같이 검색어가 3글자이고 18글자의 대상 문자열이 제시된다면 비교 대상은 16개가 만들어진다. 하지만 각 비교 대상에서 문자열 비교는 1번의 해시값 비교로 줄어들기 때문에 전체 비교 횟수는 감소하게 된다. 물론 해시값을 생성하는 해시 함수의 연산이 추가되지만 추가되는 연산 시간이 각 글자 단위의 비교에 필요한 연산 시간보다 짧다면 전체적인 검색 시간은 단축될 수 있다. 이런 이유로 해시 함수는 연산이 간단하면서도 중복되지 않는 해시값을 생성할 수 있어야 한다.

10. 윗글을 통해 알 수 있는 내용으로 적절하지 않은 것은?
- ① 검색은 저장되어 있는 문자열 전체를 대상으로 검색어가 포함되어 있는지 확인한다.

- ② 검색은 필요에 따라 각기 다른 문자열에 동일한 해시값을 생성하는 해시 함수를 사용한다.
- ③ 검색은 저장되어 있는 문자열의 부분 문자열과 검색어를 비교하는 알고리즘을 활용한다.
- ④ 자동 완성은 사용 빈도를 고려하여 입력되는 문자가 포함된 문자열을 후보로 제시한다.
- ⑤ 자동 완성은 휴대 전화와 같이 문자 입력이 불편한 경우 문자 입력을 편리하게 할 수 있는 방법이다.

11. [A]를 이해한 내용으로 적절한 것은?

- ① 검색어의 길이가 짧아진다면 비교 대상의 개수가 줄어들어 해시값 비교 횟수가 증가할 수 있겠군.
- ② 대상 문자열에 반복되는 글자가 많다면 해시값이 작아져서 해시 함수의 연산 시간이 단축될 수 있겠군.
- ③ 검색어보다 긴 대상 문자열의 개수가 늘어난다면 비교 대상이 늘어나 해시값 비교 횟수가 증가할 수 있겠군.
- ④ 대상 문자열이 1개일 경우 검색어의 길이가 짧아진다면 비교 대상의 길이가 줄어들어 해시값 비교 횟수가 감소할 수 있겠군.
- ⑤ 대상 문자열이 2개일 경우 검색어의 길이가 길어진다면 비교 대상의 개수가 늘어나 해시 함수의 연산 시간이 증가할 수 있겠군.

12. ㉦에 <보기>의 조건을 모두 추가하여 검색한다고 할 때, 이에 대한 설명으로 적절하지 않은 것은? [3점]

<보 기>

[조건]

- 검색어에 문장 부호가 포함되지 않는 경우 문장 부호가 있는 부분 문자열은 비교 대상에서 제외한다.
- 검색어에 띄어쓰기가 포함되는 경우 띄어쓰기의 위치가 일치하지 않는 부분 문자열은 비교 대상에서 제외한다.

- ① ‘우리 ㉢글’로 검색할 경우 띄어쓰기의 위치가 일치하는 비교 대상 3개가 만들어진다.
- ② ‘우리 ㉢글’로 검색할 경우의 비교 횟수보다 ‘우리글’로 검색할 경우의 비교 횟수가 더 많다.
- ③ ‘우리글’로 검색할 경우 비교 대상은 ‘㉢우리’, ‘우리나’, ‘리나라’ 등과 같이 3글자로 된 비교 대상들이 만들어진다.
- ④ ‘우리글’로 검색할 경우 부분 문자열 ‘한글’, ‘글: ㉢’, ‘: ㉣우’에는 문장 부호가 포함되어 있기 때문에 비교하지 않는다.
- ⑤ ‘우리글’로 검색할 경우 일치하는 문자열을 찾을 수 있지만 ‘우리 ㉢글’로 검색할 경우는 일치하는 문자열을 찾을 수 없다.

13. ㉠~㉥의 사전적 의미로 적절하지 않은 것은?

- ① ㉠: 어떠한 의사를 말이나 글로 나타내어 보임.
- ② ㉡: 어떤 사물이나 현상 가운데 함께 들어 있거나 함께 넣음.
- ③ ㉢: 다른 사람의 말이나 행동, 형편 따위를 잘 알아서 긍정하고 이해함.
- ④ ㉣: 사물이 생겨남. 또는 사물이 생겨 이루어지게 함.
- ⑤ ㉤: 사물을 인식하여 논리나 기준 등에 따라 판정을 내림.