

| 카카오톡 봇 강좌 | >

[고급] 문자열 탐색 알고리즘 (KMP)



윤동원 챗봇 마스터 🍷 1:1 채팅

2020.06.18. 13:43 조회 96

댓글 5 URL 복사 ⋮



KMP Algorithm

- 시간 복잡도
 - 문자열의 길이 = N
 - 찾는 단어의 길이 = M
 - $O(N + M)$
- 문자열을 Brute-Force 방식으로 탐색하게 되면 문자열이 길어질수록, 찾는 단어가 길수록 시간이 오래 걸립니다.
- 예를 들어 길이가 10000인 Text에서 길이가 50인 word를 모두 찾는다면 Brute-Force방식에서는 5만번의 탐색이 필요합니다.
 - [\[JS\]문자열 탐색\(Brute-Force 방식\)](#)
- 이 문제를 해결하기 위해 Knuth, Morris, Prett 세 사람이 새로운 문자열 탐색 알고리즘을 만들었습니다. 만든 사람들의 이름 첫 글자를 따서 **KMP Algorithm**이라고 불립니다.
- KMP알고리즘은 접두사(prefix)와 접미사(suffix)를 이용해 접두사와 접미사를 이용해 현재 탐색중인 부분에서 접미사와 접두사가 일치하는 부분으로 점프해서 넘어가는 식으로 탐색을 합니다.
- 따라서 길이가 10000인 Text에서 길이가 50인 word를 모두 찾는다면 10050번의 탐색이 필요하게 됩니다.
- 저도 그랬지만 처음 KMP 알고리즘을 접하시는 분들은 쉽게 이해가 가지 않으실거라고 생각합니다.
- 제가 설명하는 재주가 부족해 KMP알고리즘을 이해하는데 큰 도움이 됐던 링크들을 첨부하겠습니다.
 - [KMP 문자열 매칭 알고리즘 - 동빈나님 유튜브](#)
 - [KMP : 문자열 검색 알고리즘 - 멍멍멍님 블로그](#)

문자열 탐색 코드(Brute-Force 방식)

```
function buildPatternTable(word) {
  const patternTable = [0];
  let prefixIndex = 0;
  let suffixIndex = 1;

  while (suffixIndex < word.length) {
    if (word[prefixIndex] === word[suffixIndex]) {
```

```

        patternTable[suffixIndex] = prefixIndex + 1;
        suffixIndex++;
        prefixIndex++;
    } else if (prefixIndex === 0) {
        patternTable[suffixIndex] = 0;
        suffixIndex++;
    } else {
        prefixIndex = patternTable[prefixIndex - 1];
    }
}

return patternTable;
}

function KMP(text, word) {
    if (word.length === 0) {
        return 0;
    }

    let textIndex = 0;
    let wordIndex = 0;

    const patternTable = buildPatternTable(word);

    while (textIndex < text.length) {
        if (text[textIndex] === word[wordIndex]) {
            // We've found a match.
            if (wordIndex === word.length - 1) {
                return textIndex - word.length + 1;
            }
            wordIndex++;
            textIndex++;
        } else if (wordIndex > 0) {
            wordIndex = patternTable[wordIndex - 1];
        } else {
            wordIndex = 0;
            textIndex++;
        }
    }

    return -1;
}

// test code

console.log(KMP("asdfasdfasdfasdfadsfaaaaa", "aaaaa")); >
27

console.log(KMP("asdfasdfasdfasdfadsfaaaaa", "aaaaaa")); >
-1

```

문자열 탐색 코드(KMP Algorithm)

```

function buildPatternTable(word) {
    const patternTable = [0];
    let prefixIndex = 0;
    let suffixIndex = 1;

    while (suffixIndex < word.length) {
        if (word[prefixIndex] === word[suffixIndex]) {
            patternTable[suffixIndex] = prefixIndex + 1;
            suffixIndex++;
            prefixIndex++;
        } else if (prefixIndex === 0) {
            patternTable[suffixIndex] = 0;

```

```

        suffixIndex++;
    } else {
        prefixIndex = patternTable[prefixIndex - 1];
    }
}

return patternTable;
}

function KMP(text, word) {
    if (word.length === 0) {
        return 0;
    }

    let textIndex = 0;
    let wordIndex = 0;

    const patternTable = buildPatternTable(word);

    while (textIndex < text.length) {
        if (text[textIndex] === word[wordIndex]) {
            if (wordIndex === word.length - 1) {
                return textIndex - word.length + 1;
            }
            wordIndex++;
            textIndex++;
        } else if (wordIndex > 0) {
            wordIndex = patternTable[wordIndex - 1];
        } else {
            wordIndex = 0;
            textIndex++;
        }
    }

    return -1;
}

// test code

console.log(KMP("asdfasdfasdfasdfadsfaaaaa", "aaaaa"));
> 27

console.log(KMP("asdfasdfasdfasdfadsfaaaaa", "aaaaaaa"));
> -1

```

출처 : <https://velog.io/@yujo/JSKnuth-Morris-PrettKMP-Algorithm>



윤동원님의 게시글 더보기 >

❤ 좋아요 2 💬 댓글 5

🔗 공유 | 신고

댓글 등록순 최신순 🔖

댓글알림 ☐



윤 위해 웃어줘
저도 저거 봤었는데 ㅋㅋ

2020.06.18. 13:51 답글 쓰기



KeonWoo PARK
잘 읽었습니다. 감사합니다.

2020.06.18. 16:23 답글 쓰기



재승



우와

2020.06.18. 18:38 답글 쓰기



JusticeHui

<https://cafe.naver.com/nameeyee/728>

2020.06.18. 18:42 답글 쓰기



JusticeHui

kmp를 "잘" 이해하려면 오토마타를 알아야 돼서... 쉽지는 않을 것 같네요

항상 최악의 경우를 생각해야하는 상황(정보올림피아드 등)이 아니라면 보이어-무어 알고리즘도 매우 좋은 성능을 내요

2020.06.18. 18:47 답글 쓰기

Hibot

댓글을 남겨보세요



등록

글쓰기

답글

목록

▲ TOP

'카카오톡 봇 강좌' 게시판 글

전체 [고급] 말머리 글

이 게시판 새글 구독하기 ☐

[응용] 메봇에서 스크립트가 로드된 상태의 웹페이지 파싱하기 [7]	산하	2020.06.18.
[고급] Rabin-Karp Algorithm 🤖 [4]	윤동원	2020.06.18.
[고급] 문자열 탐색 알고리즘 (KMP) 🤖 [5]	윤동원	2020.06.18.
[중급] 소스 줄이기(표준내장객체의 확장) [5]	qbvkde	2020.06.16.
[응용] 배열 응용 - 다양한 경우에서의 메서드 사용 🤖 [9]	도미 doami2005	2020.06.14.

1 2 3

전체보기

이 카페 인기글

카페 오픈챗 인원제한 늘렸어요

성빈

♡0 💬7

버튼 누르면 intent가 함수를 실행하게하려
면

jomin398

♡0 💬18

로코가 뭐길래 사람들이 환장하는걸까?

쿨칠

♡1 💬18



카카오톡이 이런걸 추천하다니

ACTIVE
오픈이발 쓰는사람 계정 영정먹이는 방법

Kiri

♡0 💬14



등급제 폐지 건의

- 1
- 2
- 3
- 4
- 5



웹툰



아키네이터



메신저R카톡방 들어와 채팅치면 환영해주