

| 카카오톡 봇 강좌 | >

[초급] 강좌 #14 - 정규식



도미 doami2005 챗봇 창시자 🌞 1:1 채팅

2020.03.20. 23:57 조회 493

💬 댓글 17 URL 복사 ⋮

시작하기 전에, 이 글은 자바스크립트를 전혀 모르는 사람도, 아는 사람도 읽어주셨으면 합니다.

자바스크립트의 시작부터 끝까지 복습하는 과정입니다.

이미 아시는 분들은 복습이라 생각하며 읽어주시면 되고, 댓글에 피드백 해주셨으면 합니다.

저번 강좌 (반복문)

- <https://m.cafe.naver.com/nameeyee/13806>

7. 정규식

[1] 정규식이란?

정규 표현식은 문자열에 나타나는 특정 문자 조합과 대응시키기 위해 사용되는 패턴입니다.

줄여서 정규식이라고 합니다.

정규식은 replace, split, match 등에 들어갈 수 있습니다.

이번 강좌에선 정규식을 가장 많이 쓰는 replace 를 사용해 보겠습니다.

[2] 형태

정규식의 기본 형태는 이렇습니다.

/패턴/

패턴에는 단순 글자와, 특정 기능을 하는 특수문자가 들어갈 수 있습니다.

[3] 사용법

```
a = "가나다라마"
```

```
a.replace(/가나/, "")
```

```
-> "다라마"
```

replace 문은 해당 패턴을 사용하여 오른쪽 것으로 바꿔버리는데, 원본 문자열을 바꾸진 않으며 복제본 값을 반환합니다.

```
a = "가나다라마"
```

```
a.replace(/가나/, "")
```

```
a
```

```
-> "가나다라마"
```

a.replace(/가나/, "") 에서는 가나를 한번만 바꿀 수 있지만, / 옆에 붙는 문자를 사용하면 이를 바꿀 수 있습니다.

바로 g, i 입니다.

```
"가나다라가나가나".replace(/가나/, "")  
-> "다라가나가나"
```

```
"가나다라가나가나".replace(/가나/g, "")  
-> "다라"
```

g 를 붙이면 가나 를 모두 공백으로 바꿔버립니다.

i 는 대소문자를 구분하지 않는 문자입니다.

```
str = "aAaAb"
```

```
str.replace(/A/, "")  
-> "aaaAb"
```

```
str.replace(/A/i, "")  
-> "AaaAb"
```

```
str.replace(/A/gi, "")  
-> "b"
```

위처럼 둘을 같이 쓸 수도 있습니다.

[4] 패턴

이제부터 패턴에 들어가는 특수문자들을 알아보겠습니다.

정규식 패턴에 들어가는 특수문자들은 수가 어마어마하게 많습니다.

그래서 외우는 것 보다는 필요할 때 찾아보는 것이 더 좋습니다.

이번 강좌엔 (), [], ^, \s, \S, +, -, \d, \D, \$ 들을 알아보겠습니다.

먼저 (), [] 는 범위를 나타냅니다.

그리고 + 는 여러개를 나타낼 때에 씁니다.

(~~)+ 하면 ~~ 가 1개 이상 있는 것을 찾습니다.

그리고 [] 는 안에 들어가는 모든 문자들을 인식합니다.

[ab] -> ab 가 아닌 a와 b 각각을 찾음

그리고 [] 안에는 - 가 들어갈 수 있습니다.

이 또한 범위를 뜻하는데요,

[0-9]+ 는 모든 숫자를 일컫습니다. (\d 와 효과가 같음)

\D 는 숫자가 아닌 모든 것을 찾습니다. (\d 와 반대)

\s 는 모든 공백문자들을 찾습니다. (\S 와 반대)

따라서 [\s\S]+, [\d\D] 는 모든 문자들을 인식합니다.

그리고 + 는 범위를 뜻하는 기호입니다. 이것도 [] 안에 특이한 수 있습니다.

그리고 `<` 는 언제나 붙어있는 기호입니다. 이것도 `[]` 안에 들어갈 수 있습니다.

`[^0-9]+ = \wD`

`[0-9]+ = \wd`

`$` 는 개인적으로 중요하다고 생각되는 특수기호입니다.

`()` 로 씌운 내용물들은 오른쪽 결과에서 `$`숫자 로 불러올 수 있습니다.

자세한 건 예제에서 보세요.

특수기호들은 이스케이프 문자라고 불리는 `\` 를 통해 일반 문자로 변환시킬 수 있습니다.

[5] 예제

우리가 처음 카톡봇 파싱을 배울 때, 정규식이란 단어를 들어봤을 겁니다.

해당 식은 `replace(/<([>]+)>/g, "")` 인데요,

하나하나 해석을 해보자면,

`<([>]+)>` 에서

`<` 는 문자 `<` 를 의미하고, `>` 도 문자 `>` 를 의미합니다.

따라서 `<` 와 `>` 사이에 있는 `[>]+` 입니다. (괄호는 범위.)

`[>]+` 를 해석해봅시다. `^` 는 아까 반대의 의미라고 했었죠.

따라서 `[>]` 는 `>` 가 아닌 모든 것을 의미하고,

`+` 는 여러개를 의미합니다.

따라서 `"< 와 > 사이에 있는 > 가 아닌 모든 것"` 이라고 해석할 수 있겠습니다.

`==` 태그 제거

한번 "1가 2나 3다 4라 5마" 라는 문자열을,

1위 : 가

2위 : 나

3위 : 다

4위 : 라

5위 : 마

로 `replace` 문과 정규식을 이용하여 바꿔보겠습니다.

`str = "1가 2나 3다 4라 5마"`

`(str+" ").replace(/([0-9])([^\0-9]+)/g, "$1위 : $2\n")`

`$` 의 기능에 대해 설명드리겠습니다.

아까 말했듯이 `$`숫자 는 괄호로 씌운 것을 그때그때 가져옵니다.

`$1` 은 첫번째로 괄호에 씌워진 `[0-9]` 를,

`$2` 는 두번째로 괄호에 씌워진 `[^\0-9]+` 를 가져옵니다.

그리고 `str` 에 " " 을 더한 이유는 숫자,숫자가아닌문자+,띄어쓰기 를 패턴에 인식하게 했기 때문입니다.

[6] 마침

지금까지 자바스크립트의 정규식에 대해 알아보았습니다.
읽어주셔서 감사합니다.

이번 강좌는 조금 이해가 힘들었을 수도 있습니다.
제가 참조한 문서로 도움이 되실 수도 있으니 찾아보세요.

다음 강좌는 "반복문 (응용)" 입니다. 다음 강좌에 꼭 추가되었으면 하는 내용을 댓글에 적어주세요

[7] 참조된 문서

- <https://developer.mozilla.org/ko/docs/Web/JavaScript/Guide/%EC%A0%95%EA%B7%9C%EC%8B%9D>
- <https://beomy.tistory.com/21>

/* 도미의 강좌 시리즈 */

[자료형]

- <https://cafe.naver.com/nameyee/13341>

[변수]

- 초급) <https://cafe.naver.com/nameyee/10808>
- 응용) <https://cafe.naver.com/nameyee/13765>

[주석]

- <https://cafe.naver.com/nameyee/13384>

[조건문]

- <https://cafe.naver.com/nameyee/13414>

[연산자]

- <https://cafe.naver.com/nameyee/13458>

[반복문]

- <https://cafe.naver.com/nameyee/13806>

[try문]

- <https://cafe.naver.com/nameyee/4540>

[함수]

- (???)

[배열]

- (???)

- <https://m.cafe.naver.com/nameyee/9524>

[객체]

- 초급) <https://cafe.naver.com/nameyee/4908>
- 응용) <https://cafe.naver.com/nameyee/8232>
- 심화) <https://cafe.naver.com/nameyee/9501>

[정규식]

- <https://cafe.naver.com/nameyee/13862>

[파싱]

- Utils
- (???)
- Jsoup
- (???)

[Tip]

eval 의 위험성과 사용법

- <https://cafe.naver.com/nameyee/4341>

파싱 시 Utils 와 Jsoup 의 장단점

<https://cafe.naver.com/nameyee/4342>

- <https://cafe.naver.com/nameyee/4342>

전체보기 하는 방법

<https://cafe.naver.com/nameyee/13539>



도미 doami2005님의 게시글 더보기 >

❤ 좋아요 3 💬 댓글 17

🔗 공유 | 신고

댓글 등록순 최신순 ↻

댓글알림 ☐



여노

형 요즘에도 홍우 다녀?

2020.03.21. 00:21 답글쓰기



도미 doami2005 작성자

오늘도 갔다옴

2020.03.21. 00:21 답글쓰기



여노

도미 doami2005 어라 나돈디 형 정인다닌댓나?

2020.03.21. 00:24 답글쓰기



도미 doami2005 작성자

여노 9층멘토

2020.03.21. 00:26 답글쓰기



여노

도미 doami2005 내일도 홍우 감? 나 내일 정인가는당

2020.03.21. 00:29 답글쓰기



도미 doami2005 작성자

여노 ㅇㅇ5~7시

2020.03.21. 00:31 답글쓰기



여노

도미 doami2005 9~12:30,4~6:30 아슬아슬하게 안겪치네(?)

2020.03.21. 00:33 답글쓰기



닉넴421

와우 강의 감사드립니다.

2020.03.21. 00:24 답글쓰기



구름

태그 제거에 관해서 보충 설명을 하자면

"<HTML>".replace(/<[^>]+>/g,"");

1. <HTML> //제거되기 전
2. HTML> //문자열 < 삭제
3. > //>를 제외하고 전부 삭제니 >앞까지 전부 삭제
4. //>를 삭제하고 아무것도 없음

위와 같습니다.

2020.03.21. 01:23 답글쓰기



도미 doami2005 작성자

? 패턴은 순서대로 감지만 합니다 바꾸는 건 그 다음입니다
삭제 X 감지 O

1. <HTML> //제거되기 전
2. HTML> //문자열 < 감지
3. > 가 나오기 전까지 감지 (HTML)
4. > 감지

5. 감지된 문자열들을 "" 로 변경

2020.03.21. 01:25 답글쓰기



구름

도미 doami2005 replace 전체를 기준으로 쓴 글입니다.
태그 제거라고 했군...

2020.03.21. 01:26 답글쓰기



도미 doami2005 작성자

구름 저도 replace 전체를 기준으로 쓴 댓글입니다.
정규식 패턴을 replace 함수에 대입하는거라 감지 후 삭제합니다.

2020.03.21. 01:28 답글쓰기



구름

도미 doami2005 활활... 고였어

2020.03.21. 01:29 답글쓰기



AlphaDo

아무도 알려주지 않아
나름대로 노가다해서 구한 성질(?)

RegExp.\$(숫자) 로 마지막으로 검사한 정규식의 대상(?) 이 되는 곳에서 값을 하나씩 불러옵니다.

```
Ex) if (/(\.)팡(.)/.test("앙팡디팡")) {  
  replier.reply(RegExp.$1); //앙  
  replier.reply(RegExp.$2); //디  
}
```

RegExp.\$' 랑 RegExp.\$` 도 있는데
이건 알아서 찾아봐요(?)
유용한 거시애오.

(RegExp["\$"] 이렇게 쓰면 라이노에서 예러 안걸림)

2020.03.21. 10:20 답글쓰기



큐브마인

+(1개 이상) 뿐만 아니라
*(0개 또는 2개 이상)
?(0개 또는 1개)도 있으면 좋을 것 같네요

좋은 강의 감사합니다

2020.03.21. 13:12 답글쓰기



good

좋은 강의 감사합니다

2020.03.21. 17:01 답글쓰기



웃는 꽃게

잘 보고 갑니당 ㅎㅎ

2020.03.24. 21:10 답글쓰기

Hibot

댓글을 남겨보세요



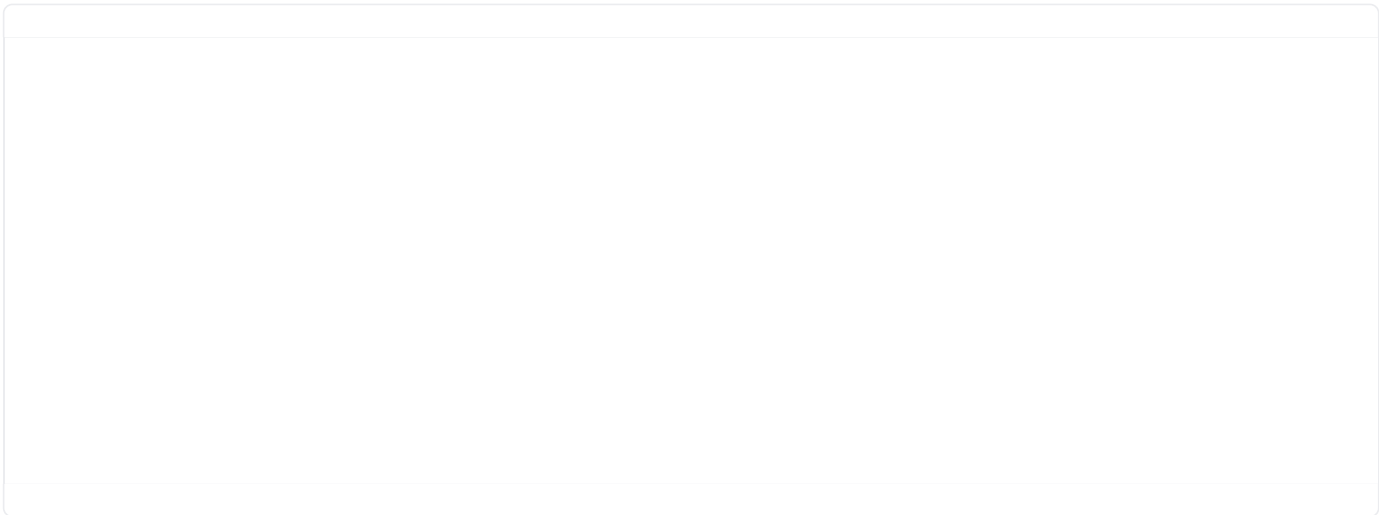
등록

글쓰기

답글

목록

▲ TOP



'| 카카오톡 봇 강좌 |' 게시판 글

전체 [초급] 말머리 글

이 게시판 새글 구독하기 ☐

jsoup 왕기초 - 암호화폐 가격 파싱 🤖 [10]	이슬두병	2020.03.25.
뭐함이가 알려주는 기초강좌 #1 [13]	뭐함	2020.03.25.
[초급] 강좌 #14 - 정규식 [17]	도미 doami2005	2020.03.20.
[카카오링크 강좌] 2. 상현굿님 위성사진 사용법 🤖 [21]	그린스크린	2020.03.20.
[메봇R] 사진저장을 포함한 채팅기록 #2 🤖 [22]	프로그	2020.03.19.


1 2 3

전체보기

이 카페 인기글

path질문

젤리파덕
♡0 💬2



혹시 여기서 변수를 저장하는 부분이 어딘지
아시는분 계시나요..??
넵!!!!!!!!!!!!!!
[확습] 모니카는 여기에 잘 있을
넵!!!!!!!!!!!!!!
[카카오톡 봇] 양산형 자동학습 봇 소스 공유

안녕하세요

tomohong
♡1 💬3

kakao.py에 대해서

undefined
♡0 💬13




태양, 달 정보 구현 완료

요즘 나타나는 새로운 홍보 수법 중 하나가

Dark Tornado
♡0 💬7

자바스크립트

도유니
♡0 💬10



링크를 포함한 글 홍보시 링크 분석(?)하기

1 2 3 4 5