

3 | HTML 분석

- 정규 표현식
 - 정규 표현식이라는 이름은 정규 문자열을 식별하는데서 유래되었다.
 - 정규 표현식은 문자열이 주어진 규칙에 일치하는지, 일치하지 않는지 판단할 수 있다.
 - 글자 a를 최소한 한 번 쓰시오
 - 그 뒤에 b를 정확히 다섯 개 쓰시오
 - 그 뒤에 c를 짝수 번 쓰시오
 - 마지막에 d가 있어도 되고 없어도 됩니다.
 - => aaaabbbbbccccd, aabbbbbbcc 등이 성립한다.
 - $aa^*bbbb(cc)^*(d|)$

3 | HTML 분석

■ 정규 표현식

- 규칙1 : 이메일 주소의 첫 번째 부분에는 다음 중 최소한 하나가 포함 되어야합니다. 대문자, 소문자, 숫자 0-9, 마침표(.), 플러스 기호(+), 밑 줄기호
 - `[A-Za-z0-9\._+]`
- 규칙2:그 다음에 @가 나타나야 한다.
 - `@`
- 규칙3:그 다음에는 반드시 대문자나 소문자가 최소한 하나 있어야 한다.
 - `[A-Za-z]`
- 규칙4: 그 다음에는 마침표가 온다.
 - `\.`
- 규칙5: 이메일 주소는 com, org, edu, net중 하나로 끝난다.
 - `(com|org|edu|net)`
- `[A-Za-z0-9\._+]+@[A-Za-z]+\.(com|org|edu|net)`

3 | HTML 분석

■ 정규 표현식 기호

기호	의미	예제	일치하는 문자열 예제
*	바로 앞에 있는 문자. 대괄호로 묶인 문자들이 0번 이상 나타납니다.	<code>a*b</code>	<code>aaaaaa, aaabbbbb, bbbbbbb</code>
+	바로 앞에 있는 문자. 대괄호로 묶인 문자들이 1번 이상 나타납니다.	<code>a+b+</code>	<code>aaaaaab, aabbbb, abbbbbb</code>
[]	대괄호 안에 있는 문자 중 하나가 나타납니다.	<code>[A-Z]*</code>	<code>APPLE, CAPITALS</code>
()	정규표현식을 평가할 때에는 하위표현식이 먼저 평가	<code>(a*b)*</code>	<code>aaabaab, abaaab</code>
{m,n}	m번 이상 n번 이하 문자들이 나타납니다.	<code>A{2,3}b{2,3}</code>	<code>aabbb, aaabbb, aabb</code>
[^]	대괄호 안에 있는 문자를 제외한 문자가 나타납니다.	<code>[^A-Z]</code>	<code>apple, banana</code>
	표현식중 하나가 나타납니다.	<code>b(a i e)d</code>	<code>bad, bid, bed</code>
.	문자 하나가 나타납니다.	<code>b.d</code>	<code>bad, bzd, b\$d</code>
^	바로 뒤에 있는 문자 혹은 하위 표현식이 문자열 맨 앞에 나타납니다.	<code>^a</code>	<code>apple, asdf, a</code>
\	특수 문자를 원래 의미대로 쓰게하는 문자입니다.	<code>\. \ \\</code>	<code>. \</code>
\$	하위 표현식이 마지막이라는 뜻입니다.	<code>[A-Z]*[a-z]*\$</code>	<code>ABCabc, zzzyx, Bob</code>
?!	‘포함하지 않는다’는 뜻입니다.	<code>^((?![A-Z]).)*\$</code>	<code>No-caps-here</code>