|  |  |
| --- | --- |
| Character class | Description |
| .(Dot) | Any character (may or may not match line terminators) |
| \d | A digit: [0-9] |
| \D | A non-digit: [^0-9] |
| \s | A whitespace character: [ space, \t (tab), \n (new line), \x0B (end of line), \f (form feed), \r (carriage) ] |
| \S | A non-whitespace character: [^\s] |
| \w | A word character: [a-zA-Z\_0-9] |
| \W | A non-word character: [^\w] |

String targetString = "A b$9;";

String regex = "\\W";

Pattern pattern = Pattern.compile(regex);

Matcher matcher = pattern.matcher(targetString);

String replacedStr = matcher.replaceAll("");

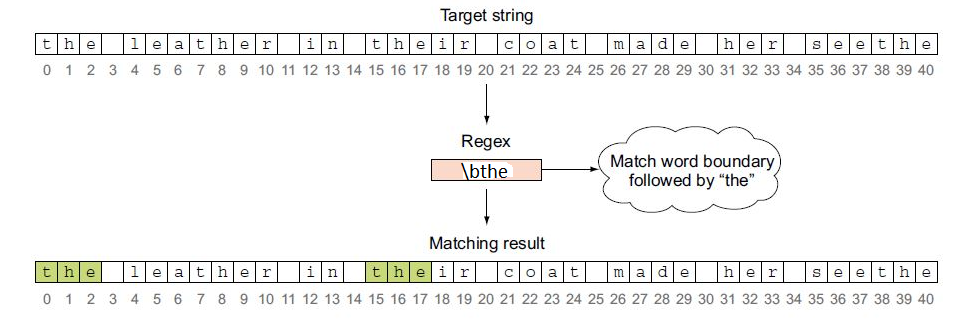
System.out.println(replacedStr);//**Ab9**

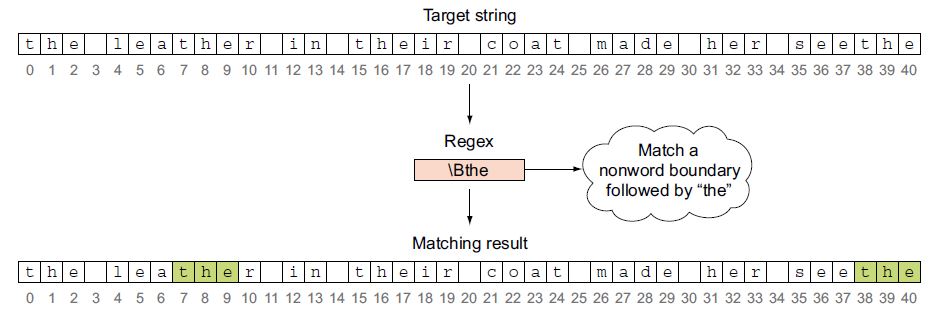
EXAM TIP Because String objects are immutable, calling replaceAll() won’t change the contents of String referred to by the variable target- String in the preceding code example. replaceAll() creates and returns a new String object with the replaced values. Watch out for questions based on it on the exam.

*Matching boundaries*

Say you want to find all occurrences of the word *the* in your book. To do that, you’d need to search for the text *the*. Well, the same is true when you want to match *the*, which can be part of another word, for example, *their*, *leather*, or *seethe*. So you need a way to limit your searches to the start or end of a word. *Matching boundaries* can help you with this. You can match boundaries including the start of a line, a word, a non-word, or the end of a line by using regex patterns. Even though the boundary constructs ^ (beginning of line) and $ (end of line) aren’t explicitly included in the exam objectives, you might see them in answer options that are incorrect.

|  |  |
| --- | --- |
| Boundary construct | Description |
| \b | A word boundary |
| \B | A non-word boundary |
| ^ | Beginning of a line |
| $ | End of a line |





\bthe means a word which will start with the and any characters after that

\Bthe means a word which contains the in any place but not at beginning.

the\b means a word that ends with the.

\bthe|the\b means a word that starts either with the or ends with the.