

正規表現の先読み について

Kazuhiro NISHIYAMA

*Ruby/Rails勉強会@関西 60th
2014/03/29*

目的

```
/(?=foo)bar/
```

はなぜマッチしないのか？

- `(?=foo)` の部分が先読み
- `foo` と `bar` の両方を含む文字列にマッチすることを期待する？

正規表現の先読みとは？

ある位置から続く文字列が
ある部分式にマッチするならば
その位置にマッチする
という正規表現

[http://docs.ruby-lang.org/
ja/2.1.0/doc/
spec=2fregexp.html](http://docs.ruby-lang.org/ja/2.1.0/doc/spec=2fregexp.html)

簡単に言うと

アンカーの一種で汎用的なもの

アンカーとは？

アンカーは幅0の文字列にマッチするメタ文字列

- ^ 行頭にマッチ
- \$ 行末にマッチ
- \A 文字列の先頭にマッチ
- \z 文字列の末尾にマッチ

つまり

- 幅を持たない
 - gsub などで \$& に入らない
- 後ろに続くパターンとマッチ開始位置が同じ

使用例 数字3桁区切り

```
parts[0].gsub!(  
  /(\d)(?=(\d\d\d)+(?!\d))/,  
  "\\1#{options[:delimiter]}")
```

https://github.com/rails/rails/blob/4-0-stable/activesupport/lib/active_support/number_helper.rb#L299
(number_to_delimited)

デモ

<https://gist.github.com/znz/9835956#file-regexp-201403-rb>