

# 99 problems... and regex

Mike Dowler



About me

IF YOU'RE HAVIN' PERL  
PROBLEMS I FEEL  
BAD FOR YOU, SON—



I GOT 99  
PROBLEMS,



SO I USED  
REGULAR  
EXPRESSIONS.



NOW I HAVE  
100 PROBLEMS.

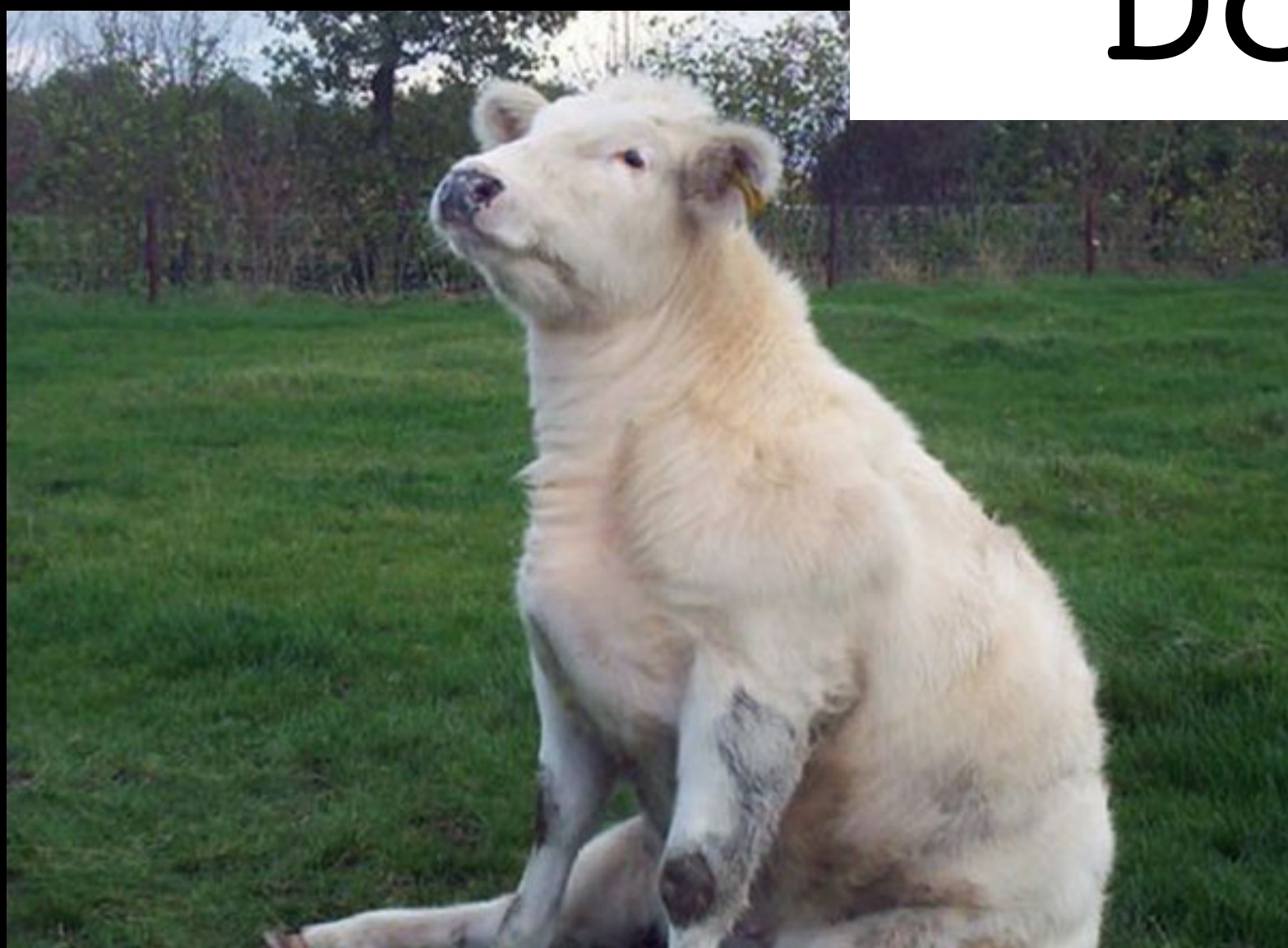


Regex, huh?  
What is it good for?





Dog?





# Scenario #1: Version numbers

How do we check if a software version meets or exceeds a given value?

4.9.0.504

5 ✓

4.9.9.9.9 ✓

4.9.0.0.504 ✗

4.9.0.503 ✗

4 ✗

4.10.0 ✓

# Start simple

4.9.0.504

4\ .9\ .0\ .504

# Start simple

4.9.0.504

4\ .9\ .0\ .504

3.14.9.0.504.7



# Start simple

4.9.0.504

`^4\.9\.0\.504$`

# Start simple

4.9.0.504

`^4\.9\.0\.504`

# Build it up - one digit

4.9.0.504



# Build it up - one digit

4.9.0.504

`^4\.9\.0\.50[456789]`

# Build it up - one digit

4.9.0.504

`^4\.9\.0\.50[456789]`

`^4\.9\.0\.50[4-9]`

# Build it up - two digits

4 . 9 . 0 . 5 0 4



# Build it up - two digits

4.9.0.504

`^4\\.9\\.0\\.5[1-9][0-9]`

# Build it up - two digits

4.9.0.504

```
^4\.9\.0\.5[1-9][0-9]
```

```
^4\.9\.0\.5[1-9]\d
```

# Build it up - one or two digits

```
^4\.9\.0\.50[4-9]
```

**OR**

```
^4\.9\.0\.5[1-9]\d
```



# Build it up - one or two digits

```
^4\.9\.0\.50[4-9]
```

**OR**

```
^4\.9\.0\.5[1-9]\d
```

```
^4\.9\.0\. (50[4-9]|5[1-9]\d)
```

# Build it up - three digits

4.9.0.504

# Build it up - three digits

4.9.0.504

$^4 \backslash .9 \backslash .0 \backslash . (50[4-9] | 5[1-9] \backslash d | \underline{[6-9] \backslash d \backslash d})$



# Build it up - three digits

4.9.0.504

$^4 \backslash .9 \backslash .0 \backslash . (50[4-9] | 5[1-9] \backslash d | \underline{[6-9] \backslash d \backslash d})$

$^4 \backslash .9 \backslash .0 \backslash . (50[4-9] | 5[1-9] \backslash d | \underline{[6-9] \backslash d \{2\}})$

# Build it up - beyond three digits

4.9.0.504

```
^4\.9\.0\.(50[4-9]|5[1-9]\d|[6-9]\d{2}|\d{4})
```

# Build it up - beyond three digits

4.9.0.504

```
^4\.9\.0\.(50[4-9]|5[1-9]\d|[6-9]\d{2}|\d{4})
```

```
^4\.9\.0\.(50[4-9]|5[1-9]\d|[6-9]\d{2}|\d{4,})
```

# Quantifiers

$$? = \{0, 1\}$$
$$* = \{0, \}$$
$$+ = \{1, \}$$

# Version numbers

4.9.0.504

```
^(4\. (9\. (0\. (50[4-9]|5[1-9]\d|  
[6-9]\d{2}|\d{4}))|[1-9])|[1-9]\d)|[5-9]|  
[1-9]\d)
```



# Scenario #2:

## Log file parsing

How can we extract (semi-)unpredictable data from predicatable contexts?

```
<backupConfig>
  <backupSets>
    <backupSet id="x">
      <name>yyyyyy</name>
      ...
    </backupSet>
    <backupSet id="x">
      <name>yyyyyy</name>
      ...
    </backupSet>
```

# Find the string - part 1

```
^\s*<backupSet id="\d+">
```

`\s` matches space, tab or newline

`\d` matches any digit

# Find the string - part 2

```
^\s*<backupSet id="\d+">\s*<name>.+<\//name>
```

. matches anything *except* newline

/ needs to be escaped

Find the **data** in the string

# Find the **data** in the string

```
^\s*<backupSet id="\d+">\s*<name>.+<\/name>
```



# Find the **data** in the string

```
^\s*<backupSet id="\d+">\s*<name>.+<\/name>
```

```
^\s*<backupSet id="(\d+)">\s*<name>(.+)<\/name>
```

# Find the **data** in the string

```
^\s*<backupSet id="\d+"\>\s*<name>.+</name>
```

```
^\s*<backupSet id="(\d+)">\s*<name>(.+)</name>
```

```
^\s*<backupSet id="(P<id-cap>\d+)">\s*<name>(P<name-cap>.+)</name>
```

# Practice!

<https://regex101.com/>

regular expressions 101

@regex101 donate contact bug reports & feedback wiki

SAVE & SHARE

Save Regex %+\$

FLAVOR

PCRE (PHP) ✓

ECMAScript (JavaScript)

Python

Golang

TOOLS

Code Generator

Regex Debugger

REGULAR EXPRESSION

3 matches, 87 steps (~1ms)

/ ^ ( 4 \. ( 9 \. ( 0 \. ( 5 0 [ 4 - 9 ] | 5 [ 1 - 9 ] \d | [ 6 - 9 ] \d { 2 } | \d { 4 } ) | [ 1 - 9 ] ) | [ 1 - 9 ] \d ) | [ 5 - 9 ] | [ 1 - 9 ] \d ) / g m

TEST STRING

SWITCH TO UNIT TESTS

5  
4.9.9.9.9.9  
4.9.0.0.504  
4.9.0.503  
4  
4.10

SUBSTITUTION

EXPLANATION

/ ^ ( 4 \. ( 9 \. ( 0 \. ( 5 0 [ 4 - 9 ] | 5 [ 1 - 9 ] \d | [ 6 - 9 ] \d { 2 } | \d { 4 } ) | [ 1 - 9 ] ) | [ 1 - 9 ] \d ) | [ 5 - 9 ] | [ 1 - 9 ] \d ) / g m

asserts position at start of a line

1st Capturing Group  
( 4 \. ( 9 \. ( 0 \. ( 5 0 [ 4 - 9 ] | 5 [ 1 - 9 ] \d | [ 6 - 9 ] \d { 2 } | \d { 4 } ) | [ 1 - 9 ] ) | [ 1 - 9 ] \d ) | [ 5 - 9 ] | [ 1 - 9 ] \d )

1st Alternative

MATCH INFORMATION

Match 1  
Full match 0-1 5  
Group 1. 0-1 5

Match 2  
Full match 2-7 4.9.9

QUICK REFERENCE

Search reference

All Tokens

Common Tokens

General Tokens

Anchors

Zero or one of a a?

Zero or more of a a\*

One or more of a a+

Exactly 3 of a a{3}

3 or more of a a{3,}

Between 3 an... a{3,6}

# Practice!

<https://regex101.com/>

regular expressions 101

@regex101 donate contact bug reports & feedback wiki

SAVE & SHARE

Save Regex %+\$

FLAVOR

PCRE (PHP) ✓

ECMAScript (JavaScript)

Python

Golang

TOOLS

Code Generator

Regex Debugger

REGULAR EXPRESSION

3 matches, 87 steps (~1ms)

```
^/^(4\.(9\.(0\.(50[4-9]|5[1-9]\d|[6-9]\d{2}|\d{4})|
[1-9])|[1-9]\d)|[5-9]|1[1-9]\d)/gm
```

TEST STRING

5  
4.9.9.9.9.9  
4.9.0.0.504  
4.9.0.503  
4  
4.10

SUBSTITUTION

EXPLANATION

^ asserts position at start of a line

1st Capturing Group  
(4\.(9\.(0\.(50[4-9]|5[1-9]\d|[6-9]\d{2}|\d{4})|
[1-9])|[1-9]\d)|[5-9]|1[1-9]\d)

1st Alternative

MATCH INFORMATION

Match 1

Full match 0-1 5

Group 1. 0-1 5

Match 2

Full match 2-7 4.9.9

QUICK REFERENCE

Search reference

All Tokens

Common Tokens

General Tokens

Anchors

Meta Characters

Zero or one of a a?

Zero or more of a a\*

One or more of a a+

Exactly 3 of a a{3}

3 or more of a a{3,}

Between 3 and 6 a{3,6}

# Practice!

<https://regex101.com/>

regular expressions 101 [@regex101](#) [donate](#) [contact](#) [bug reports & feedback](#) [wiki](#)

**SAVE & SHARE**

[Save Regex](#) %+\$

**FLAVOR**

☒ PCRE (PHP) ✓

☐ ECMAScript (JavaScript)

☐ Python

☐ Golang

**TOOLS**

[Code Generator](#)

[Regex Debugger](#)

**REGULAR EXPRESSION** 3 matches, 87 steps (~1ms)

```
^/^(4\.(9\.(0\.(50[4-9]|5[1-9]\d|[6-9]\d{2}|\d{4}))|
```

**TEST STRING** SWITCH TO UNIT TESTS

```
5
4.9.9.9.9.9
4.9.0.0.504
4.9.0.503
4
4.10
```

**EXPLANATION**

^ asserts position at start of a line

**1st Capturing Group**

(4\.(9\.(0\.(50[4-9]|5[1-9]\d|[6-9]\d{2}|\d{4}))|

**1st Alternative**

**MATCH INFORMATION**

Match 1

Full match	0-1	5
Group 1.	0-1	5

Match 2

Full match	2-7	4.9.9
------------	-----	-------

**QUICK REFERENCE**

Search reference

All Tokens	Zero or one of a	a?
Common Tokens	Zero or more of a	a*
General Tokens	One or more of a	a+
anchors	Exactly 3 of a	a{3}
	3 or more of a	a{3,}
	Between 3 an...	a{3,6}

# Practice!

<https://regex101.com/>

regular expressions 101

@regex101 donate contact bug reports & feedback wiki

SAVE & SHARE

Save Regex %+\$

FLAVOR

PCRE (PHP) ✓

ECMAScript (JavaScript)

Python

Golang

TOOLS

Code Generator

Regex Debugger

REGULAR EXPRESSION

3 matches, 87 steps (~1ms)

```
^/^(4\.(9\.(0\.(50[4-9]|5[1-9]\d|[6-9]\d{2}|\d{4}))|
[1-9])|[1-9]\d|[5-9]|1[1-9]\d)/gm
```

TEST STRING

```
5
4.9.9.9.9.9
4.9.0.0.504
4.9.0.503
4
4.10
```

SUBSTITUTION

EXPLANATION

^ asserts position at start of a line

1st Capturing Group

1st Alternative

MATCH INFORMATION

Match 1

Full match 0-1 5

Group 1. 0-1 5

Match 2

Full match 2-7 4.9.9

QUICK REFERENCE

Search reference

All Tokens

Common Tokens

General Tokens

Anchors

Zero or one of a a?

Zero or more of a a\*

One or more of a a+

Exactly 3 of a a{3}

3 or more of a a{3,}

Between 3 an... a{3,6}



# Practice!

<https://regex101.com/>

regular expressions 101

@regex101 donate contact bug reports & feedback wiki

SAVE & SHARE

Save Regex %+\$

FLAVOR

PCRE (PHP) ✓

ECMAScript (JavaScript)

Python

Golang

TOOLS

Code Generator

Regex Debugger

REGULAR EXPRESSION

3 matches, 87 steps (~1ms)

/ ^ ( 4 \. ( 9 \. ( 0 \. ( 5 0 [ 4 - 9 ] | 5 [ 1 - 9 ] \d | [ 6 - 9 ] \d { 2 } | \d { 4 } ) | [ 1 - 9 ] ) | [ 1 - 9 ] \d | [ 5 - 9 ] | [ 1 - 9 ] \d ) / g m

TEST STRING

5  
4.9.9.9.9.9  
4.9.0.0.504  
4.9.0.503  
4  
4.10

SUBSTITUTION

EXPLANATION

/ ^ ( 4 \. ( 9 \. ( 0 \. ( 5 0 [ 4 - 9 ] | 5 [ 1 - 9 ] \d | [ 6 - 9 ] \d { 2 } | \d { 4 } ) | [ 1 - 9 ] ) | [ 1 - 9 ] \d | [ 5 - 9 ] | [ 1 - 9 ] \d ) / g m

asserts position at start of a line

1st Capturing Group  
( 4 \. ( 9 \. ( 0 \. ( 5 0 [ 4 - 9 ] | 5 [ 1 - 9 ] \d | [ 6 - 9 ] \d { 2 } | \d { 4 } ) | [ 1 - 9 ] ) | [ 1 - 9 ] \d | [ 5 - 9 ] | [ 1 - 9 ] \d )

1st Alternative

MATCH INFORMATION

Match 1

Full match 0-1 5

Group 1. 0-1 5

Match 2

Full match 2-7 4.9.9

QUICK REFERENCE

Search reference

All Tokens

Common Tokens

General Tokens

Anchors

Meta Characters

Zero or one of a **a?**

Zero or more of a **a\***

One or more of a **a+**

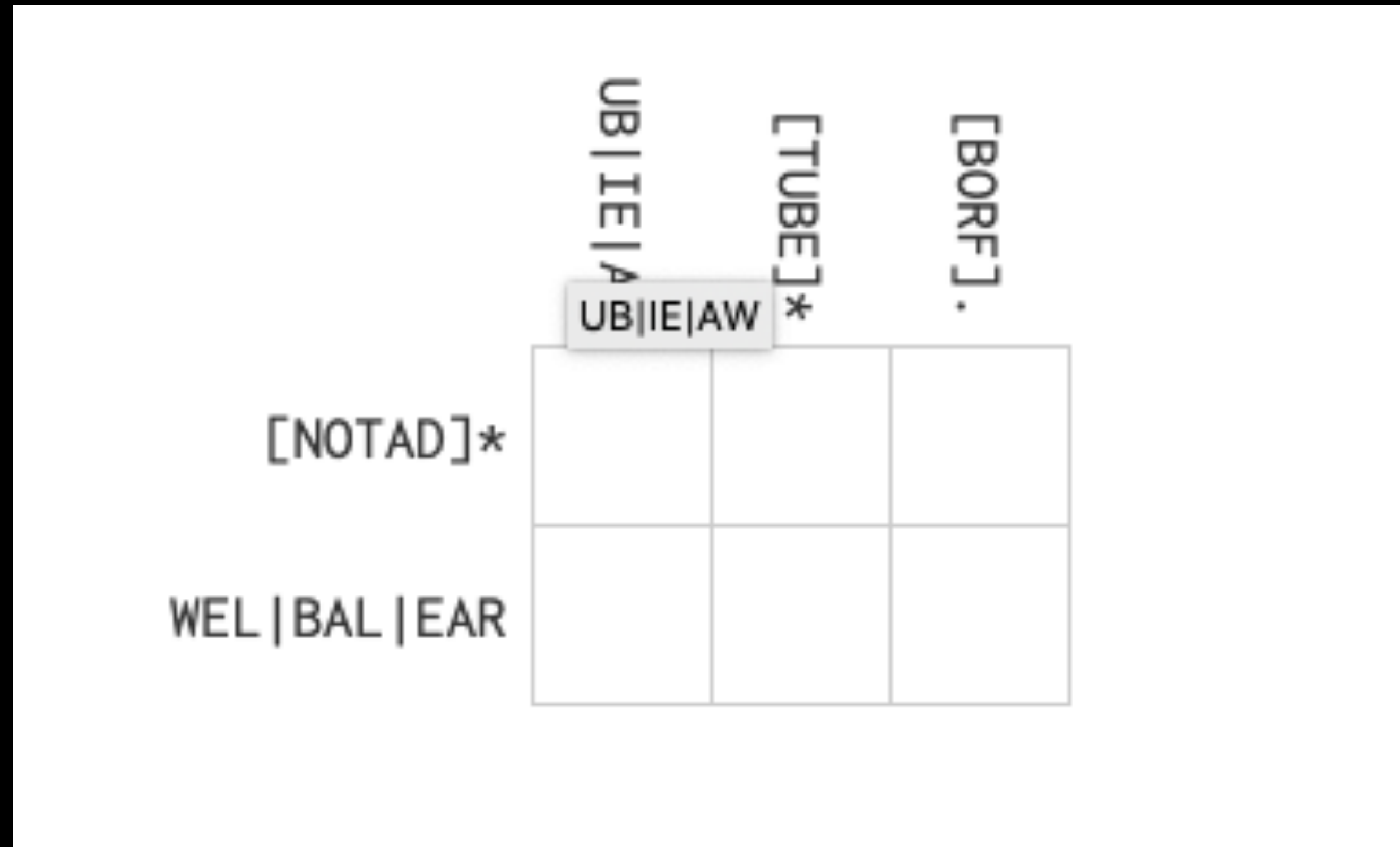
Exactly 3 of a **a{3}**

3 or more of a **a{3,}**

Between 3 and 6 of a **a{3,6}**

# Practice!

<https://regexcrossword.com/>



# Putting regex to work - Jamf

Test group

Computer Group

Criteria

☐ Show in Jamf Pro Dashboard

AND/OR		CRITERIA	OPERATOR	VALUE	
	( ▼	Application Title	is ▼	some_application.app	▼
and ▼	▼	Application Version	does not match regex ▼	^4\.(9\.(0\.(50[4-9] 5[1-9]\d 6-9]\d(2))\d{4	) ▼

# Putting it into practice - python 3

```
>>> import re

>>> result = re.match( r'^\s*<backupSet id="( ?P<id-  
cap>\d+)">\s*<name>( ?P<name-cap>.+)</name>', applog )

>>> result.group( 'id-cap' )
'66'

>>> result.group( 'name-cap' )
'Mike's Backup'
```

# Putting it into practice - zsh

```
% regex="<backupSet id=\"([[:digit:]]+)\">"

% if [[ $applog =~ $regex ]]; then
    echo $MATCH           // <backupSet id="66"
    echo ${match[1]}      // 66
fi
```

# Summary

- Methodology
- Positional anchors - ^ and \$
- Alternatives - [...]
- Classes - \d, \s, \w
- Quantifiers - {x,y}, ?, +, \*
- Groups - (...)

## Questions?