



Terence Eden's Blog



## [Regular Expressions make me feel like a powerful wizard](#) [- and that's not a good thing](#)

By @edent on 2023-02-06 · [#Computer Science](#) [#programming](#) [#regex](#) · [14 comments](#) ·  
600 words · Read ~278 times.



(This is a rant because I'm exhausted after debugging something. If you've made RegEx your whole personality, I'm sorry.)

The other day I had to fix a multi-line Regular Expression (RegEx). After a few hours of peering at it with a variety of tools, I finally understood the problem. Getting that deep into the esoteric mysteries made me feel like a powerful wizard with complete mastery of my domain. And I think that's dangerous.

I'm sure we've all read a story about a witch or wizard who distractedly substitutes eye-of-newt with iron-ute with disastrous consequences. Humans are easily confused. And confusion leads to unexpected mistakes.

Look, most humans are very bad at reading compiled code. Without any external tools - can you tell me what the following code does?

```
0000000 c031 d88e c08e 15be b47c ac0e 003c 0474
0000010 10cd f7eb 48f4 6c65 6f6c 202c 6f57 6c72
0000020 2164 0a0d 0000 0000 0000 0000 0000 0000
0000030 0000 0000 0000 0000 0000 0000 0000 0000
*
00001f0 0000 0000 0000 0000 0000 0000 0000 aa55
0000200
```

No. Of course not<sup>1</sup>. That's why we write code in a more human readable language and then compile it to computer readable instructions.

Regular Expressions are a sort-of halfway house. They're *slightly* readable by humans - but written in such a terse vocabulary as to be mostly unintelligible without concentration. There's no space for comments. Different engines have variable support for all their functions. They are a symbolic language with unhelpfully indecipherable and inconsistent symbols.

As a result, once the RegEx becomes more than trivially complex they're hard for most humans to understand. That makes them difficult to debug. It also makes it difficult to add or remove functionality.

I genuinely - and possibly misguidedly - believe that even something like

`^\w+([-+.']\w+)*@\w+([-.]\w+)*\.\w+([-.]\w+)*$` might just as well be written in [BrainFuck](#).

My contention is that *almost all* RegExs would be better served by more human readable code and that the very existence of [RegEx101.com](#) ought to bring shame on our industry.

Here are some positive use-cases for RegEx:

- ⇒ You want to show off how smart you are.
- ⇒ You need maximum efficiency when combing through a billion lines of text.
- ⇒ You have a desire to build something hard to debug.
- ⇒ You don't have lots of printer paper and need to make your code as terse as possible.
- ⇒ You think if/else and switch/case statements are the mark of a diseased mind.
- ⇒ You don't trust compilers.

I know what you're thinking: "This guy's too stupid to get regular expressions!" Yes. Yes I am. So are most people.

What I'm getting at is that source code is designed to be read and edited by busy and distracted humans. We should be writing intelligible code for each other and letting computers do the boring work of making it more efficient.

You don't have to agree with me. That's fine. But, perhaps you'll take note of the famous maxim from [the "Wizard" book](#):



a computer language is not just a way of getting a computer to perform operations but rather that it is a novel formal medium for expressing ideas about methodology. Thus, programs must be written for people to read, and only incidentally for machines to execute.

[Structure and Interpretation of Computer Programs](#)

We are not wizards. Nor should we strive to be. The alchemists fell.

🐾 Footnotes 🐾

01. You can [read the original code which is MIT Licenced.](#) ↩



## 14 thoughts on “Regular Expressions make me feel like a powerful wizard - and that's not a good thing”

 [taste of taboo](#) 🐼📷 says:

[2023-02-06 12:39](#)

[@Edent](#) regex are a total pain for adhd brains

[Reply](#)



 [Steven Pears](#) says:

[2023-02-06 12:53](#)

I wrote Perl for a year at uni before I found ASP

Pro: I'm the guy who is asked to write a regex in 2 minutes that saves someone an hour of writing code just to search GBs of logs.

Con: I'm the guy who is asked to translate regexes when they're found without a comment attached.

[Reply](#)

 [barubary@infosec.exchange](#) says:

[2023-02-06 12:55](#)

[@Edent](#) There's no space for comments.... unless you're using Perl, in which case you can use spaces, indentation, comments, etc in your regexes as you like. [#perl](#)

[perl](#)

[Reply](#)

 [Emily S](#) says:

[2023-02-06 13:05](#)

[@Edent](#) I'm pretty sure this one of those lessons that one must learn on the road to becoming a senior engineer.

[Reply](#)



 **Andy Mabbett** says:

[2023-02-06 13:28](#)

Not that I disagree with you, but I find that:

<https://regex101.com/>

helps - not least because it gives you a human-readable version of your regex in an adjacent pane.

[Reply](#)

 **Evgeny Kuznetsov** says:

[2023-02-06 13:54](#)

Wow! Had I known about that [RegEx101](#) thing, my life would have been much simpler. Then again, in that case I would probably not think Regular Expressions were *write-only code*, which they of course are.

[Reply](#)



 **James O'Malley** says:

[2023-02-06 14:06](#)

Couldn't agree more with this. And it's for this reason, rather than my being an idiot whose brain can't learn them, why my code always contain dozens of `str_replace()`s and `explode()`s.

[Reply](#)



 **Neil Young** says:

[2023-02-06 14:28](#)

Agree with this. There's a use case for "immediate" regular expressions as sort of keyboard shortcuts, but if they're going to be read by anyone else, they're obtuse.

[Reply](#)

 **AP** says:

[2023-02-06 14:42](#)

Perl introduced the possibility (through the `/x` and `/xx` modifiers) to ignore spaces, tabs, linefeeds and even comments in the regular expression body. It allows to "unconvolute" what would have been an unreadable mess into something understandable and maintainable. Every programming language should allow this in its regular expressions.

[Reply](#)

 **Sam J Sharpe** says:

[2023-02-06 15:07](#)

[@Edent](#) is it bad that I can read that regex and I think it's attempting to validate an email address?

[Reply](#)

➤ **Richard Morton says:**

[2023-02-06 16:06](#)

Thinking off the top of my head, is there a way to have a plain language style code that then gets compiled to a regex, and then include the plain language in the code comments or documentation?

[Reply](#)

➤ **Alan says:**

[2023-02-06 17:10](#)

Wholeheartedly agree! I shudder every time I come across a RegEx that *might* be the problem, even worse if it has no unit tests,

[Reply](#)

 [Pat Mächler](#)  **says:**

[2023-02-06 19:02](#)

"if/else and switch/case statements are the mark of a diseased mind."

Indeed!

While loops and gotos+labels are completely sufficient.

[Reply](#)

 [chessmango](#) **says:**

[2023-02-07 14:55](#)

[@Edent](#) I agree that a dedicated, less-clunky solution should be used where possible, but the sheer ubiquity of RegEx is what makes it invaluable as far as I'm concerned. Means I can toss a string into vscode to check for specifics, or use the same in some 'find' type box in a browser even (in some cases). It's awful to write and read, but it's there 😊

[Reply](#)

## Leave a Reply

Your email address will not be published. Required fields are marked \*

Comment \*

Name \*

Email \*

Website

Notify me of follow-up comments by email.

Notify me of new posts by email.

Post Comment

To respond on your own website, enter the URL of your response which should contain a link to this post's permalink URL. Your response will then appear (possibly after moderation) on this page. Want to update or remove your response? Update or delete your post and re-enter your post's URL again. ([Learn More](#))

URL/Permalink of your article

Ping me!

**Found this post useful? Click the icons to support this blog**



[More ways to support my blog](#)

## Get new posts by email

Enter your email address to subscribe to this blog and receive brand new posts by email.

(Or subscribe to this  [Atom Feed.](#))

Join 427 other subscribers.

## Search Blog Posts

  

## Explore The Archives

2023			2022		
<u>January</u> 31 posts	<u>February</u> 9 posts	March	<u>January</u> 30 posts	<u>February</u> 23 posts	<u>March</u> 15 posts
April	May	June	<u>April</u> 19 posts	<u>May</u> 19 posts	<u>June</u> 19 posts
July	August	September	<u>July</u> 19 posts	<u>August</u> 18 posts	<u>September</u> 12 posts
October	November	December	<u>October</u> 8 posts	<u>November</u> 30 posts	<u>December</u> 31 posts

2021			2020		
<u>January</u> 31 posts	<u>February</u> 28 posts	<u>March</u> 31 posts	<u>January</u> 31 posts	<u>February</u> 29 posts	<u>March</u> 31 posts
<u>April</u> 30 posts	<u>May</u> 31 posts	<u>June</u> 30 posts	<u>April</u> 30 posts	<u>May</u> 31 posts	<u>June</u> 30 posts
<u>July</u> 31 posts	<u>August</u> 31 posts	<u>September</u> 30 posts	<u>July</u> 31 posts	<u>August</u> 31 posts	<u>September</u> 30 posts
<u>October</u> 31 posts	<u>November</u> 30 posts	<u>December</u> 31 posts	<u>October</u> 31 posts	<u>November</u> 30 posts	<u>December</u> 31 posts

2019			2018		
<u>January</u> 31 posts	<u>February</u> 12 posts	<u>March</u> 17 posts	<u>January</u> 8 posts	<u>February</u> 4 posts	<u>March</u> 6 posts
<u>April</u> 12 posts	<u>May</u> 12 posts	<u>June</u> 10 posts	<u>April</u> 14 posts	<u>May</u> 5 posts	<u>June</u> 6 posts
<u>July</u> 7 posts	<u>August</u> 5 posts	<u>September</u> 6 posts	<u>July</u> 6 posts	<u>August</u> 13 posts	<u>September</u> 14 posts
<u>October</u> 14 posts	<u>November</u> 30 posts	<u>December</u> 17 posts	<u>October</u> 8 posts	<u>November</u> 30 posts	<u>December</u> 4 posts



2017			2016		
<u>January</u> 12 posts	<u>February</u> 9 posts	<u>March</u> 8 posts	<u>January</u> 10 posts	<u>February</u> 10 posts	<u>March</u> 11 posts
<u>April</u> 4 posts	<u>May</u> 10 posts	<u>June</u> 5 posts	<u>April</u> 9 posts	<u>May</u> 8 posts	<u>June</u> 9 posts
<u>July</u> 5 posts	<u>August</u> 6 posts	<u>September</u> 3 posts	<u>July</u> 6 posts	<u>August</u> 9 posts	<u>September</u> 4 posts
<u>October</u> 4 posts	<u>November</u> 30 posts	December	<u>October</u> 2 posts	<u>November</u> 30 posts	<u>December</u> 14 posts

2015			2014		
<u>January</u> 8 posts	<u>February</u> 11 posts	<u>March</u> 10 posts	<u>January</u> 13 posts	<u>February</u> 13 posts	<u>March</u> 14 posts
<u>April</u> 4 posts	<u>May</u> 9 posts	<u>June</u> 3 posts	<u>April</u> 14 posts	<u>May</u> 8 posts	<u>June</u> 7 posts
<u>July</u> 7 posts	<u>August</u> 9 posts	<u>September</u> 10 posts	<u>July</u> 9 posts	<u>August</u> 5 posts	<u>September</u> 5 posts
<u>October</u> 2 posts	<u>November</u> 30 posts	<u>December</u> 4 posts	<u>October</u> 1 post	<u>November</u> 30 posts	<u>December</u> 20 posts

2013			2012		
<u>January</u> 25 posts	<u>February</u> 17 posts	<u>March</u> 15 posts	<u>January</u> 14 posts	<u>February</u> 8 posts	<u>March</u> 13 posts
<u>April</u> 18 posts	<u>May</u> 11 posts	<u>June</u> 14 posts	<u>April</u> 15 posts	<u>May</u> 10 posts	<u>June</u> 16 posts
<u>July</u> 6 posts	<u>August</u> 14 posts	<u>September</u> 6 posts	<u>July</u> 8 posts	<u>August</u> 8 posts	<u>September</u> 6 posts
<u>October</u> 4 posts	<u>November</u> 30 posts	<u>December</u> 14 posts	<u>October</u> 6 posts	<u>November</u> 30 posts	<u>December</u> 31 posts

2011			2010		
<a href="#">January</a> 13 posts	<a href="#">February</a> 11 posts	<a href="#">March</a> 11 posts	<a href="#">January</a> 6 posts	<a href="#">February</a> 15 posts	<a href="#">March</a> 12 posts
<a href="#">April</a> 12 posts	<a href="#">May</a> 8 posts	<a href="#">June</a> 8 posts	<a href="#">April</a> 13 posts	<a href="#">May</a> 4 posts	<a href="#">June</a> 3 posts
<a href="#">July</a> 6 posts	<a href="#">August</a> 5 posts	<a href="#">September</a> 11 posts	<a href="#">July</a> 15 posts	<a href="#">August</a> 8 posts	<a href="#">September</a> 11 posts
<a href="#">October</a> 7 posts	<a href="#">November</a> 30 posts	<a href="#">December</a> 17 posts	<a href="#">October</a> 9 posts	<a href="#">November</a> 30 posts	<a href="#">December</a> 9 posts

2009			2008		
<a href="#">January</a> 1 post	<a href="#">February</a> 5 posts	<a href="#">March</a> 3 posts	<a href="#">January</a> 2 posts	February	<a href="#">March</a> 2 posts
<a href="#">April</a> 7 posts	<a href="#">May</a> 12 posts	<a href="#">June</a> 8 posts	<a href="#">April</a> 3 posts	<a href="#">May</a> 2 posts	June
<a href="#">July</a> 10 posts	<a href="#">August</a> 10 posts	<a href="#">September</a> 12 posts	<a href="#">July</a> 1 post	<a href="#">August</a> 3 posts	<a href="#">September</a> 1 post
<a href="#">October</a> 22 posts	<a href="#">November</a> 31 posts	<a href="#">December</a> 15 posts	<a href="#">October</a> 3 posts	<a href="#">November</a> 2 posts	<a href="#">December</a> 1 post

2007			2006		
January	February	March	January	February	March
April	May	June	<a href="#">April</a> 1 post	May	June
July	August	September	July	August	September
October	<a href="#">November</a> 4 posts	<a href="#">December</a> 5 posts	October	<a href="#">November</a> 1 post	December

2005			2004		
January	February	<a href="#">March 1 post</a>	January	February	March
April	May	June	April	<a href="#">May 5 posts</a>	<a href="#">June 3 posts</a>
July	August	<a href="#">September 1 post</a>	<a href="#">July 1 post</a>	August	September
October	November	December	October	November	December

2003			2002		
January	February	<a href="#">March 2 posts</a>	January	<a href="#">February 1 post</a>	March
April	May	June	<a href="#">April 3 posts</a>	May	June
July	August	September	July	August	September
October	November	December	October	November	December

2001			2000		
January	February	March	January	February	<a href="#">March 1 post</a>
April	May	June	April	May	June
<a href="#">July 1 post</a>	August	September	July	August	September
<a href="#">October 1 post</a>	November	December	October	<a href="#">November 1 post</a>	December

1999			1997		
January	February	March	<a href="#">January 1 post</a>	February	March
April	May	June	April	May	June
July	August	<a href="#">September 1 post</a>	July	August	September
October	November	<a href="#">December 1 post</a>	October	November	December

1995			1987		
January	February	<a href="#">March 1 post</a>	January	February	March
April	May	June	April	May	June
July	August	September	July	August	September
October	November	December	October	November	<a href="#">December 1 post</a>

© Terence Eden [Contact Me](#) [Subscribe](#) [Citations](#) [Support My Blog](#)

[Bespoke Computing Consultancy](#) [About Me](#)

ISSN 2753-1570

