

読みにくいソースコードの理解に影響を与える要因の分析

杉山裕一 森崎修司 (名古屋大学)

・ 従来研究

Cソースコード難読化の国際コンテスト (IOCCC: International Obfuscated C Code Contest) の受賞作品を分析し、プログラマが読みにくいコードパターンを分析しています[1]. 難読化したソースコードと同一の実行結果を得られるソースコードを73人の学生に読んでもらい、実行結果の正解率を比較しています. 表1はその結果であり、上にあるコードパターンほど難読化したソースコードの正解率が低かったこと(効果量が大きかったこと)を表しています.

コードパターン	例	効果量	p値
文字エンコーディング	<code>printf("%d", 013)</code>	0.63	2.93e-14
ステート中のプリプロセッサ	<code>int V1 = 1 #define M1 1 +1;</code>	0.54	8.53e-11
マクロ演算子の優先順位	<code>#define M1 64-1 2*M1</code>	0.53	1.77e-07
代入文	<code>V1 = V2 = 3;</code>	0.52	3.78e-10
制御文中での論理演算	<code>V1 && F2();</code>	0.48	5.62e-09
インクリメント演算子の後置き	<code>V1 = V2++;</code>	0.45	6.98e-08
型変換	<code>(double)(3/2)</code>	0.42	5.17e-07
逆添字	<code>1["abc"]</code>	0.40	1.52e-06
「?」演算子	<code>V2 = (V1==3)?2:V2</code>	0.36	1.74e-05
中置演算子の優先順位	<code>0 && 1 2</code>	0.33	5.90e-05
コンマ演算子	<code>V3 = (V1+=1, V1)</code>	0.30	2.46e-04
インクリメント演算子の先置き	<code>V1 = ++V2;</code>	0.28	6.89e-04
制御文中での数値演算	<code>if (4 % 2)</code>	0.24	4.27e-03
変数の再目的化	<code>argc = 7;</code>	0.22	6.66e-03
中括弧の省略	<code>if (V) F(); G();</code>	0.22	8.64e-03

表1:正解率が低かったコードパターン

・ 本研究の狙い

文献[1]で示されたコードパターンのうち、MISRA-C 2014にも該当するコードパターン9種類を対象に、以下を分析しようとしています.

- ・ コードパターンの知識を持っているプログラマの正解率
- ・ プログラマの経験が正解率や読解時間に影響するか
- ・ 意思決定方法(モデルベース/モデルフリー)[2]の傾向はソースコード読解にも同様の影響を与えるか

図1はコードパターンの一例です. 左側の知識を問う問題を解けたプログラマはコードパターンを知っていると判断し、右側の知識の活用を問う問題での正解率や読解時間を分析します.

図1:ソースコードの例

知識を問う問題

```
void main() {  
    int V1 = 4;  
  
    if (V1 == 3) V1++; V1++;  
  
    printf("%d\n", V1);  
}
```

知識の活用を問う問題

```
void main() {  
    int V1;  
    int V2;  
    int V3 = 0;  
    int V4 = 1;  
  
    for(V1=0;V1<5;V1++){  
        for(V2=0;V2<3;V2++){  
            V3++;V4++;  
        }  
    }  
  
    printf("%d, %d\n", V3, V4);  
}
```

・ 協力者募集

右のQRコードからもしくは以下のURLにアクセスしてください.

全体で40~60分程度かかります. 文献[2]の内容の調査のため、Windows, Macで動作するmillisecond社製のInquisitというツールのインストールが必要です.

<https://modelcheckform-241916.appspot.com>



参考文献

- [1] D. Gopstein, J. Iannacone, Y. Yan, L. DeLong, Y. Zhuang, M.K.-C. Yeh, J. Cappos, Understanding Misunderstandings in Source Code, In Proc. of the 11th Joint Meeting on Foundations of Software Engineering, pp. 129-139(2017)
- [2] N. D. Daw, S. J. Gershman, B. Seymour, P. Dayan, R. J. Dolan, Model-Based Influences on Humans' Choices and Striatal Prediction Errors, Neuron, vol. 69, no. 6, pp. 1204-1215(2011)

