## Giving nightmares to reverse engineers

JULES POON

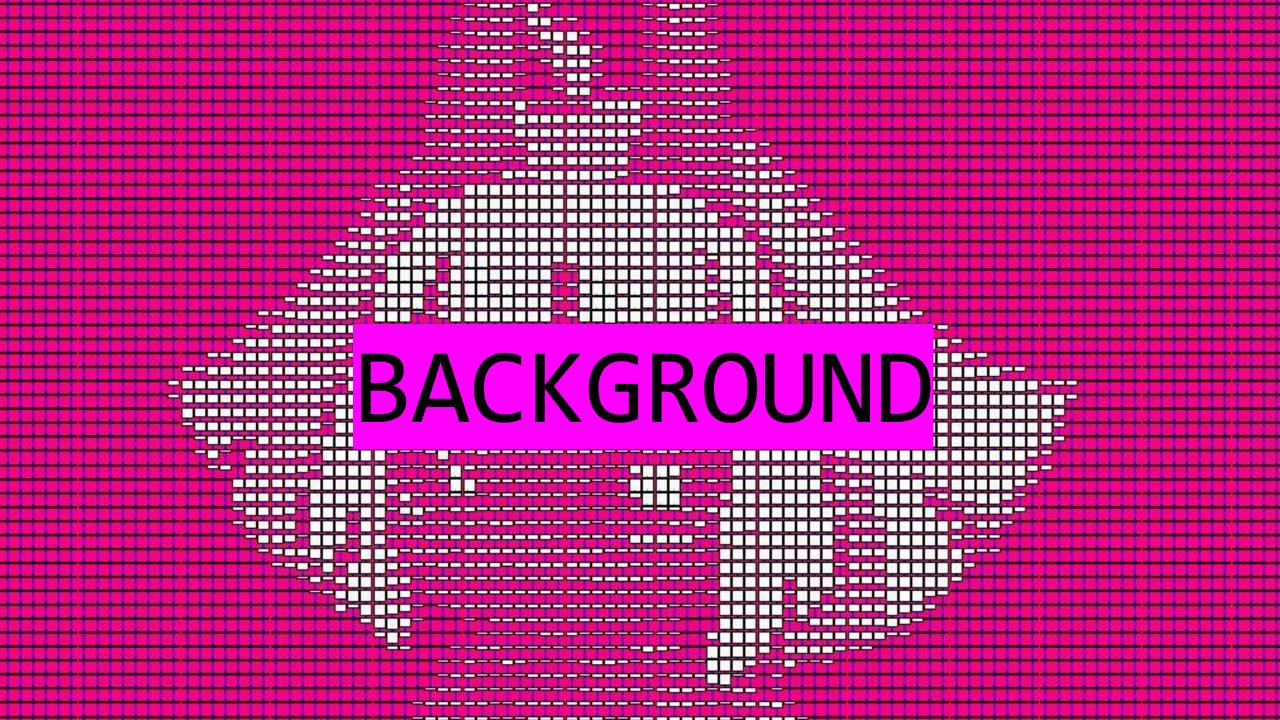
GEEKCAMP 2023

#### important information about me

- 1. Student @ Somewhere
- 2. Likes Math and PL

# BIGENALS. for Intern

# ARTFUSCATOR A PRODUCT PITCH



#### PROGRAMS: HOW THEY WORK

```
#if defined(RANDALL WAS HERE)
# define Py UNREACHABLE() \
   Py_FatalError( \
        "If you're seeing this, the code is in what I thought was\n" \
        "an unreachable state.\n\n" \
        "I could give you advice for what to do, but honestly, why\n" \langle
        "should you trust me? I clearly screwed this up. I'm writing\n" \
        "a message that should never appear, yet I know it willn''
        "probably appear someday.\n\n" \
        "On a deep level, I know I'm not up to this task.\n" \
        "I'm so sorry.n" \
        "https://xkcd.com/2200")
#elif defined(Py DEBUG)
# define Py UNREACHABLE() \
   Py FatalError( \
        "We've reached an unreachable state. Anything is possible.n" \
        "The limits were in our heads all along. Follow your dreams.n" \
        "https://xkcd.com/2200")
#elif defined(__GNUC__) && (__GNUC__ > 4 || (__GNUC__ == 4 && __GNUC_MINOR__
>= 5))
# define Py_UNREACHABLE() __builtin_unreachable()
#elif defined( clang ) || defined( INTEL COMPILER)
# define Py UNREACHABLE() builtin unreachable()
```

<---C PROGRAM (curtursy of cpython)

#### PROGRAMS: HOW TI

#### ---> Machine code In a PE file X8664

#### Can run on most intel CPUs

00001260	48	89	40	24	08	48	83	EC	58	48	8B	44	24	60	48	89	н		L.	\$		н			х	н		D	\$		н	
30001270								E1																						\$		
0001280								FF																								
0001290	32	C0	EB	52	48	8B	44	24	28	48	8B	4C	24	38	48	2B														8		
300012A0	C8	48	8B	C1	48	89	44	24	40	48	8B	54	24	40	48	8B														@		
900012B0	4C	24	28	E8	E8	F9	FF	FF	48	89	44	24	30	48	83	7C														H		
900012C0	24	30	00	75	04	32	C0	EB	1D	48	8B	44	24	30	8B	40	\$	0		u		2				н		D	\$	0		@
900012D0	24	25	00	00	00	80	85	CØ	74	04	32	C0	EB	<b>0</b> 8	BØ	01																
300012E0	EB	04	32	C0	EB	00	48	83	C4	58	C3	сс	сс	сс	сс	сс			2				н			х						
000012F0	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	СС																
0001300	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс																
00001310	88	4C	24	<b>0</b> 8	48	83	EC	28	E8	92	F1	FF	FF	85	C0	75		L	\$		н			(								u
00001320	02	EB	17	0F	B6	44	24	30	85	C0	74	02	EB	0C	33	C0						D	\$	0			t				3	
00001330	48	8D	0D	21	72	00	00	48	87	01	48	83	C4	28	C3	СС	н			I	r			Н			Н			(		
00001340	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	СС																
00001350	88	54	24	10	88	4C	24	08	48	83	EC	28	0F	B6	05	FD		т	\$			L	\$		н			(				
00001360	71	00	00	85	C0	74	0D	0F	B6	44	24	38	85	C0	74	04	q					t				D	\$	8			t	
00001370	BØ	01	EB	16	0F	B6	4C	24	30	E8	BD	F1	FF	FF	0F	B6							L	\$	0							
00001380	4C	24	30	E8	D7	F0	FF	FF	BØ	01	48	83	C4	28	C3	СС	L	\$	0								Н			(		
0001390	сс	СС	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	СС																
00013A0	48	89	4C	24	<b>0</b> 8	48	83	EC	48	48	8B	05	B8	71	00	00	н		L	\$		Н			н	н				q		
000013B0	48	89	44	24	30	48	83	7C	24	30	FF	75	2C	48	8B	4C	н		D	\$	0	Н		I	\$	0		u		н		L
000013C0	24	50	E8	05	1A	00	00	85	C0	75	0C	48	8B	44	24	50	\$	Ρ								u		н		D	\$	Ρ
000013D0	48	89	44	24	20	EB	<b>0</b> 9	48	C7	44	24	20	00	00	00	00	н		D	\$				Н		D	\$					
000013E0	48	8B	44	24	20	EB	31	EB	2F	48	8B	54	24	50	48	8D	н		D	\$			1		/	Н		т	\$	Ρ	н	
000013F0	0D	73	71	00	00	E8	C6	19	00	00	85	C0	75	0C	48	8B			q										u		Н	
30001400	44	24	50	48	89	44	24	28	EB	<b>0</b> 9	48	C7	44	24	28	00	D	\$	Ρ	н		D	\$	(			н		D	\$	(	
30001410	00	00	00	48	8B	44	24	28	48	83	C4	48	C3	сс	сс	СС				н		D	\$	(	Н			н				
00001420	СС	сс	сс	сс	сс	сс	сс	сс	сс	сс	СС	сс	сс	сс	сс	CC																
30001430	сс	СС	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	СС																
00001440	48	89	4C	24	<b>0</b> 8	48	83	EC	38	48	8B	05	30	71	00	00														q		
00001450	48	89	44	24	20	48	83	7C	24	20	FF	75	0E	48	8B	4C														H		
00001460	24	40	E8	6B	19	00	00	EB	1D	EB	1B	48	8B	44	24	40														D		
00001470	48	89	44	24	28	48	8B	54	24	28	48	8D	ØD	FF	70	00	Н		D	\$	(	Н		Т	\$	(	Н				р	

#### **REVERSE ENGINEERS**

#### People who read machine code to understand a **binary's functionality**

- Malware researchers
- Software crackers
- Masochists

#### REVERSE

#### People w understa function

- Malware
- Softwar
- Masoch

001260	48	89	4C	24	<b>0</b> 8	48	83	EC	58	48	8B	44	24	60	48	89	н		L	\$		н			х	Н		D	\$		Н	
001270	44	24	38	48	8D	05	86	E1	FF	FF	48	89	44	24	28	48	D	\$	8	Н							н		D	\$	(	н
001280	8B	4C	24	28	E8	37	FB	FF	FF	0F	B6	C0	85	C0	75	04		L	\$	(		7									u	
001290	32	CØ	EB	52	48	8B	44	24	28	48	8B	4C	24	38	48	2B	2			R	н		D	\$	(	н		L	\$	8	н	+
0012A0	C8	48	8B	C1	48	89	44	24	40	48	8B	54	24	40	48	8B		Н			Н		D	\$	@	Н		Т	\$	@	Н	
0012B0	4C	24	28	E8	E8	F9	FF	FF	48	89	44	24	30	48	83	7C	L	\$	(						н		D	\$	0	н		
0012C0	24	30	00	75	04	32	C0	EB	1D	48	8B	44	24	30	8B	40	\$	0		u		2				н		D	\$	0		a
0012D0	24	25	00	00	00	80	85	C0	74	04	32	C0	EB	<b>0</b> 8	BØ	01	\$	%							t		2					
0012E0	EB	04	32	C0	EB	00	48	83	C4	58	C3	сс	сс	сс	сс	сс			2				н			х						
0012F0	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс																
001300	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс																
001310	88	4C	24	08	48	83	EC	28	E8	92	F1	FF	FF	85	C0	75		L	\$		н			(								u
001320	02	EB	17	ØF	B6	44	24	30	85	C0	74	02	EB	0C	33	C0						D	\$	0			t				3	
001330	48	8D	0D	21	72	00	00	48	87	01	48	83	C4	28	C3	сс	н			!	r			н			н			(		
001340	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс																
001350	88	54	24	10	88	4C	24	08	48	83	EC	28	ØF	B6	05	FD		т	\$			L	\$		н			(				
001360	71	00	00	85	C0	74	0D	0F	B6	44	24	38	85	C0	74	04	q					t				D	\$	8			t	
001370	BØ	01	EB	16	0F	B6	4C	24	30	E8	BD	F1	FF	FF	0F	B6							L	\$	0							
001380	4C	24	30	E8	D7	F0	FF	FF	BØ	01	48	83	C4	28	C3	сс	L	\$	0								н			(		
001390	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс																
0013A0	48	89	4C	24	<b>0</b> 8	48	83	EC	48	48	8B	05	B8	71	00	00	н		L	\$		н			н	н				q		
0013B0	48	89	44	24	30	48	83	7C	24	30	FF	75	2C	48	8B	4C	Н		D	\$	0	н		L	\$	0		u		н		L
0013C0	24	50	E8	05	1A	00	00	85	C0	75	0C	48	8B	44	24	50	\$	Р								u		н		D	\$	Р
0013D0	48	89	44	24	20	EB	<b>0</b> 9	48	C7	44	24	20	00	00	00	00	н		D	\$				Н		D	\$					
0013E0	48	8B	44	24	20	EB	31	EB	2F	48	8B	54	24	50	48	8D	Н		D	\$			1		/	н		т	\$	Ρ	н	
0013F0	ØD	73	71	00	00	E8	C6	19	00	00	85	C0	75	ØC	48	8B			q										u		н	
001400	44	24	50	48	89	44	24	28	EB	<b>0</b> 9	48	C7	44	24	28	00	D	\$	Ρ	н		D	\$	(			н		D	\$	(	
001410	00	00	00	48	8B	44	24	28	48	83	C4	48	C3	сс	сс	СС				н		D	\$	(	н			Н				
001420	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	СС																
001430	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс	сс																
001440	48	89	4C	24	08	48	83	EC	38	48	8B	05	30	71	00	00	Н		L	\$		н			8	Н			0	q		
001450	48	89	44	24	20	48	83	7C	24	20	FF	75	0E	48	8B	4C	Н		D	\$		н		I	\$			u		н		L
001460	24	40	E8	6B	19	00	00	EB	1D	EB	1B	48	8B	44	24	40	\$	@		k								н		D	\$	a
001470	48	89	44	24	28	48	8B	54	24	28	48	8D	ØD	FF	70	00	Н		D	\$	(	н		т	\$	(	н				D	

code to

#### Binary/machine isntructions can be **DISSASSEMBLED** into a

# **1. HIGH LEVEL 2. HUMAN-READABLE**programming language known as

## ASSEMBLY

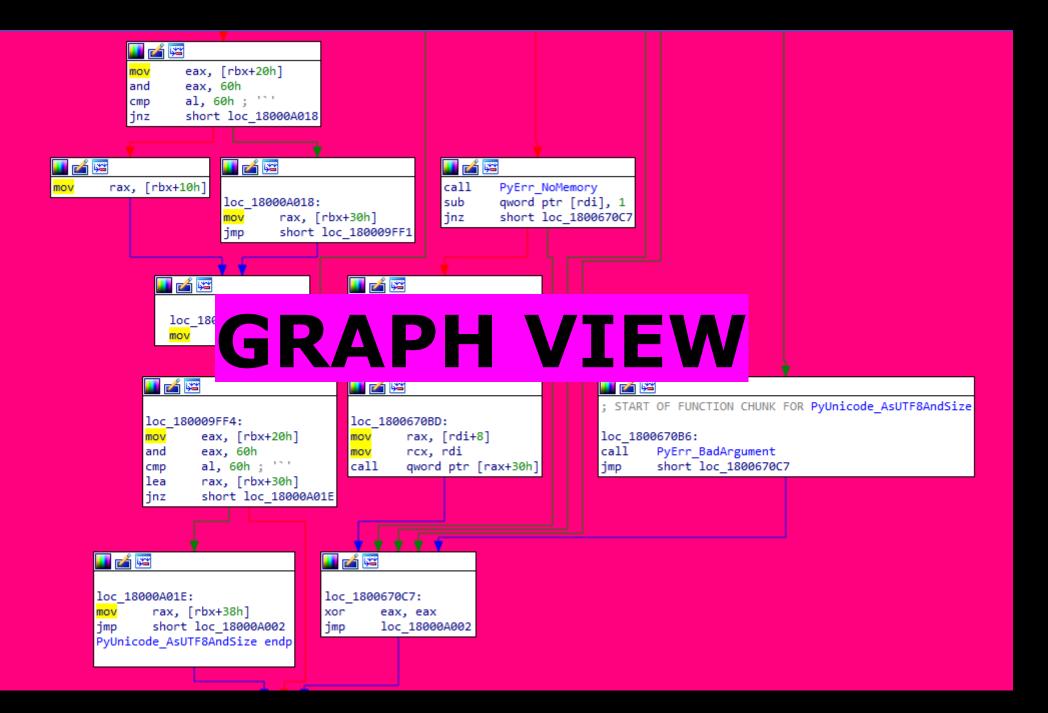
```
.text:0000000180009F94 ; FUNCTION CHUNK AT .text:00000001800670B6 SIZE 0000008C BYTES
.text:0000000180009F94
.text:0000000180009F94
                                                 [rsp+arg 0], rbx
                                        mov
.text:0000000180009F99
                                        mov
                                                [rsp+arg 8], rsi
                                        push
                                                rdi
.text:0000000180009F9E
.text:0000000180009F9F
                                        sub
                                                rsp, 20h
text:0000000180009FA3
                                        moν
                                                rax, [rcx+8]
.text:0000000180009FA7
                                        mov
                                                rsi, rdx
text:0000000180009FAA
                                        mov
                                                rbx, rcx
                                                dword ptr [rax+0A8h], 10000000h
text:0000000180009FAD
                                        test
.text:0000000180009FB7
                                        iz
                                                loc 1800670B6
                                                byte ptr [rcx+20h], 80h
text:0000000180009FBD
                                        test
text:0000000180009FC1
                                        iz
                                                loc 1800670CE
text:0000000180009FC7
text:000000180009FC7 loc 180009FC7:
                                                                 ; CODE XREF: PyUnicode_AsUTF8AndSize+5D144↓j
text:0000000180009FC7
                                        mov
                                                eax, [rbx+20h]
text:0000000180009FCA
                                                eax, 60h
                                        and
                                                           15.1
                                                al, 60h ;
text:0000000180009FCD
                                        CMD
text:0000000180009FCF
                                        lea
                                                rax, [rbx+30h]
text:0000000180009FD3
                                        inz
                                                short loc 18000A012
text:0000000180009FD5
text:000000180009FD5 loc 180009FD5:
                                                                 ; CODE XREF: PyUnicode AsUTF8AndSize+82↓j
text:0000000180009FD5
                                        test
                                                rax, rax
.text:0000000180009FD8
                                        jz
                                                loc 1800670DD
.text:0000000180009FDE
text:0000000180009FDE loc 180009FDE:
                                                                 ; CODE XREF: PyUnicode AsUTF8AndSize+5D198↓j
text:0000000180009FDE
                                                                 ; PyUnicode AsUTF8AndSize+5D1A9↓j
text:0000000180009FDE
                                                rsi, rsi
                                        test
                                                short loc 180009FF4
.text:0000000180009FE1
                                        jz
.text:0000000180009FE3
                                                eax, [rbx+20h]
                                        mov
text:0000000180009FE6
                                        and
                                                eax, 60h
                                                al, 60h ; '`'
.text:0000000180009FE9
                                        CMD
.text:0000000180009FEB
                                        jnz
                                                short loc 18000A018
text:0000000180009FED
                                        mov
                                                rax, [rbx+10h]
text:0000000180009FF1
text:0000000180009FF1 loc 180009FF1:
                                                                 ; CODE XREF: PyUnicode_AsUTF8AndSize+88↓j
.text:0000000180009FF1
                                                [rsi], rax
                                        mov
.text:0000000180009FF4
                                                                 ; CODE XREF: PyUnicode_AsUTF8AndSize+4D^j
.text:0000000180009FF4 loc 180009FF4:
text:0000000180009FF4
                                        mov
                                                eax, [rbx+20h]
                                                eax, 60h
text:0000000180009FF7
                                        and
.text:0000000180009FFA
                                        cmp
                                                al, 60h ;
                                                rax, [rbx+30h]
text:0000000180009FFC
                                        lea
text:000000018000A000
                                        inz
                                                short loc 18000A01E
```

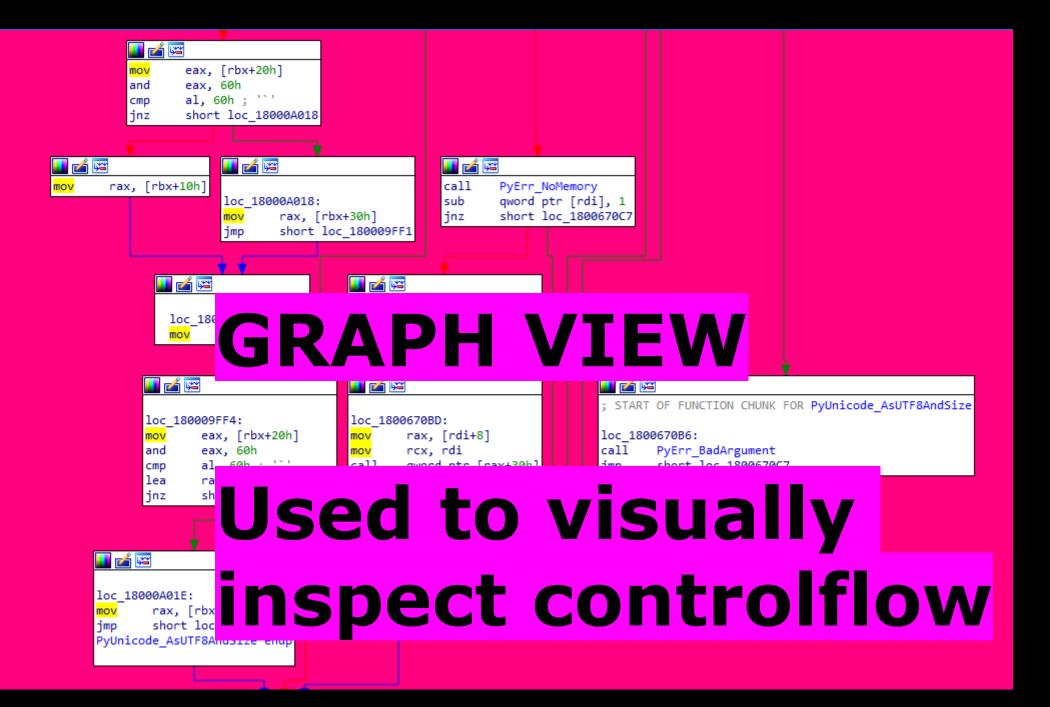
#### REVERSE ENGINEERS (RE)

# Tools to help them read more code at once

- IDA:

Interactive DisAssembler THe adobe photoshop for RE

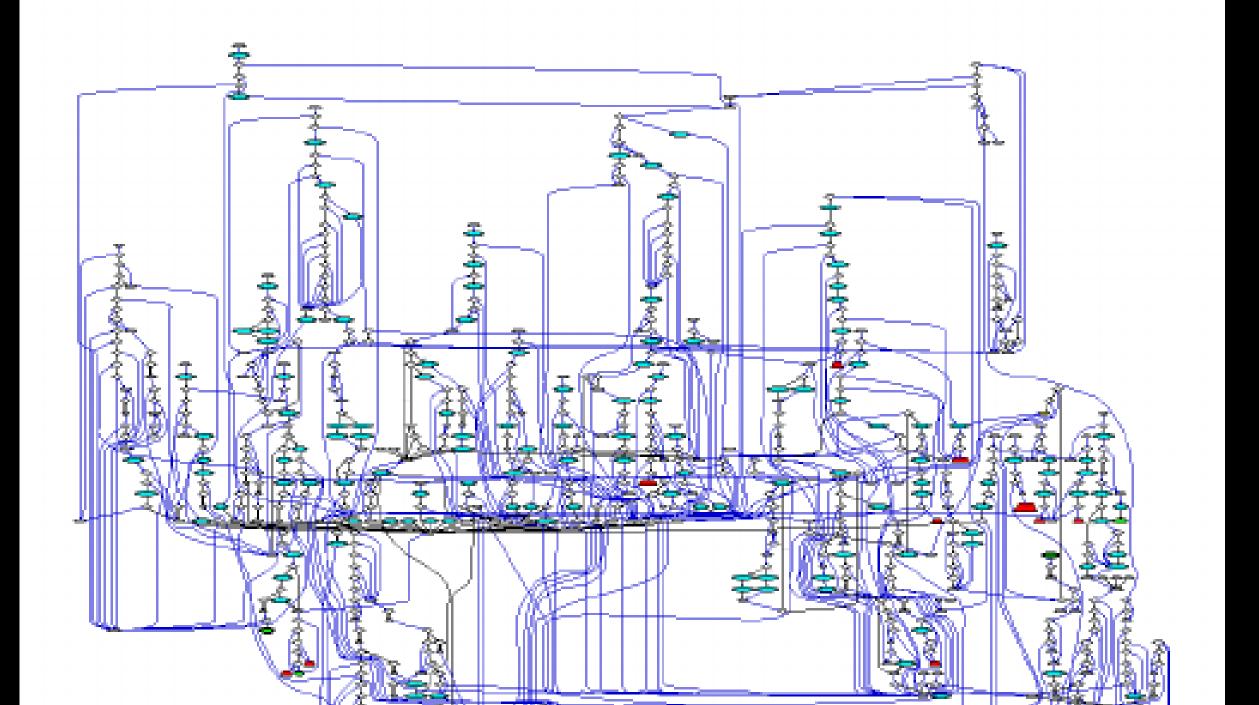


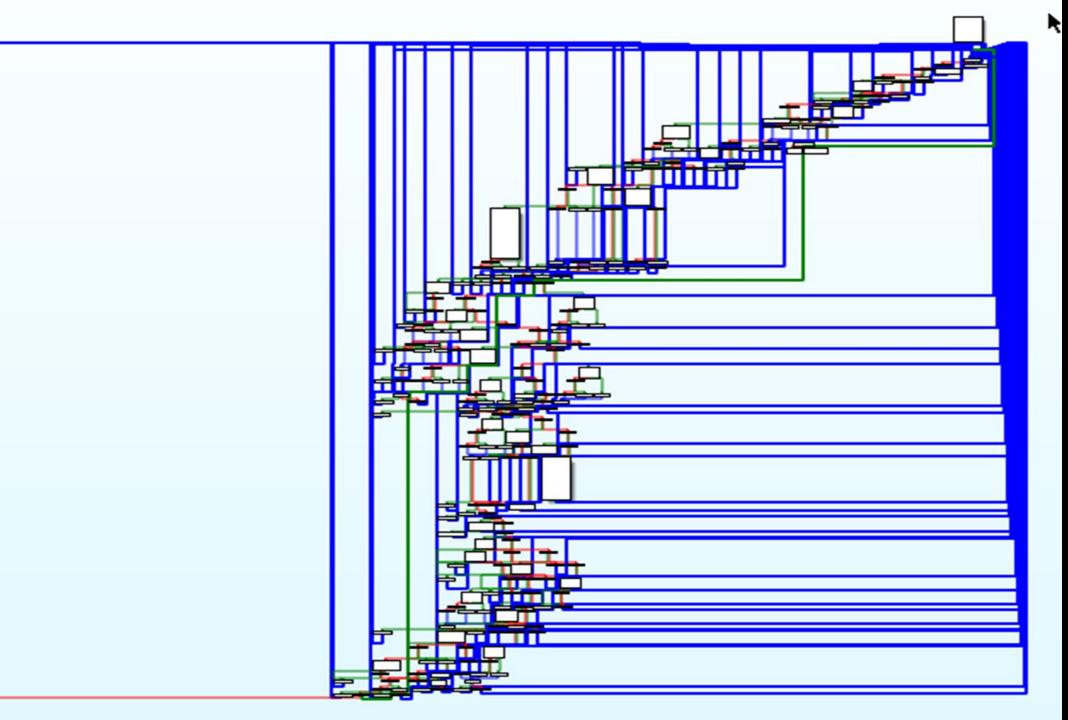


#### REVERSE ENGINEERS (RE)

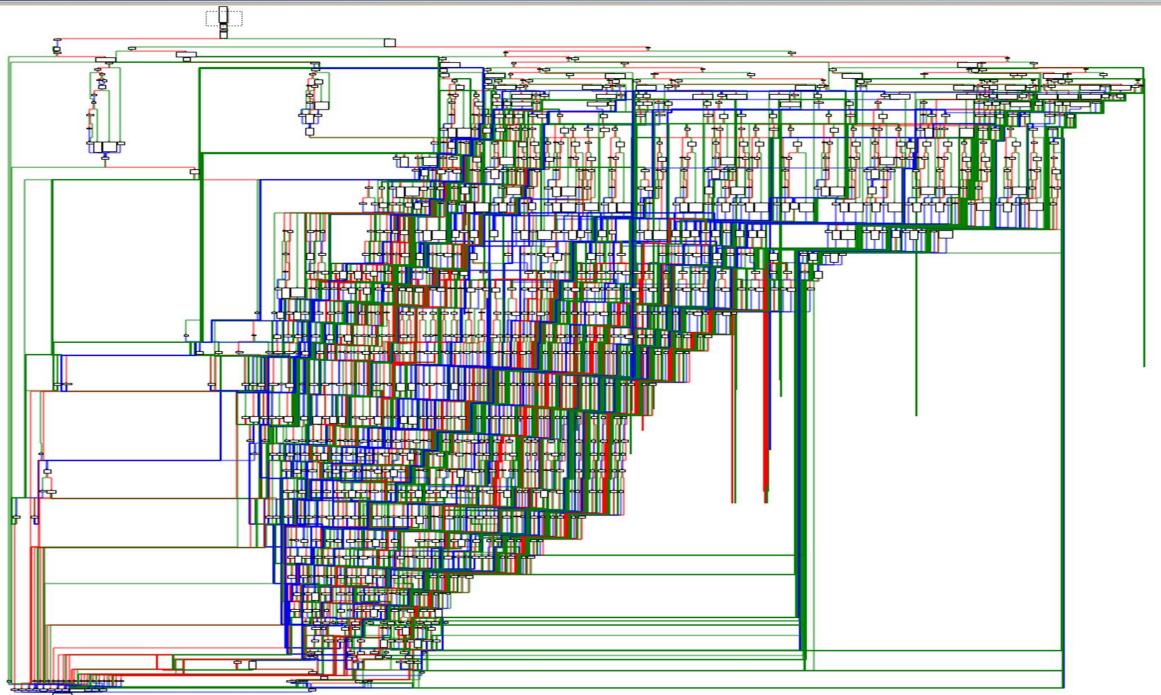
#### Reverse Engineering is a cat and mouse game

- Bad actors/Shitty companies keep making reversing harder





Graph overview



×

#### REVERSE ENGINEERS (RE)

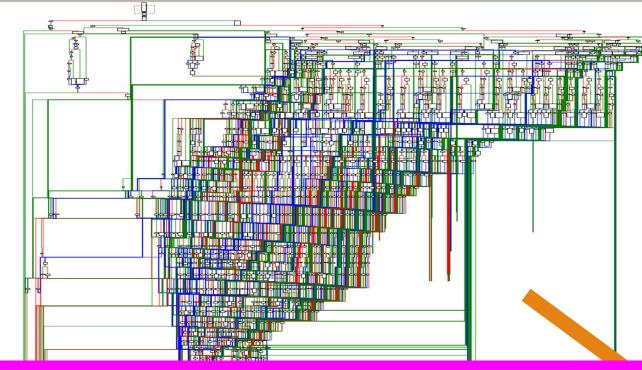
#### **Reverse Engineers are selectively breeded to**

achieve higher and higher levels of psychological resistance

# Introducing ARTFUSCATOR

# Introducing ARTFUSCATOR

#### A NOVEL APPROACH TO PSYCHOLOGICAL WARFARE AGAINST REVERSE ENGINEERS



#### What's better than complex control flow

### flow Than 'no' controlflow at all

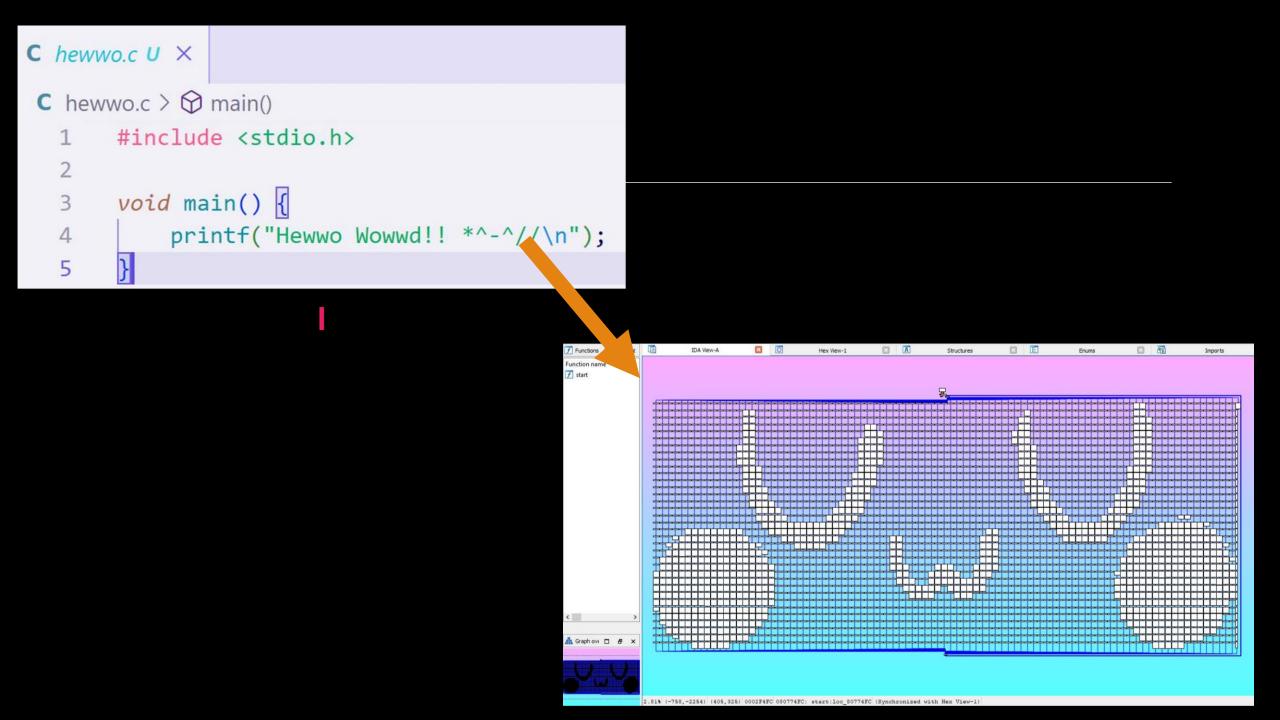
#### A Radically New Experience



#### Artfuscator

1. A C Compiler

2. Compiles entire program into A SINGLE CONTROLFLOW GRAPH ART



#### Motivation

#### Motivation

- covid

#### PREVIOUS WORKD **HITT** ------

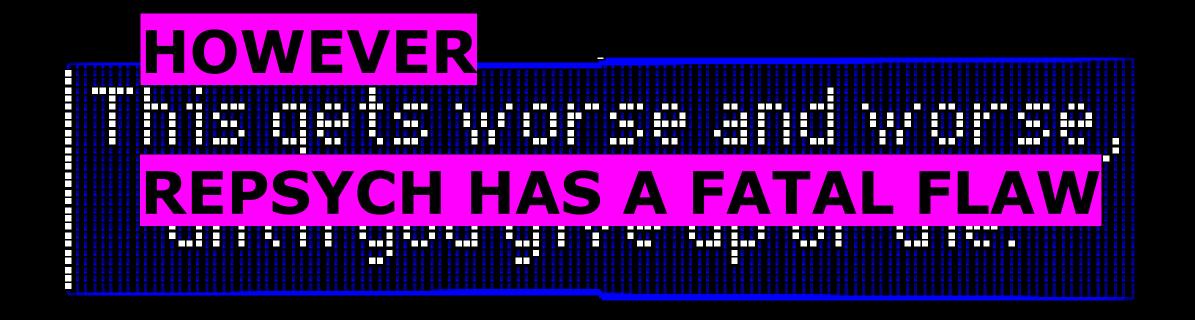
#### **REPSYCH by christopher domas**



#### **REPSYCH by XOREAXEAXEAX**

## 1. Pioneer in Psychological Warfare in Reverse Engineering 2. Invented controlflow graph art

#### **REPSYCH by XOREAXEAXEAX**



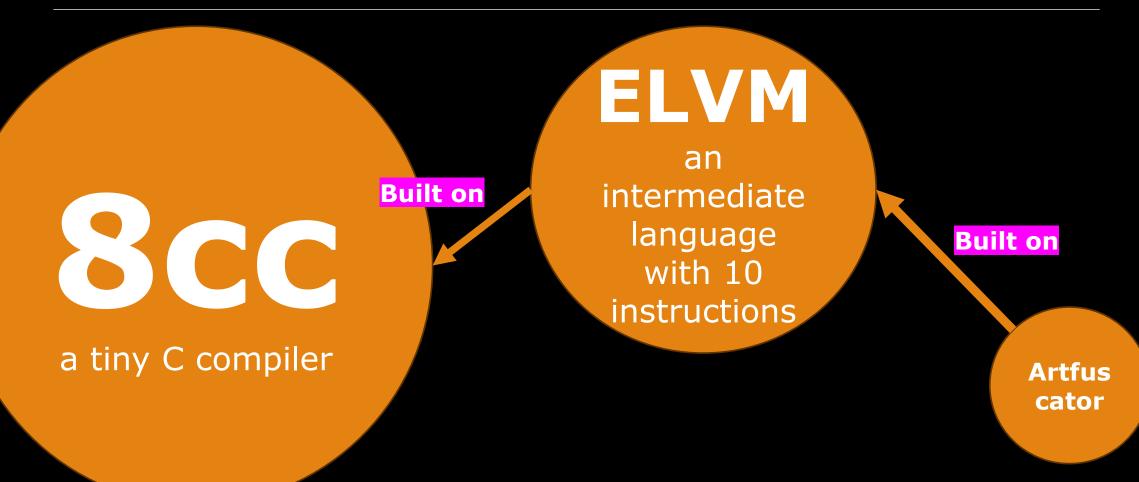
#### **REPSYCH by XOREAXEAXEAX**

#### THE GRAPH ART HAS NO FUNCTIONALITY

### ARTFUSCAT OR FIXES **THIS**

## "Monumental achievement in compiler engineering" (Source: me)

"Monumental achievement in compiler engineering" (Source: me) Artfuscator is built on the work of **GIANTS** 

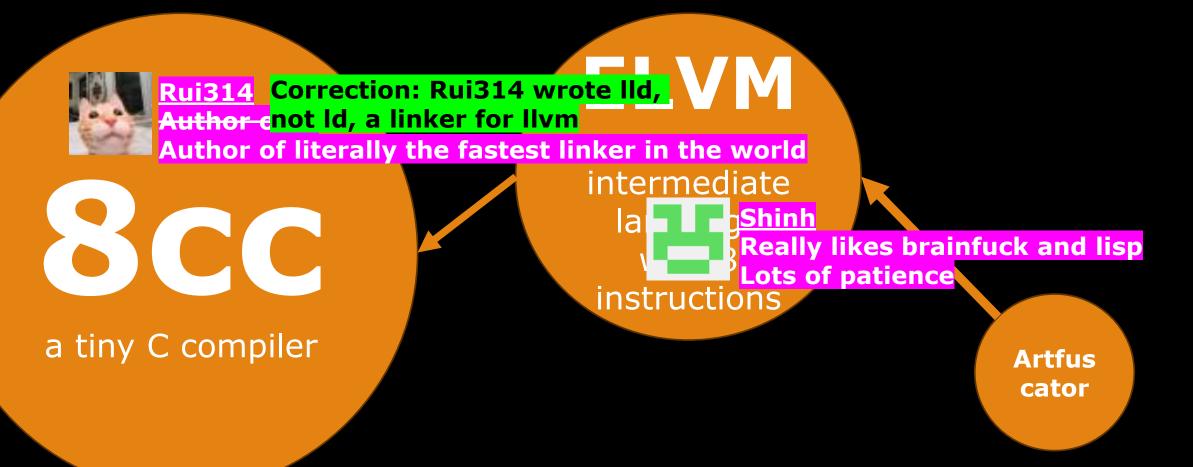


Rui314 Correction: Rui314 wrote IId, Author enot Id, a linker for Ilvm Author of literally the fastest linker in the world

a tiny C compiler

intermediate language with 10 instructions

Artfus cator



Rui314 Correction: Rui314 wrote IId, Author cnot Id, a linker for Ilvm Author of literally the fastest linker in the world

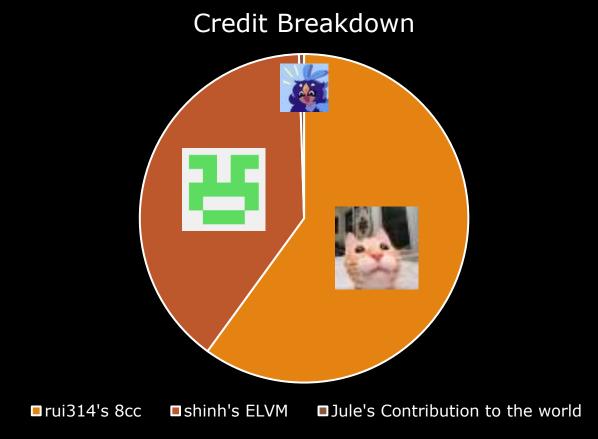
BCC

a tiny C compiler

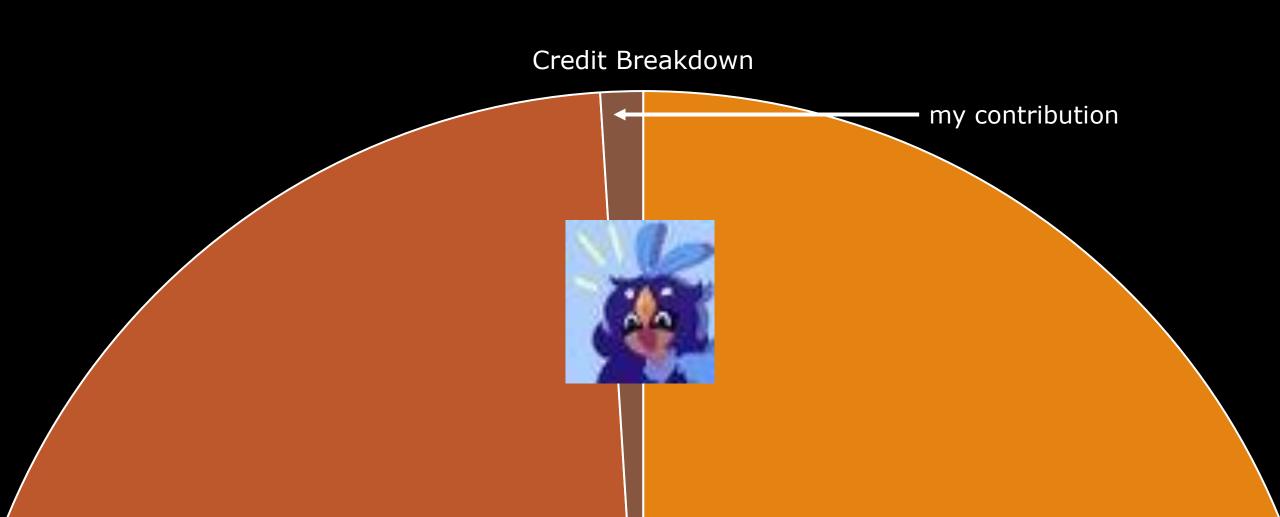
late and set of patience instructions

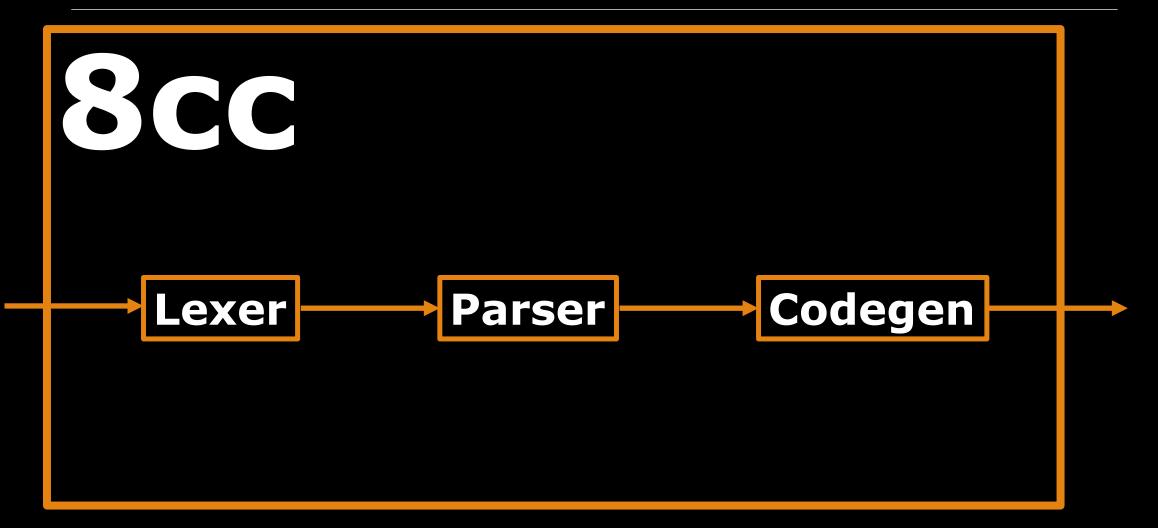
intermediate

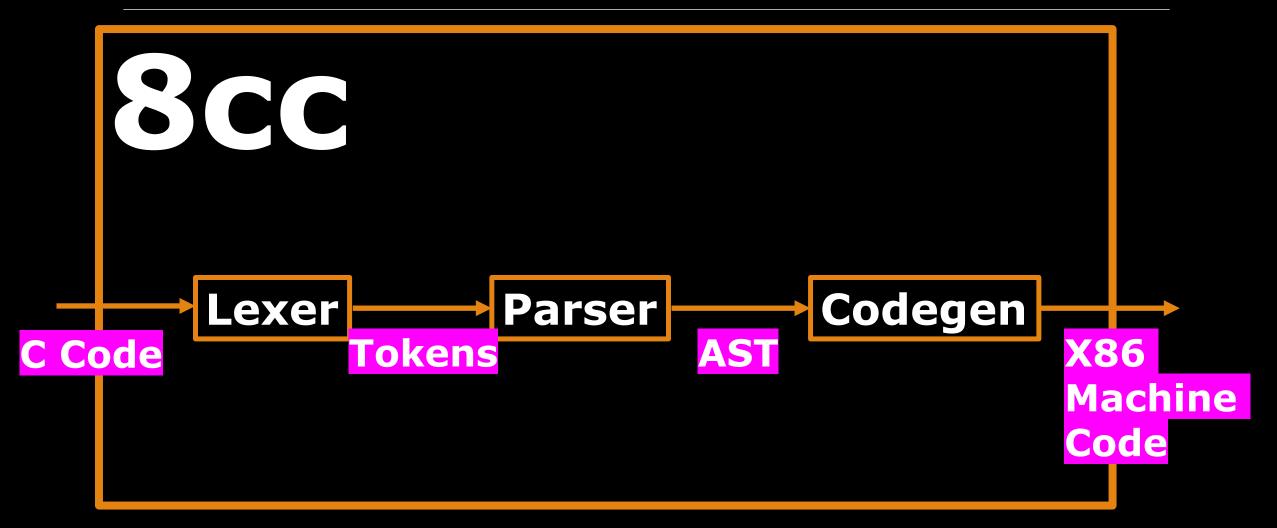
#### Credit breakdown

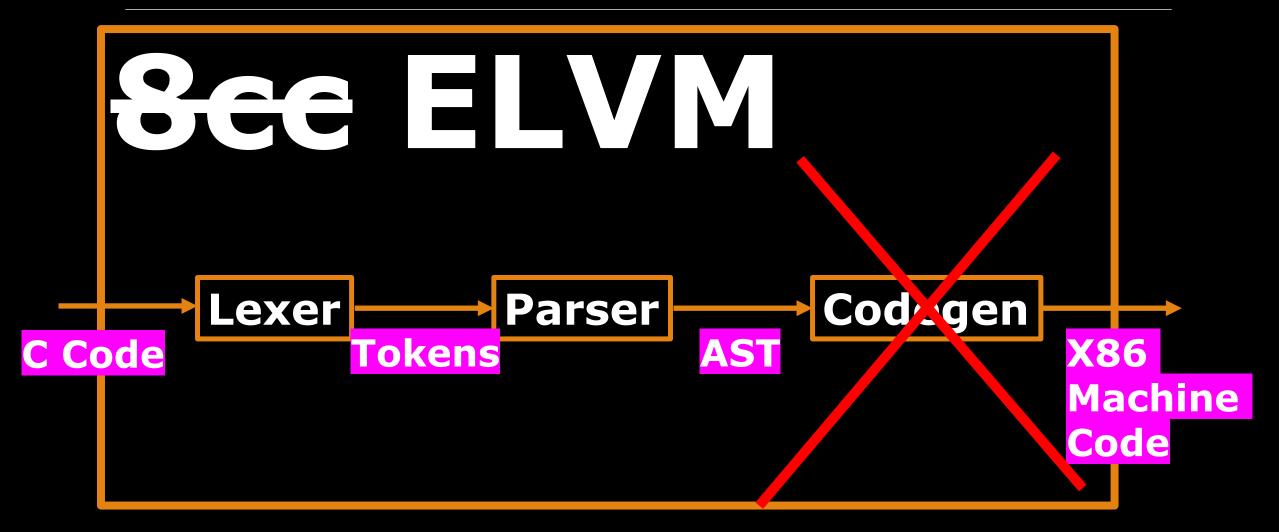


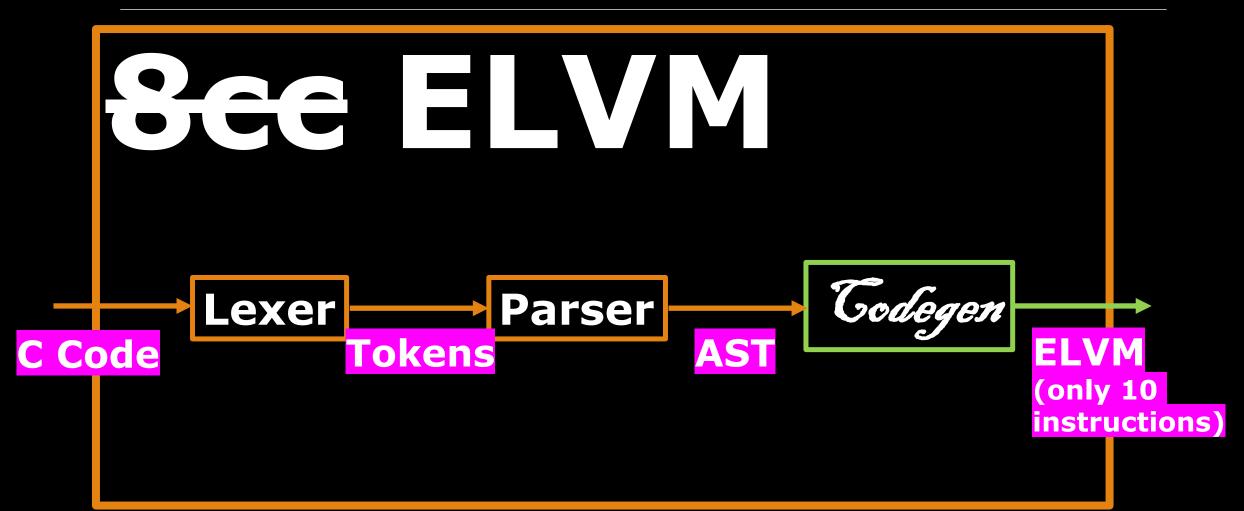
#### Credit breakdown











# **Bee ELVA Artfus**

instructions



## Artfuscator



Codegen ELVM (only 10 instructions)

## Artfuscator



#### **MY OWN Artfuscator Internals** CODEGEN Job of my codegen: Convert a program made up of **10 instructions into an artful** controlflow graph mov, add, sub, load, store, setcc, jcc, putc, getc, exit

#### **Problem:**

MÝ OWN CODEGEN

## To control the graph fully we can't "have" controlflow

#### Can't have the program's controlflow interferering with our art



#### MÝ OWN CODEGEN

## Artfuscator Internals Solution?

MÝ OWN CODEGEN

## Do complex code transformations to remove controlflow entirely?

## Spend 100s of hours crafting the perfect passes?

#### **Solution?**

MÝ OWN CODEGEN

## Do complex code transformations to ren To much effortely?

Spend 100s of hours crafting the perfect passes?

#### **Solution:**

MÝ OWN CODEGEN

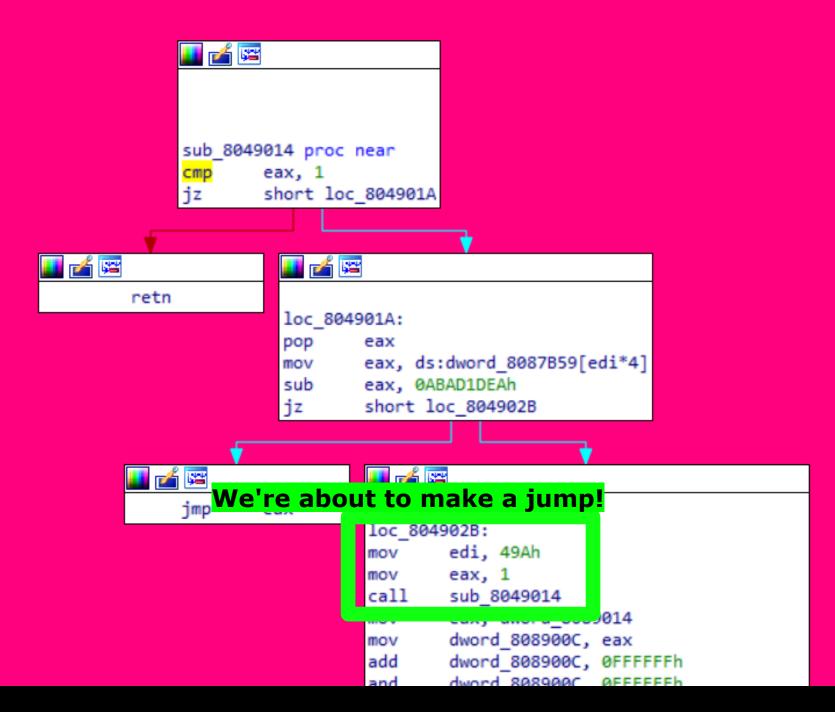
## We trick disassemblers by converting all jmp into call

#### **Solution:**

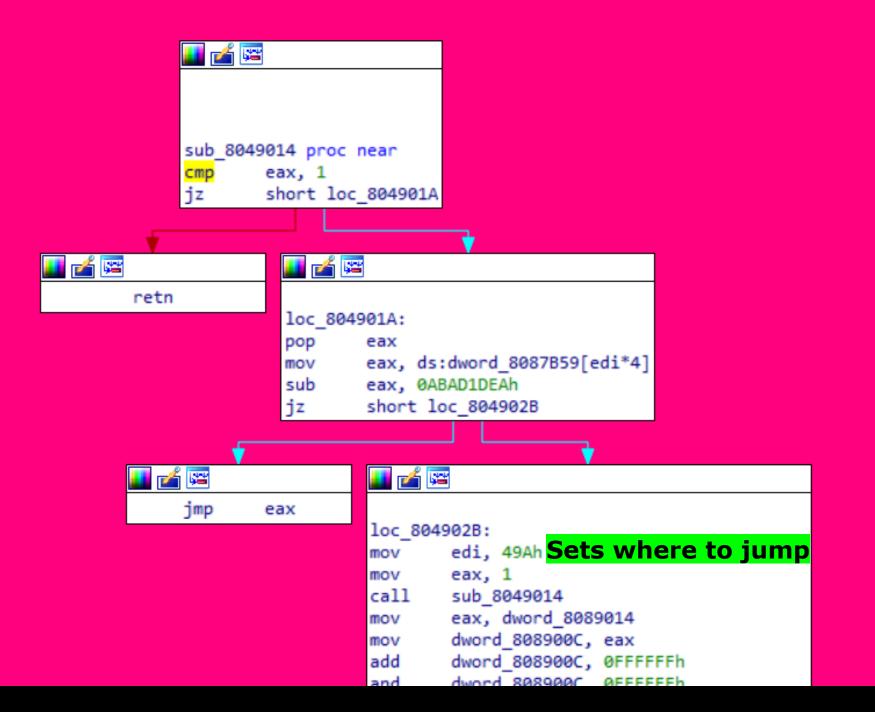
MÝ OWN CODEGEN

## We trick disassemblers by converting all jmp into call

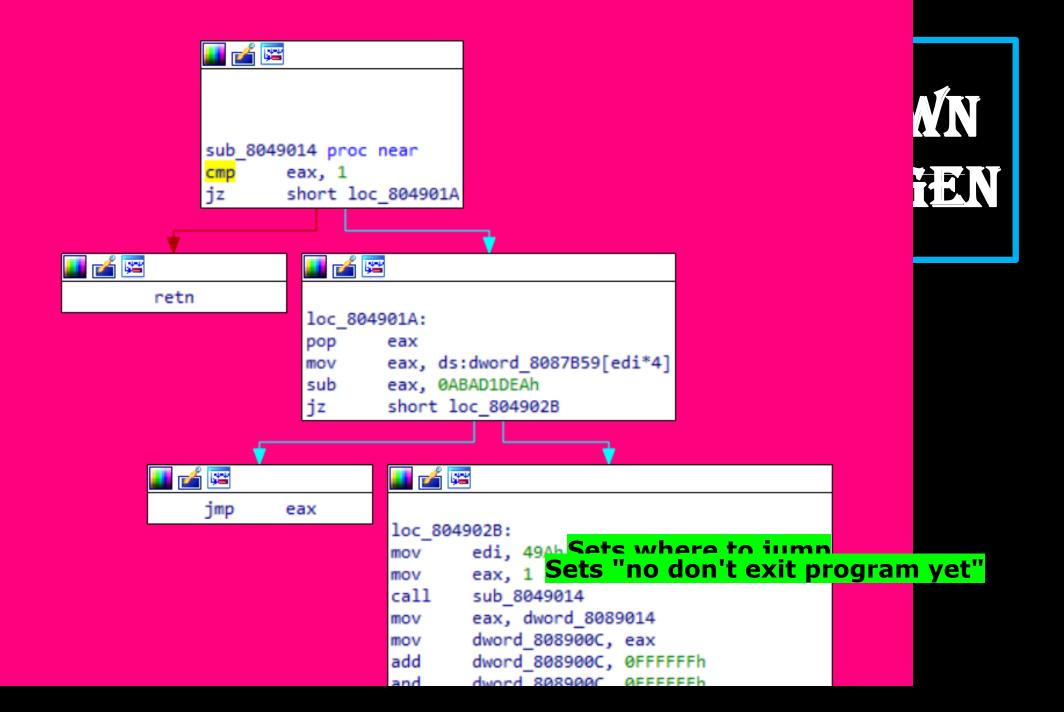
#### Works bcuz disassemblers don't split a block on a call

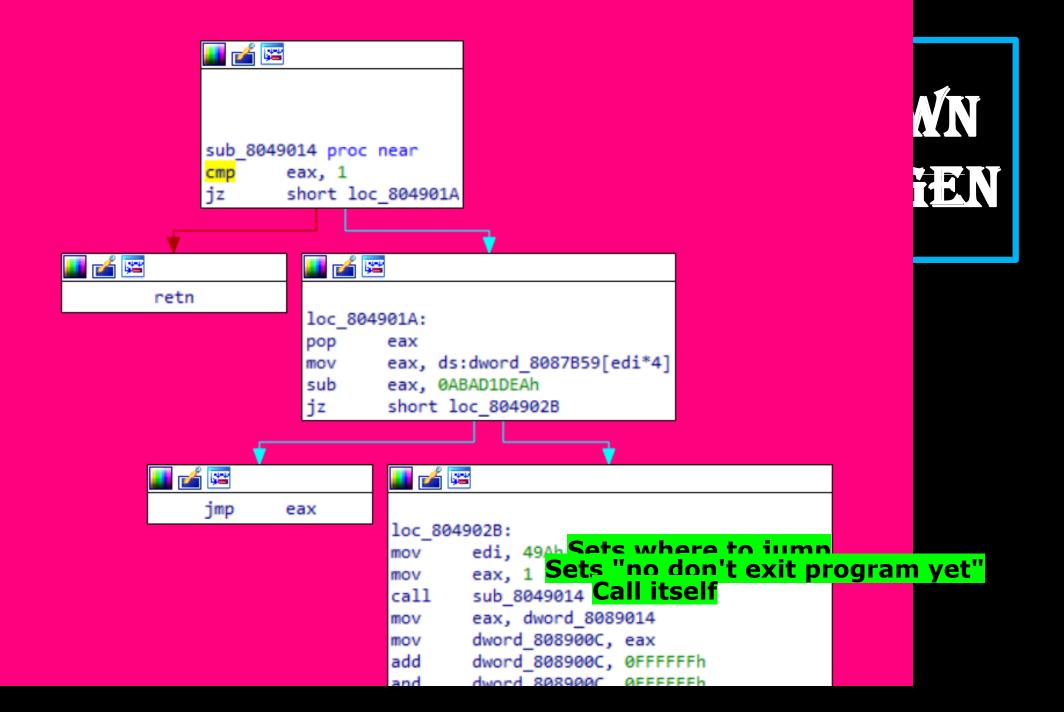


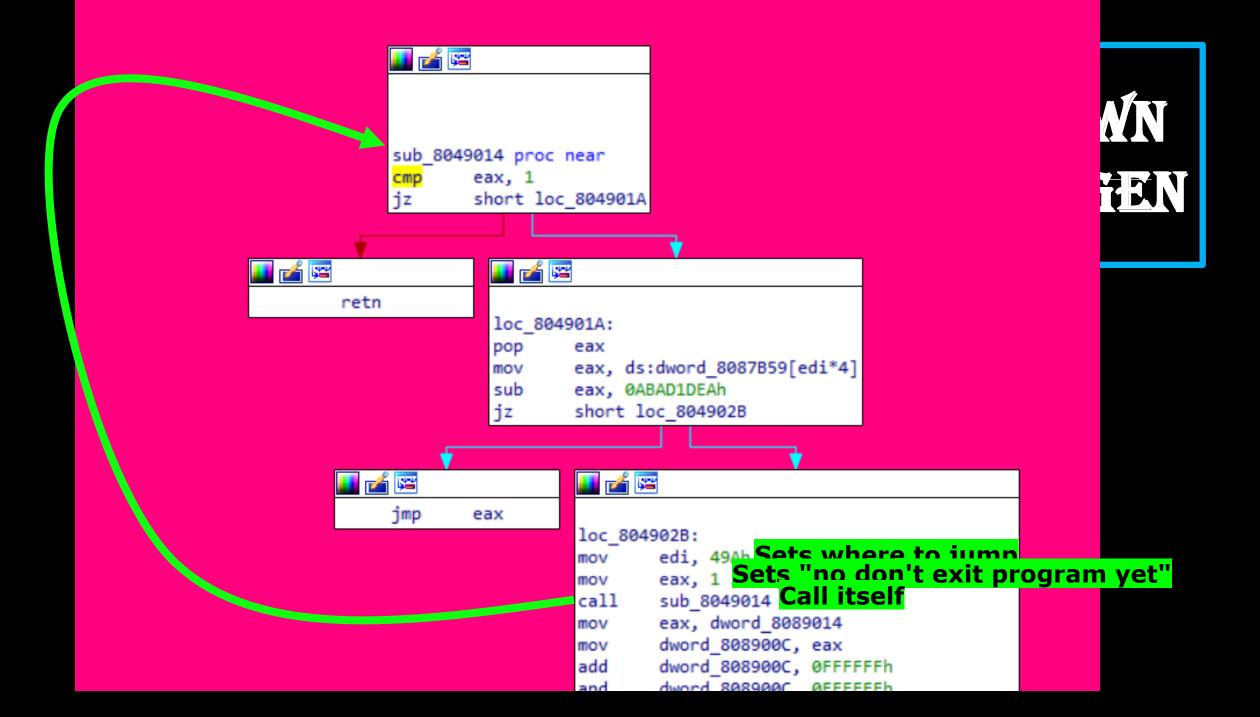
ŃN

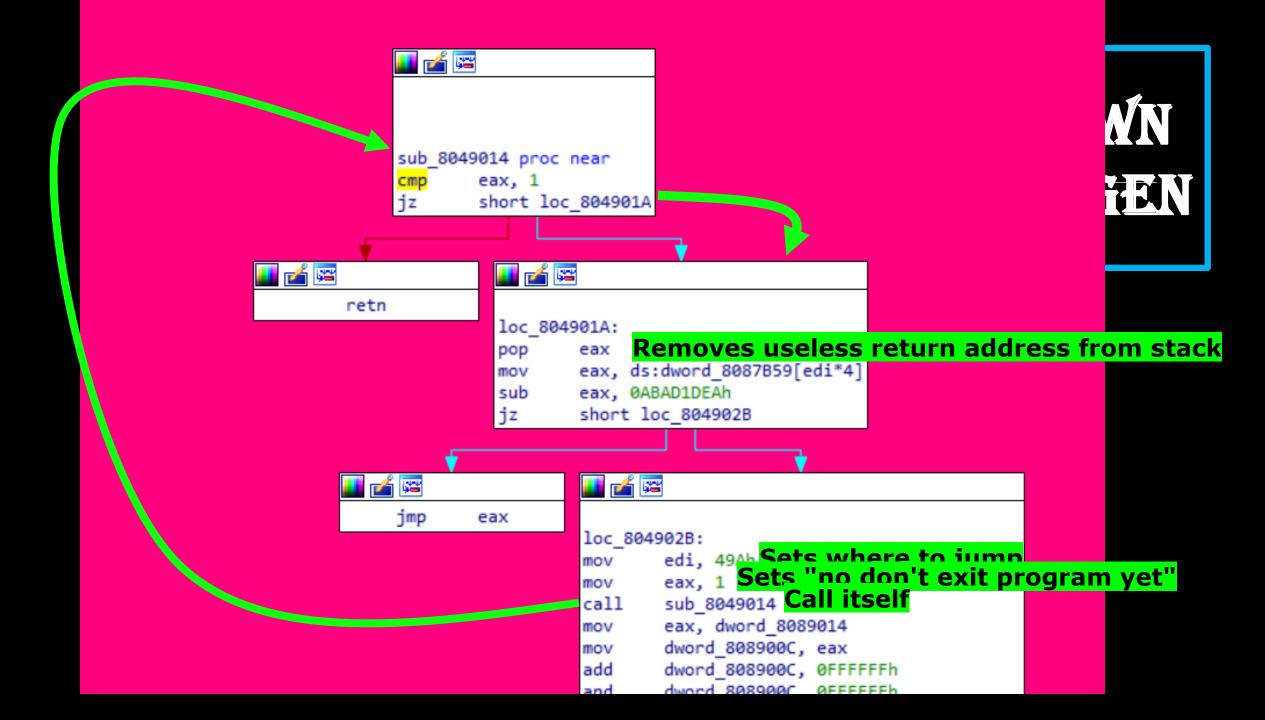


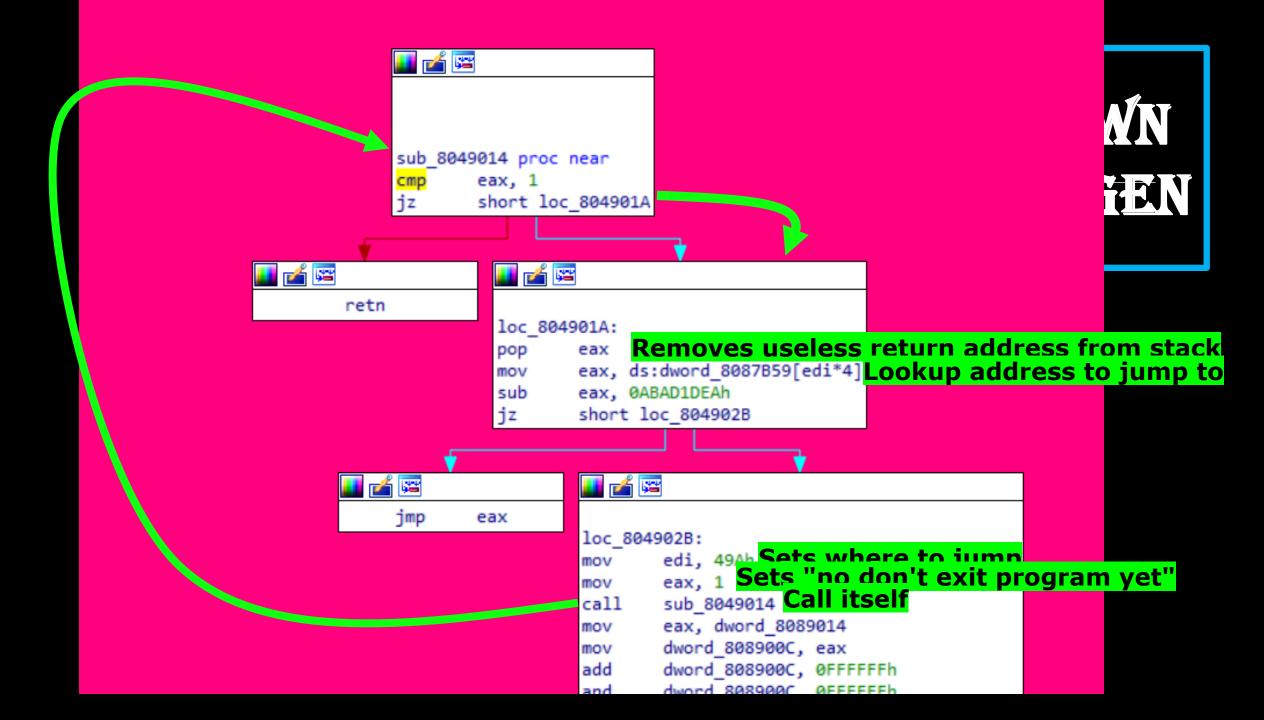
ŃN

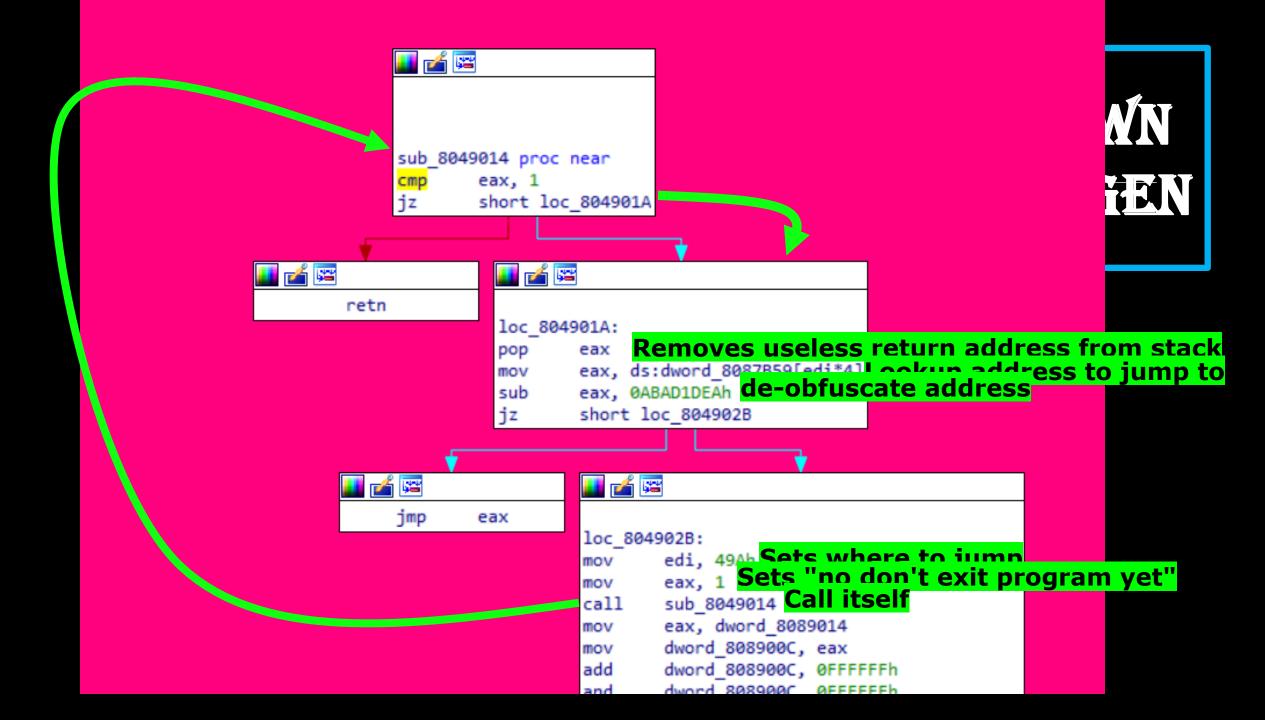


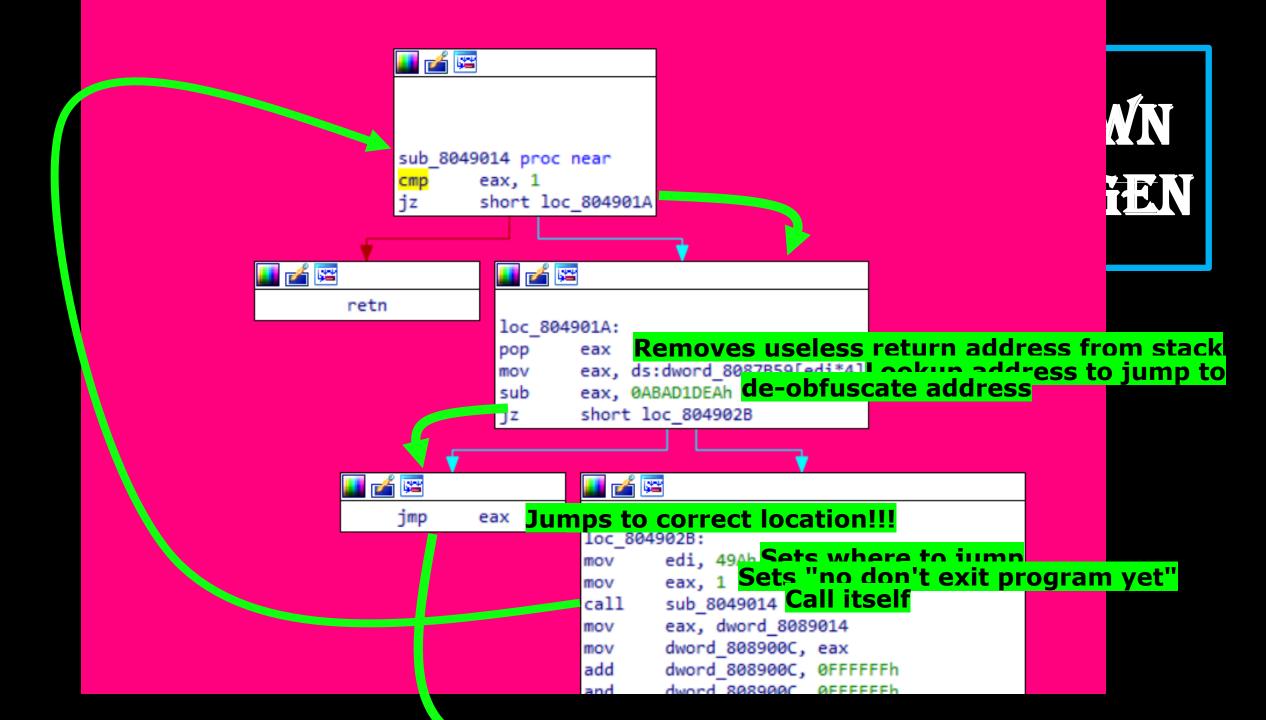


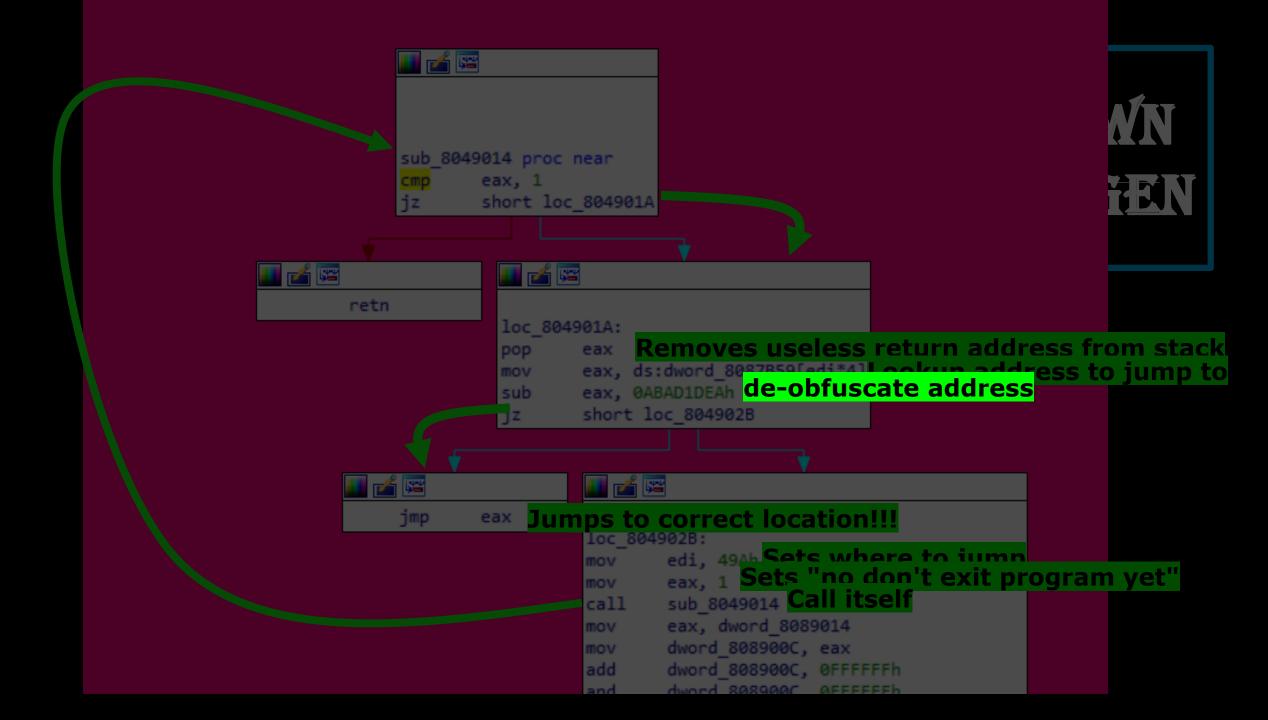


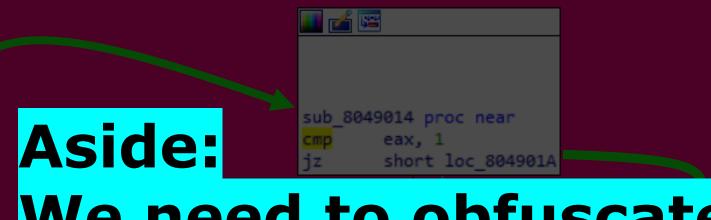












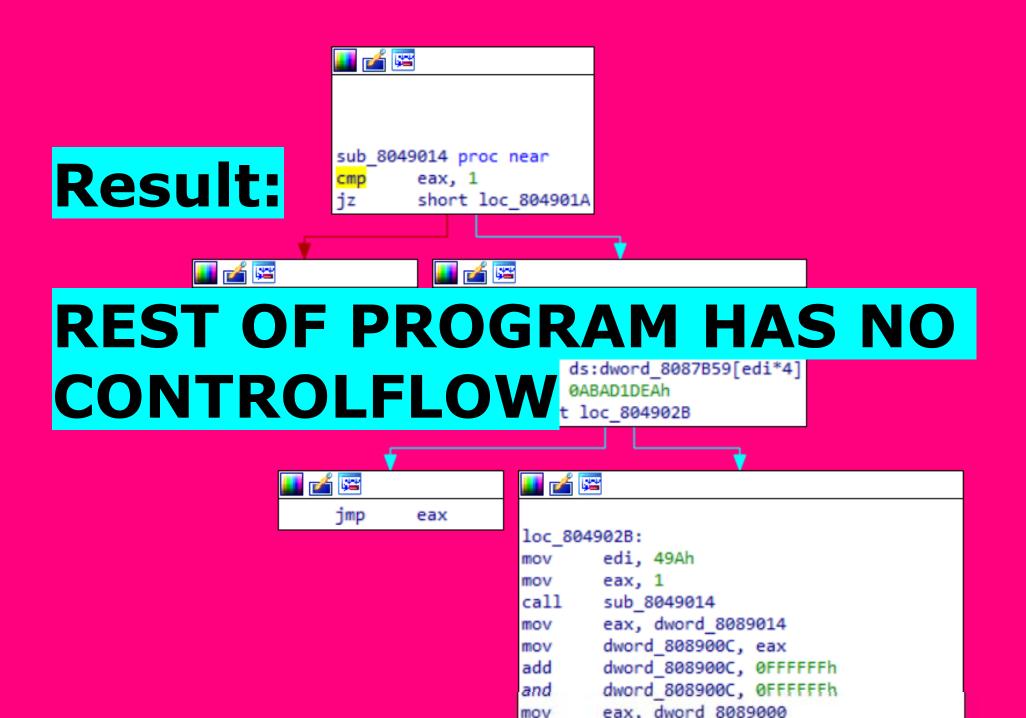
retn



#### We need to obfuscate the address

loc_80		
рор	eax Removes useless return address from state	
mov	eax, ds:dword 8087859[edi*4] ookup address to jump	to
sub	eax, ds:dword_80278505edit41 ookup address to jump eax, 0ABAD1DEAh de-obfuscate address	
ΤZ	short loc 804902B	

#### because otherwise IDA can find the jump target and draw an arrow to there, WHICH WE DONT WANT



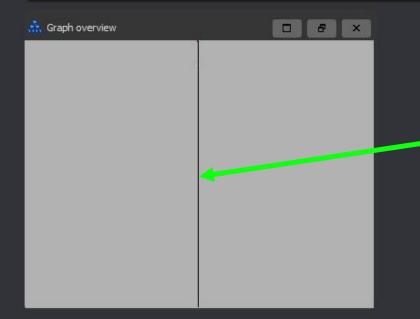


Jules 10/13/2022 12:37 PM You've heard of controlflow flattening, get ready for controlflow lengthening



Lisp

/uwu\$ ./a.art.exe
> (defun uwu (x y) (+ x y))
(lambda (x y) (+ x y))
> (defun owo (x y z w) (uwu (uwu x y) (uwu z w)))
(lambda (x y z w) (uwu (uwu x y) (uwu z w)))
> (owo 1 2 3 4)
10
>



# This is the graph!

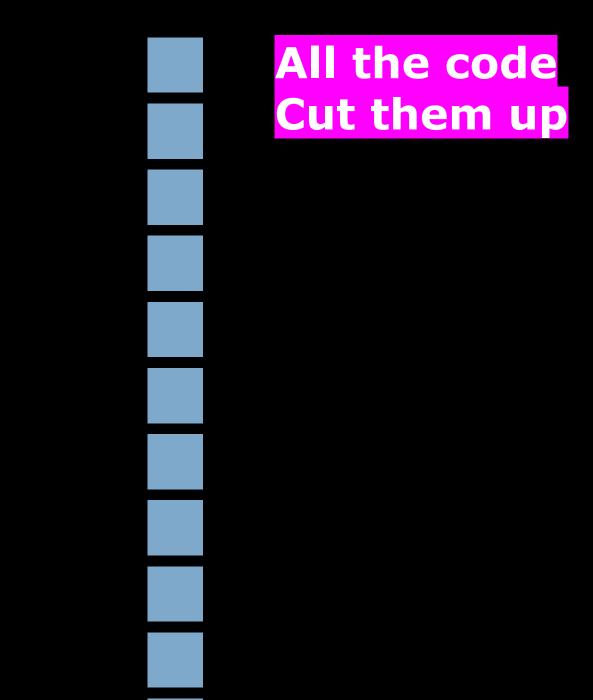


епідтатті 10/13/2022 12:40 рм НАННАНАНАНАНАНАНАНАНАНАНАНАНА

u did it

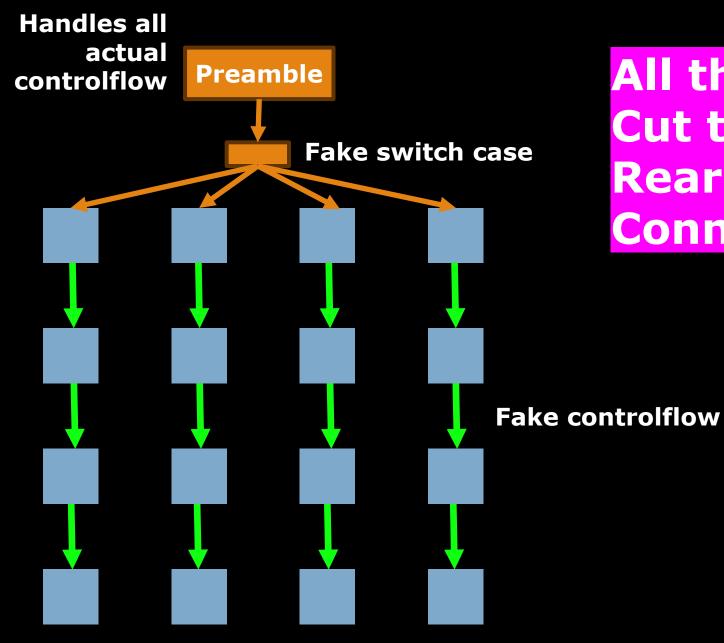
madlad

#### All the code



All the code Cut them up Rearrange into a picture

Brightness of pixel given by length of each "piece"



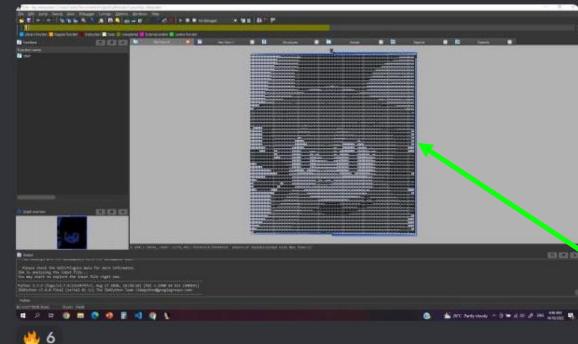
#### All the code Cut them up Rearrange into a picture Connect by adding jmps



Jules 10/14/2022 4:43 AM

#### lisp-interpreter.zip 113.01 кв

#### Ok am going to sleep





nigmatrix 10/14/2022 9:09 AM



sansders 10/14/2022 10:24 AM oh wow you can potentially do bad apple on ida oh nvm you already did LOL



ariana 10/14/2022 10:50 AM and this is how re challenges are made at 5am

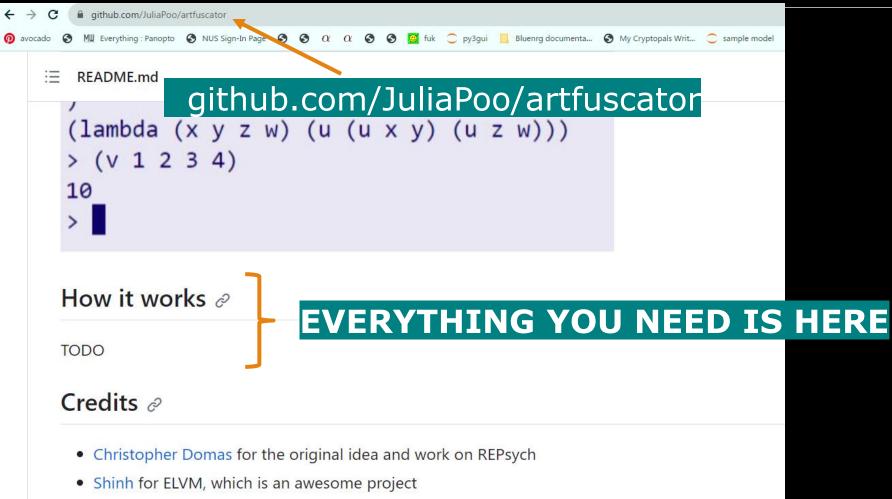
## This is the graph!

#### \*\*\*\*\*\*\* Project Informatior ╡<u>╝╗</u>╗╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪╪<mark>┊╝╝╝╝</mark>╤╤╪╪╪╪╪

#### Documentation for Artfuscator



#### Documentation for Artfuscator



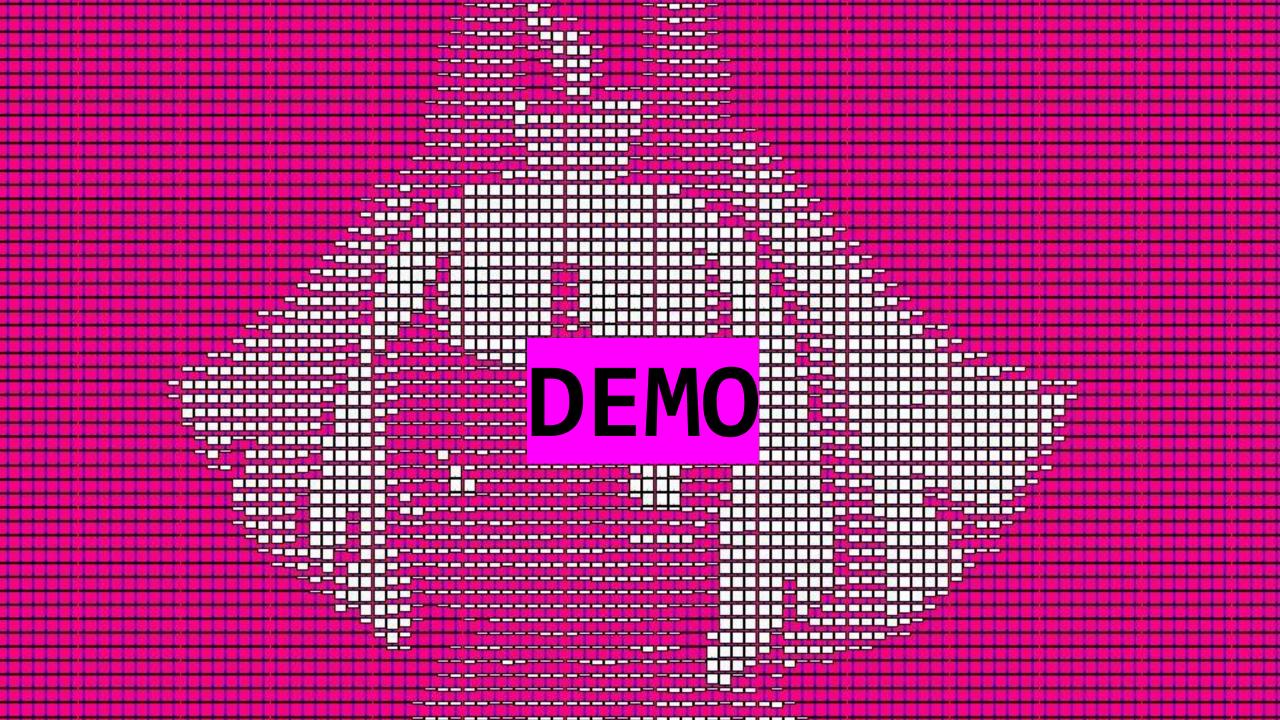
### Licensing

The MIT License does not hold me LIABLE FOR ANY CLAIM, DAMAGES, PSYCHOLOGICAL DAMANGES OR OTHER LIABILITY, ARISING FROM THE USE OF THIS PROJECT

Limitations

× Liability

× Warranty



## re: important information about me

# e0958745(0).nus.edu