# **SYNACKTIV**

### Demystifying Objective-C internals A practical approach

Sthack 23/05/2025

PUBLIC







#### Victor Cutillas

- Security researcher @ Synacktiv
  - Reverse engineering
  - Exploitation
- Mostly working on iOS





#### Context

- Objective-C 101
- Compilation
- Metadata
- Reversing tips
- Objective-C helper plugin

#### Conclusion



# <u>Context</u>

### <u>Context</u>



#### What is the aim of this talk?

- It is **not** to be technically exhaustive
  - You should be able to dive deeper yourself afterwards
- It is **not** to discuss which tool is the best
  - We are focusing on understanding **static analysis** methodology
- It is to give iOS newcomers keys to analyze Objective-C code
  - What is the role of the runtime?
  - How to efficiently reverse engineer Objective-C implementations?
  - How to recover types of manipulated variables?
- An IDA plugin is also presented to ease analysis

### <u>Con</u>text



#### Objective-C appeared in 1984

- Derived from Smalltalk
- Used in NeXTSTEP, the historical code base of Darwin (macOS)

### Apple language of reference until 2014

- New code bases now use Swift
- As of iOS 18.4, among 3800+ libraries in the shared cache
  - 90% rely on Objective-C code
- Almost all core services are implemented using Objective-C

### <u>Con</u>text



#### It is a superset of the C language

- Compatible with the C syntax
- It is also possible to write Objective-C++

#### It is object-oriented

Notion of classes, inheritance

#### ... and reflective

Reflection is the ability of a process to examine, introspect, and modify its own structure and behavior.

Reflective programming – Wikipedia

This implies metadata!



# **Objective-C 101**



### Basic terminology

**Objective-C 101** Reverser edition!

- Class: set of methods, properties & instance variables
- Instance variable: a field of a class
- Property: instance variable with compiler-generated getter/setter

#### Objective-C is message-based

- Protocol: defines a required set of methods & properties
  - Similar to a C++ interface
- Selector: pointer to a C-style string discriminating a method
- Message passing: alternative name for method calling





#### NSObject root class

- All Objective-C classes inherit from this class
  - The only other root class is NSProxy (out of this talk scope)
- Implemented by Foundation
- Provides a default set of methods



#### NSObject root class

- Lifetime
  - + alloc
  - retain
  - release

- Introspection
  - + isSubclassOfClass:
  - isKindOfClass:
  - respondsToSelector:
- init valueForKey:
- Some of them are designed to be overridden, some examples:
  - description
  - valueForUnknownKey:

- Describing
  - description
  - debugDescription









#### Objects lifetime

- Automatic Reference Counting (ARC) is used everywhere
- No explicit calls to retain/release are made in the code
  - The compiler triggers an error if the developer uses them
- Lifetime-related calls are automatically generated by the compiler



13

### Objective-C runtime

**Objective-C 101** 

Source code is available on GitHub\*

Reverser edition!

- Objective-C code links to the runtime library
  - libobjc.A.dylib
- The runtime exposes functions called by the compiler

...

objc\_msgSend()

objc\_opt\_isKindOfClass()

objc\_release()

objc\_autoreleasePoolPush()

- objc\_retain()
- It also exposes a user API: <objc/runtime.h>
  - More on that later



#### Foundation framework

**<u>Obj</u>ective-C 101** Reverser edition!

- Implements base objects, below are a few examples
- Language features
  - Arrays, dictionaries, sets
  - Strings, numbers
- Operating system interactions
  - Filesystem
  - Networking
- Inter-process communication (NSXPC)

- Archiving, unarchiving
- Errors
- Process management



# **<u>Compilation</u>**

### Toolchain

- Apple uses the LLVM toolchain
  - It is opensource\*
  - Apple maintains its own version with proprietary patches
- Clang is therefore the officially supported compiler
  - It is used by Apple's IDE, Xcode
  - It can also be directly used in the command line

#### SYNACKTIV

# **Compilation**

### Code generation

- Consider Objective-C as being internally transpiled to C
- Here is a method from UIKitCore class UIView

```
@implementation UIView
- (void)_registerAuxiliaryChildEnvironmentForTraitInvalidations:(id)obj
{
    if (self->_traitChangeRegistry == nil)
        self->_traitChangeRegistry = [[_UITraitChangeRegistry alloc] init];
    [self->_traitChangeRegistry
        registerAuxiliaryChildEnvironmentForTraitInvalidations:obj];
}
@end
```

Let's understand its generated implementation step by step

#### SYNACKTIV

### **Compilation**

### Code generation

Method calls are converted to objc\_msgSend() calls

```
- (void) registerAuxiliaryChildEnvironmentForTraitInvalidations: (id)obj
{
    if (self-> traitChangeRegistry == nil)
        self->_traitChangeRegistry = [[ UITraitChangeRegistry alloc] init];
    [self->_traitChangeRegistry
        registerAuxiliaryChildEnvironmentForTraitInvalidations:objl;
               - (void)_registerAuxiliaryChildEnvironmentForTraitInvalidations: (id)obj
                    if (self->_traitChangeRegistry == nil)
                       self->_traitChangeRegistry = [[_UITraitChangeRegistry alloc] init];
                    objc_msgSend(self->_traitChangeRegistry,
                                 "registerAuxiliaryChildEnvironmentForTraitInvalidations:",
                                 obj);
```

#### Code generation

This previous method call

- ... is made on an object
- ... invokes the method identified by its selector
- ... passes one method argument

- Note: the code is a bit simplified for readability
  - objc\_msgSend() must actually be cast to be called
  - The selector must be wrapped

```
- (void)_registerAuxiliaryChildEnvironmentForTraitInvalidations:(id)obj
{
    if (self->_traitChangeRegistry == nil)
        self->_traitChangeRegistry = [[_UITraitChangeRegistry alloc] init];
    (void (*)(id, SEL, id) objc_msgSend)(
        self->_traitChangeRegistry,
        @selector(registerAuxiliaryChildEnvironmentForTraitInvalidations:),
        obj);
}
```

#### Code generation

- The example UIView method
  - (void)\_registerAuxiliaryChildEnvironmentForTraitInvalidations: (id)obj;
  - ... is called on an instance of a class
  - ... is identified by its selector
  - ... expects one argument
  - ... does not have a return value
- Here is the corresponding C implementation declaration

void \_registerAuxiliaryChildEnvironmentForTraitInvalidations(UIView \*self, SEL sel, id obj);

This is identical to the corresponding objc\_msgSend() calls

- [[cls alloc] init] is a common object creation pattern
  - This call chain is optimized by the compiler

#### SYNACKTIV

## **Compilation**

- ARC automatically calls lifetime-related functions
  - objc\_retain(), objc\_release()
  - objc\_autorelease(), objc\_claimAutoreleasedReturnValue()

#### SYNACKTIV

## **Compilation**

- It is possible to convert Objective-C to C++ using a compiler flag
  - This may help you better understand code generation
  - Is is also pretty useful to better understand metadata structures

```
clang -arch arm64 -isysroot "$SDK_ROOT" -I"$SDK_ROOT"/usr/include -fobjc-arc -rewrite-objc code.m
```

#### Method calls optimization

- Since iOS 17, no direct calls to objc\_msgSend() are made
  - Stubs are generated and shared as much as possible
  - They are shared between shared cache libraries



### Objective-C blocks

- Anonymous functions, similar to C++ lambdas
- Very common in Apple code
  - Mostly because of interactions with libdispatch.dylib
  - This library provides a C API for executing code concurrently
- Most common use cases are
  - Lazy initialization → dispatch\_once()
  - Asynchronous function call → dispatch\_async()
  - Synchronous function call
  - Various callbacks

- $\rightarrow$  dispatch\_sync()
  - $\rightarrow$  filtering, enumeration, ...



#### Objective-C blocks

Example method from NSConcreteFileHandle (Foundation)

```
- (void)closeFile
{
    if (self->_flags & FILE_CLOSED)
        return;
    [self _cancelDispatchSources];
    dispatch_async(self->_monitoringQueue, ^(){
        self->_flags |= FILE_CLOSED;
        close(self->_fd);
    });
}
```

The block code will be executed on the given queue (a thread)

#### **SYNACKTIV**

#### **Objective-C blocks** - (void)closeFile

```
if (self-> flags & FILE CLOSED)
    return:
```

```
[self cancelDispatchSources];
dispatch_async(self->_monitoringQueue, ^(){
    self-> flags |= FILE CLOSED;
    close(self-> fd);
```

```
void closeFile(NSConcreteFileHandle *self, SEL sel)
    if (self-> flags & FILE CLOSED)
        return;
    objc_msgSend(self, "_cancelDispatchSources");
    Block_layout_closeFile block = {
        .isa = &OBJC CLASS NSStackBlock ,
        .flags = 0 \times COOOOOOO
        .invoke = closeFile_block_invoke,
        .descriptor = &closeFile_block_descriptor,
        .self = self,
    };
    dispatch async(self-> monitoringQueue, &block);
```

```
void closeFile_block_invoke(Block_layout_closeFile *block)
```

```
block->self->_flags |= FILE_CLOSED;
close(block->self->_fd);
```

});

#### SYNACKTIV

### **Compilation**

### Objective-C blocks

Blocks are objects generated by the compiler

```
struct Block_layout_closeFile {
    Class isa;
    int32_t flags;
    int32_t reserved;
    void (*invoke)(Block_layout_closeFile *block);
    Block_descriptor *descriptor;
    // Captured variables
    NSConcreteFileHandle *self;
};
```

- invoke points to a native function containing the block code
  - Blocks can also take parameters, additionally passed to this function
- descriptor provides metadata about the block (more on that later)
- The layout of the structure also includes captured variables

#### Objective-C blocks

```
(void)closeFile
—
{
    if (self->_flags & FILE_CLOSED)
        return;
    [self _cancelDispatchSources];
    dispatch_async(self->_monitoringQueue, ^(NSError *error){
        if (error == nil) {
            self->_flags |= FILE_CLOSED;
            close(self-> fd);
    });
                   void closeFile_block_invoke(Block_layout_closeFile *block, NSError *error)
                   ſ
                       if (error == nil) {
                           block->self->_flags |= FILE_CLOSED;
                           close(block->self->_fd);
                       }
                   }
```

#### Objective-C blocks

- You will encounter two main types of blocks
- Global blocks
  - Are stored in a read-only section of the Mach-O
  - Typically used for lazy initialization of global data: dispatch\_once()
- Stack blocks
  - Created at runtime
  - Used when local variables are captured





#### Objective-C data is stored in dedicated sections

- This allows dyld to properly register Objective-C information to the runtime
- Their name is prefixed with \_\_objc\_, here are some examples

Section name	Contents
objc_classlist objc_protolist	Pointers to defined classes & protocols
objc_selrefs objc_classrefs	Pointers to selectors & classes used by this Mach-O
objc_ivar	Instance variable offsets
objc_data	Structures of defined classes & metaclasses
objc_const	Constant data: typically classes & metaclasses information

#### **SYNACKTIV**

## **Compilation**

### Listing method calls

1) Go to the corresponding **method\_t** structure

- By listing references to the implementation
- Or directly searching the selector string



#### **SYNACKTIV**

### Listing method calls

2) Find all stubs referencing this selector



### Listing method calls

3) List & filter stub calls if other methods share this selector



#### Calling a method

- Let's figure out how objc\_msgSend() decides what to call
- This will walk us through the Objective-C metadata structures
- Keep in mind that all data structures are public

### Calling a method

- 1) ISA (class) pointer is used to fetch metadata about the object
- 2) Class information is then fetched



. . .

### Calling a method

- The runtime now has access to metadata lists concerning
  - ivars Instance variables (structure fields)
  - base\_props
     Class properties (*ivars* with getter and/or setter)
  - base\_meths
     Defined methods
  - base\_prots
     Protocols the class conforms to

#### SYNACKTIV

### <u>Met</u>adata

### Calling a method

- 3) Runtime iterates over the list of all declared methods
- 4) Requested selector is compared with each method selector



#### SYNACKTIV

#### Calling a method

5) When a match is found, the implementation is called



SYNACKTIV

### Calling a method

#### 6) Method not found?

• Start over by looking it up in the *super* class



- No super class? Runtime calls [obj doesNotRecognizeSelector:]
  - NSObject implementation throws an exception, leading to a crash

#### SYNACKTIV

#### Method signature

Let's take a look at this types field



This string encodes information about arguments

#### Method signature

- Type encoding is partially documented
- Describes return value and arguments: type, location



@0:8 is by definition common to all Objective-C methods

#### Method signature

- Side note: variadic Objective-C methods exist
  - Yet, no metadata is available to check that they are variadic
  - Method signatures simply stop before the variadic arguments

```
@interface NSSet
```

```
// Signature is "@24@0:8@16"
+ (NSSet *)setWithObject:(id)obj;
// Signature is identical
+ (NSSet *)setWithObjects:(id)obj, ...;
```

#### @end

### Type encoding

- Describes primitive types
  - Numbers
  - Objects (instance, class)
- ... and recursive types
  - Pointers
  - Arrays
  - Structures, unions
- A public documentation exists





#### Type encoding

Encoded structure type

^{UIContentViewElementLayoutInfo={CGSize=dd}B{NSDirectionalEdgeInsets=dddd}}

- Argument type: UIContentViewElementLayoutInfo \*
- Associated data structures

```
struct UIContentViewElementLayoutInfo {
   CGSize var1;
   BOOL var2;
   NSDirectionalEdgeInsets var3;
};
```



#### Extended type encoding

- Some metadata structures contain *extended* type encoding
  - Instance variables information
  - Protocols
  - Block descriptors
- They contain additional information about objects & structures
  - Protocol the object shall conform to
  - Expected object class
  - Structure field names





#### Extended type encoding

- Block signature
  - Here a protocol is specified for the return value

@"<NSCopying>"40@?0r^v8{\_NSRange=QQ}16^B32

id<NSCopying>(^)(Block\_layout\_f00 \*, const void \*, \_NSRange, bool \*);

- Protocol method
  - Object class names in protocols are required for NSXPC serialization





# **<u>Reversing tips</u>**

## <u>Reversing tips</u> – Recovering types

#### Some instance variables are accessed through their metadata

- Their offset in the object is fetched at runtime
  - This is not true for all instance variables
- It is then possible to deduce the associated object type

ADRP	X8, #_OBJC_IVAR_\$_UIViewtraitChangeRegistry@PAGE
LDRSW	X8, [X8,#_OBJC_IVAR_\$_UIViewtraitChangeRegistry@PAGEOFF]
LDR	X0, [X23,X8] ; X23 is v5
BL	UITraitChangeRegistry_updateAuxiliaryChildrenTraitsIfNeedec

-[\_UITraitChangeRegistry updateAuxiliaryChildrenTraitsIfNeeded](\*((\_QWORD \*)v5 + 6));

- Here, the decompiler inlines the offset value and adds it to v5
  - v5 can therefore be safely assumed to be UIView \* (or a subclass)

-[\_UITraitChangeRegistry updateAuxiliaryChildrenTraitsIfNeeded](v5->\_traitChangeRegistry);

### <u>Reversing tips</u> – Recovering types

#### Some instance variables are accessed through their metadata

Accesses to this instance variable can be easily listed

• Keep in mind that this is only valid for a subset of instance variables

	ADRP	X8, #_OBJC	_IVAR_\$_UIViewtraitChangeRe	gistry@PAG	E ; _UITraitChangeRegistry *_traitChange	Registr
	LDRSW LDR	X8, [X8,# <mark>_</mark> X8, [X0,X8	OBJC_IVAR_\$_UIViewtraitCham ]	ngeRegistry	<pre>'@PAGEOFF] ; _UITraitChangeRegistry *_tra</pre>	itChang
refs to _OBJC_	IVAR_\$_UIViewtraitChangeRegistry					
Direction T	īype Address			Text		
🗷 Up 🕡	o -[UIView_wrappe	dProcessTraitChan	ges:withBehavior:]:loc_189B4480	C <mark>ADRP</mark>	X8, #_OBJC_IVAR_\$_UIViewtraitChangeR	
🗷 Up 🥡	o -[UIView traitCol	ectionDidChangel	nternal:]+2B0	ADRP	X8, # OBJC IVAR \$ UIView. traitChangeR	
🗷 Up 🛛	o -[UIView register	ForTraitTokenChan	ges:withTarget:action:]+30	ADRP	X8, # OBIC IVAR \$ UIView. traitChangeR	
⊠Up o	o -[UIView register	ForTraitTokenChan	ges:withHandler:]+28	ADRP	X8, # OBIC IVAR \$ UIView. traitChangeR	+C↑J
zuung 🗹	o -[UIView unregiste	erForTraitChanges:	]+4	ADRP	X8, # OBIC IVAR \$ UIView. traitChangeR	
⊠Up o	o -[UIView register	AuxiliaryChildEnvi	ronmentForTraitInvalidations:]+10		X8, # OBJC IVAR \$ UIView. traitChangeR	
<b>z</b> i (	o -[UIView unregis	erAuxiliaryChildEr	nvironmentForTraitInvalidations:]	ADRP	X8, # OBJC IVAR \$ UIView. traitChangeR	
🗙 🎽 adrp						
Line 7 of 7						

This type recovery method can be automated in your analysis tool

#### Private frameworks can be imported

- Objective-C frameworks can be dynamically loaded via dlopen()
  - Use the runtime API to manipulate private classes
- Example: check if the device sends crash logs to Apple

#### **SYNACKTIV**

#### Instrumenting methods

The runtime provides a simple way to hook method calls

```
// Hooking -[NSArray description]
NSString *(*description_impl)(id, SEL) = NULL;
                                                               $./hook
NSString *description_hook(id self, SEL sel) {
                                                              0x109b380d0 description: (
    NSString *desc = description_impl(self, sel);
                                                                   42,
    printf("%p description: %s", (__bridge void *)self,
                                                                   hello
           desc.UTF8String);
    return desc:
}
int main(void) {
    Method method = class_getInstanceMethod(objc_getClass("NSArray"),
                                            sel registerName("description"));
    description_impl = method_setImplementation(method, description_hook);
    NSLog(@"Array: %@", @[ @42, @"hello" ]); // Calls -[NSArray description]
```



#### Reference counting optimization

• Since iOS 16, calls to *refcount*-related functions are specialized

id objc\_retain(id obj);
void objc\_release(id obj);

- obj parameter can now be passed in any ARM64 register
  - From X0 to X25
  - Except sensitive registers X16, X17, X18
- This reduces register manipulation in generated assembly
  - Return value is still stored in X0

#### Reference counting optimization

- Specialized flavours are exposed by the runtime
  - See header file runtime/objc-abi.h
- A decent reversing tool should support those calls
  - They are omnipresent in generated assembly code
  - Decompiler output is messy without support

_registerA	<pre>wuxiliaryChildEnvironmentForTraitInvalidations:</pre>				
SUB	SP, SP, #0×30				
STP	X20, X19, [SP,#0x20+var_10]		_objc_retain_x2:		
STP	X29, X30, [SP,#0x20+var_s0]				
ADD	X29, SP, #0x20		ANDS	X0, X2, X2	
MOV	X19, X0		B.LE	objc_retain_ret	
BL	_objc_retain_x2 ; Argument is retained		LDXR	X16, [X2]	
MOV	X1, X0				

## **<u>Reversing</u>** tips



#### Runtime uses Pointer Authentication Code (PAC)

- Cryptographic ARM64 hardware protection
  - Available since iPhone XS & XR (A12+ SoCs)
- Enables developers to protect pointers with signatures
- An attacker controlling a protected pointer must then provide
  - A valid address
  - A valid signature of this address
  - Otherwise a fatal error is triggered (kernel sends a SIGKILL)
- This *dramatically* increases exploitation difficulty

## **<u>Reversing</u>** tips



#### Runtime uses Pointer Authentication Code (PAC)

- Data pointers
  - Object class pointers (ISA) are protected
    - Prevents memory corruption & Use-After-Free bugs
  - Caches, custom virtual tables, dynamically registered metadata, etc.
- Function pointers
  - Exception handling, image loading, internal hooks, trampolines, etc.
- Distinct discriminators are used for all pointer types
  - Integers scrambled with the pointer they are signed with
  - Provides a semantic enforcement on signed pointers authentication





#### Decompiler output cleanup

- Decompiler output contains many (usually) irrelevant runtime calls
  - Lifetime objc\_retain(), objc\_release()
  - Return values objc\_autorelease(), objc\_claimAutoreleasedReturnValue()
  - Property manipulation objc\_storeStrong(), objc\_loadWeakRetained()
- They slow down the analyst when reading code
- ... and also prevent type propagation



#### Decompiler output cleanup

- Return values are made opaque when claimed
  - Ambiguous variable typing
  - Method implementations are unresolved on untyped variables

UIViewControllerViewAnimator \*from; id view; id window; id controller;

```
view = objc_claimAutoreleasedReturnValue(
    -[UIViewControllerViewAnimator view](
    from, "view"));
window = objc_claimAutoreleasedReturnValue(
    objc_msgSend(v77, "window"));
controller = objc_claimAutoreleasedReturnValue(
    +[UIWindowController windowControllerForWindow:](
      &OBJC_CLASS___UIWindowController,
      "windowControllerForWindow:",
      window));
objc_release(v78);
objc_release(v77);
```



#### Decompiler output cleanup

- Removing those calls enhances the analysis efficiency
  - Greatly improved type propagation
    - Method calls are chained
  - Fewer intermediate variables
  - Around 10% to 20% fewer lines of decompiled code

UIViewController \*from; UIWindowController \*controller;

```
controller = +[UIWindowController windowControllerForWindow:](
   &OBJC_CLASS___UIWindowController,
    "windowControllerForWindow:",
    -[UIView window](-[UIViewController view](from, "view"), "window"));
```



#### Decompiler output cleanup

- Implemented using a microcode hook (MMAT\_PREOPTIMIZED)
  - Intermediate language used by Hex-Rays decompiler
- Runtime calls are replaced by micro instructions "*neutralizing*" them
  - They replicate the caller's state after the call is made
- Example of a replacement
  - ARM64 assembly
     BL \_objc\_retain\_x3
     Generated microcode
     (all !objc\_retain <fast:"id obj" x3.8> => id x0.8
     Neutralization
     x3.8, x0.8



#### Other feature

- Propagation of cross-references to current method selector
- Release will be made very soon on Synacktiv's Github page\*



# **Conclusion**

### **Conclusion**



#### Objective-C provides a lot of metadata

- It is very reversing-friendly
- Structures can be easily recovered
- Manipulated types are pretty straightforward to determine

#### It will not disappear anytime soon

- 90% of shared cache libraries rely on its use
- It is deeply integrated inside the core of iOS

# **SYNACKTIV**

https://www.linkedin.com/company/synacktiv

https://x.com/synacktiv

X

https://bsky.app/profile/synacktiv.com



https://synacktiv.com