

PA193 - Secure coding principles and practices



Designing good and secure API
Automata-based programming

Petr Švenda svenda@fi.muni.cz

CRCS

Centre for Research on
Cryptography and Security

Overview

- Lecture:
 - How to write good & secure API
 - automata-based programming
- Labs
 - Write small security API
 - Model inner state with automata-based design

ORGANIZATIONAL

Reuse of existing code

- Code reuse is generally great thing, but..
- NOT in homework or assignments!
- It is **NOTOK**:
 - Take any code from web when you should create code completely on your own (project - parser)
 - Share code of your solution with others (homework)
- Penalization points assigned
- Next occurrences will be handled by faculty's disciplinary committee

```
#include "LDSSecurityObject.h"
#include <dirent.h>
#include <openssl/sha.h>
int main(void)
{
    LDSSecurityObject_t *lds;
    lds = (LDSSecurityObject_t*)calloc(1, sizeof *lds);
    DIR *dir;
    FILE *fp;
    char dirname[100],dirname1[100];
    char filenames[100][100];
    char correctnames[17][100];
    int countfiles =0,i;
    int count,j,cmp,flag=0;
    int foundindex[16];
    struct dirent *ent;
    if(!lds) exit(1);

    FILE *f=fopen("Sample-data/lds.bin","rb");
    if(!f) exit(1);
    unsigned char buffer[10000];
    int buflen,size;
    char *input;
    unsigned char *hashvalue;
    buflen=fread(buffer,1,10000,f);
    fclose(f);

    printf("Input the name of directory (example Sample-data)");
    scanf("%s",dirname);

    strcpy(dirname1,dirname);
    if ((dir = opendir (dirname)) != NULL)
    {
        while ((ent = readdir (dir)) != NULL)
        {
            strcpy(filenames[countfiles],ent->d_name);
            //printf ("%s\n", ent->d_name);
            //printf ("%s\n", filenames[countfiles]);
            countfiles++;
        }
        closedir (dir);
    }
    else
    {
        /* could not open directory */
        perror ("");
    }
}
```

```
#include "LDSSecurityObject.h"
#include <dirent.h>
#include <openssl/sha.h>
int main(void)
{
    LDSSecurityObject_t *lds;
    lds = (LDSSecurityObject_t*)calloc(1, sizeof *lds);
    DIR *dir;
    FILE *fp;
    char Directory[100],Directory1[100];
    char in_file_name[100][100];
    char corrcct_names[17][100];
    int no_of_files =0,i;
    int cnt,j,cmp,flag=0;

    struct dirent *ent;
    if(!lds) exit(1);

    FILE *f=fopen("Sample-data/lds.bin","rb");
    if(!f) exit(1);
    unsigned char buffer[10000];
    int buflen,size;
    char *input;
    unsigned char *hashvalue;
    buflen=fread(buffer,1,10000,f);
    fclose(f);

    printf("Enter The directory Name whos files to be veified :");
    scanf("%s",Directory);

    strcpy(Directory1,Directory);
    if ((dir = opendir (Directory)) != NULL)
    {
        while ((ent = readdir (dir)) != NULL)
        {
            strcpy(in_file_name[no_of_files],ent->d_name);
            no_of_files++;
        }
        closedir(dir);
    }
    else
    {
        /*Directory opening error*/
        perror ("");
    }
}
```

```
12/11/2015 11:27:15 4,135 bytes C, C++, C#, ObjC Source ANSI PC
int bitrates[] = {
    BITRATEFREE, BITRATEFREE, BITRATEFREE, BITRATEFREE, BITRATEFREE,
    32, 32, 32, 32, 8,
    64, 48, 40, 48, 16,
    96, 56, 48, 56, 24,
    128, 64, 56, 64, 32,
    160, 80, 64, 80, 40,
    192, 96, 80, 96, 48,
    224, 112, 96, 112, 56,
    256, 128, 112, 128, 64,
    288, 160, 128, 144, 80,
    320, 192, 160, 160, 96,
    352, 224, 192, 176, 112,
    384, 256, 224, 192, 128,
    416, 320, 256, 224, 144,
    448, 384, 320, 256, 160,
    BITRATEBAD, BITRATEBAD, BITRATEBAD, BITRATEBAD, BITRATEBAD
};

typedef struct{
```

```
18/11/2015 16:40:53 11,086 bytes C, C++, C#, ObjC Source ANSI UNIX
int bitrates[] = {
    BITRATEFREE, BITRATEFREE, BITRATEFREE, BITRATEFREE, BITRATEFREE,
    32, 32, 32, 32, 8,
    64, 48, 40, 48, 16,
    96, 56, 48, 56, 24,
    128, 64, 56, 64, 32,
    160, 80, 64, 80, 40,
    192, 96, 80, 96, 48,
    224, 112, 96, 112, 56,
    256, 128, 112, 128, 64,
    288, 160, 128, 144, 80,
    320, 192, 160, 160, 96,
    352, 224, 192, 176, 112,
    384, 256, 224, 192, 128,
    416, 320, 256, 224, 144,
    448, 384, 320, 256, 160,
    BITRATEBAD, BITRATEBAD, BITRATEBAD, BITRATEBAD, BITRATEBAD
};

typedef struct{
    /// unsigned framesync :12; //Frame synchronizer
```

```
12/11/2015 11:27:15 4,135 bytes C, C++, C#, ObjC Source ANSI PC
int readMP3header(FILE *f, MP3HEADER *h){
    MP3ID3TAG2 tag;

    //push file point to the beginning
    rewind(f);
    fread(&tag, 1, sizeof(MP3ID3TAG2), f);

    //tag id3v2 are located at the beginning of file, id3v1 at the end
    if(tag.tagid[0]=='I' && tag.tagid[1]=='D' && tag.tagid[2]=='3'){//is
        fseek(f, unpacktagsize(tag), SEEK_CUR);
    }else{//isn't tag id3v2 - go back
        rewind(f);
    }

    //I'm currently not interested in the final state of the file pointer
```

```
18/11/2015 16:40:53 11,086 bytes C, C++, C#, ObjC Source ANSI UNIX
int ReadMP3Header(FILE *f, MP3HEADER *h, unsigned int StartFlag, uint16_t framesync,
    unsigned char head[4] ;
    int cont ;
    MP3ID3TAG2 tag;
    int lc = 0;

    if ( StartFlag == 1 )
    {
        rewind(f); // set file pointer to beginning of file
        fread(&tag, 1, sizeof(MP3ID3TAG2), f);

        // Check for the tag id3v2 is present at the beginning of file,
        if(tag.tagid[0]=='I' && tag.tagid[1]=='D' && tag.tagid[2]=='3')
        { // if tag id3v2 is present then jump to end of tag
            fseek(f, unpacktagsize(tag), SEEK_CUR);

            printf("\nFile Has Id3Tag2 Present At Beginning");
        }
        else{ // if tag idv3 isn't present then go back to beginning of file
            rewind(f);
        }
    }
}
```

Projects – phase 2

- Code review
 - One other team (parser code, available in IS: `Final_versions_of_the_first_part_TO_EVALUATE`)
 - Open-source project
- https://is.muni.cz/auth/el/1433/podzim2015/PA193/um/59227670/PA193_P3_Projects_second_part_2015.pdf

PROBLEM

What is this device for?



IBM 4758 Hardware Security Module (HSM)

Hardware Security Modules (HSM)

- Hardware Security Modules are high-security devices
 - small security computer inside general purpose computer
 - RAM, CPU, storage...
 - resilience against tampering, side-channels...
 - support various cryptographic operations
 - keys are generated, stored and used directly on the device
 - additional restricted code can be uploaded (firmware)
- HSM exposes its functionality via API **API**
 - E.g., encrypt supplied data with key generated inside HSM
- HSM is trusted, accessed by not-so-trusted applications
 - HSM's API serves as wall between different levels of trust
 - Intentionally limits visibility and access

Security API

API



When we use API?

- All the time 😊
- When using function from standard library
- When using external library
- When calling system (Win32 API, POSIX...)
- When calling methods of our own class
- ...

When we design API?

- Almost all the time 😊
- Function signature is API for this function usage
- List of public methods in interface is its API
- When we create **good** API?
 - good programming habit is to create reusable modules
 - every module has its own API
 - once module will get reused, API *cannot* be changed easily

Application programming interface (API)

- Different types of API

1. Non-security API

- any library API (module/library interface)
- e.g., C++ STL, Boost library, Web API...

2. Cryptographic API

- set of functions for cryptographic operations
- e.g., Microsoft CryptoAPI, OpenSSL API...

3. Security API

- allows untrusted code to access sensitive resources in secure way
- e.g., PKCS#11 HSM module, suExec, OAuth

Language (in)dependent API

- Language **dependent** API
 - API available only for one particular language
 - ABI is relevant (calling convention, memory layout...)
- Language **independent** API
 - Not restricted to particular languages
 - E.g., Web API based on HTTP/REST/JSON
- **Language bindings**
 - Bridge between particular language and library/OS API
 - E.g., library implemented in C, but called from Python
 - Additional API in target language with small proxy code

What is API and ABI?

- API = **A**pplication **P**rogramming **I**nterface
 - source code-based specification intended to be used as an interface between software components to communicate
 - classes, interfaces, methods...
- ABI = **A**pplication **B**inary **I**nterface
 - specification of interface on binary level
 - size, binary representation and layout of data types
 - function calling conventions (stdcall, decl...)
 - how to make system calls (functions outside program memory)
 - binary formats of data produced (little/big endian...)
- API != ABI, but both are necessary

Web API

- Web API = API used to invoke method on web server
 - Usually via HTTP(S) with REST
 - Language independent API
- E.g., Twitter API
 - *POST* <https://api.twitter.com/1.1/statuses/update.json?status=At%20PA193>
- Application programming interface key (API key)
 - Code supplied by program calling an API
 - Identifies program, developer, user...
 - Can be used to control usage (e.g., limit requests...)

Quiz – where is API?

- Language: C
 - API: Functions listed in header files (*.h)
- Language: C++
 - API: public methods of class
 - API: public methods of abstract class
- Language: Java
 - API: public methods of class
 - API: methods of interface
- Twitter Web API
 - API: HTTP/REST requests, response in JSON format



© Martin Handford



PRINCIPLES OF GOOD API

Credits: Joshua Bloch

- *Joshua Bloch, How to Design a Good API and Why it Matters (Google)*
 - <http://lcsd05.cs.tamu.edu/slides/keynote.pdf>
 - video: <http://www.infoq.com/presentations/effective-api-design>
- *Reading/watching is highly recommended*
- *Many ideas taken from his presentation*
 - *demonstrated on cryptographic libraries by myself*

Principles of good API (Joshua Bloch)

1. Easy to learn
 2. Easy to use, even without documentation
 3. Hard to misuse
 4. Easy to read and maintain code that uses it
 5. Sufficiently powerful to satisfy requirements
 6. Easy to extend
 7. Appropriate to audience
- <http://lcsd05.cs.tamu.edu/slides/keynote.pdf>

Process of API design (Joshua Bloch)

1. Gather requirements
2. Start with short specification (1 page)
3. Write API early and often
4. Test and use your API
 - especially when designing SPI (Service Providers Interface)
 - write more plugins (one – NOK, two – difficult, three - OK)
5. Prepare for evolution and mistakes
 - displease everyone equally

What is Service Provider Interface?

javax.crypto

Class CipherSpi

[java.lang.Object](#)

└─ [javax.crypto.CipherSpi](#)

```
public abstract class CipherSpi
extends Object
```

This class defines the *Service Provider Interface (SPI)* for the `Cipher` class. All the abstract methods in this class must be implemented by a particular cipher algorithm.

In order to create an instance of `Cipher`, which encapsulates an instance of this `CipherSpi` class, an application calls one of the [getI](#) methods. Optionally, the application may also specify the name of a provider.

A *transformation* is a string that describes the operation (or set of operations) to be performed on the given input, to produce some output (e.g., `DES`), and may be followed by a feedback mode and padding scheme.

A transformation is of the form:

- "algorithm/mode/padding" or
- "algorithm"

(in the latter case, provider-specific default values for the mode and padding scheme are used). For example, the following is a valid transformation:

```
Cipher c = Cipher.getInstance("DES/CBC/PKCS5Padding");
```

General principles - encapsulation

- API should do one thing and do it well
- As small as possible, but not smaller
 - if in doubt, leave function out (you can add, but not remove)
- Implementation details should not leak into API
 - try to hide as much as possible from user
- Minimalize accessibility (encapsulation)
 - make public what really needs to be
 - no public fields (attributes) except constants
- Make understandable names (self-explanatory, consistent, easy to read when used)

```
if (key.length() < 80)  
    generateAlert("NSA can crack!");
```


General principles - documentation

- Document rigorously
 - JavaDoc, Doxygen...
 - specify how function should be used
 - class: what instance represents
 - Method: contract between method and client
 - preconditions, postconditions, side effects
 - Parameters: who owns (ptr), units, format...
- Specific case of documentation are Annotations
 - e.g., Microsoft SAL, pre&post conditions, Java annotations...

```
/**
 * \brief Output = HMAC-SHA-512( hmac key, input buffer )
 *
 * \param key    HMAC secret key
 * \param keylen length of the HMAC key
 * \param input  buffer holding the data
 * \param ilen   length of the input data
 * \param output HMAC-SHA-384/512 result
 * \param is384  0 = use SHA512, 1 = use SHA384
 */
```

```
void* memcpy(void* destination, const void* source, size_t num);
void* memcpy(__out_bcount(num) void* destination,
             __in_bcount(num) const void* source, size_t num);
```

Which one you like more? Why?

POLARSSL

```
/**
 * \brief      Output = HMAC-SHA-512( hmac key, input buffer )
 *
 * \param key   HMAC secret key
 * \param keylen length of the HMAC key
 * \param input  buffer holding the data
 * \param ilen  length of the input data
 * \param output HMAC-SHA-384/512 result
 * \param is384 0 = use SHA512, 1 = use SHA384
 */
void sha512_hmac( const unsigned char *key, size_t keylen,
                  const unsigned char *input, size_t ilen,
                  unsigned char output[64], int is384 );
```



OPENSSL

```
unsigned char *HMAC(const EVP_MD *evp_md, const void *key, int key_len,
                   const unsigned char *d, size_t n, unsigned char *md,
                   unsigned int *md_len);
```

OpenSSL – HMAC ☹️ (hard to understand)

```
//hmac.h
```

```
unsigned char *HMAC(const EVP_MD *evp_md, const void *key, int key_len,
const unsigned char *d, size_t n, unsigned char *md,
unsigned int *md_len);
```

```
//ossl_typ.h
```

```
typedef struct env_md_st EVP_MD;
```

```
//evp.h
```

```
struct env_md_st
```

```
{
```

```
int type;
```

```
int pkey_type;
```

```
int md_size;
```

```
unsigned long flags;
```

```
int (*init)(EVP_MD_CTX *ctx);
```

```
int (*update)(EVP_MD_CTX *ctx,const void *data,size_t count);
```

```
int (*final)(EVP_MD_CTX *ctx,unsigned char *md);
```

```
int (*copy)(EVP_MD_CTX *to,const EVP_MD_CTX *from);
```

```
int (*cleanup)(EVP_MD_CTX *ctx);
```

```
/* FIXME: prototype these some day */
```

```
int (*sign)(int type, const unsigned char *m, unsigned int m_length,
unsigned char *sigret, unsigned int *siglen, void *key);
```

```
int (*verify)(int type, const unsigned char *m, unsigned int m_length,
const unsigned char *sigbuf, unsigned int siglen,
```

```
....
```

```
} /* EVP_MD */;
```

General principles - performance

- Consider performance impact of API decisions
 - but be not influenced by implementation details
 - underlying performance issues will be fixed eventually, but API warping (for fixing past issue) remains
- Examples of bad performance decisions
 - need for frequent allocations and copy constructors
 - pass arguments by reference or pointer
 - use copy free functions
 - usage of mutable objects instead of immutable
 - use const everywhere possible
 - need for frequent re-coding (`byte[]` -> `string` -> `byte[]`)

Copy-free functions

- API style which minimizes array copy operations
- Frequently used in cryptography
 - we take block, process it and put back
 - can take place inside original memory array
- **int** encrypt(byte array[], **int** startOffset, **int** length);
 - encrypt data from *startOffset* to *startOffset + length*;
- Wrong(?) example:
 - **int** encrypt(byte array[], **int** length, byte outArray[], **int*** pOutLength);
 - note: C/C++ can still use pointers arithmetic
 - note: Java can't (we need to create new array)

Sensitive data (keys) in memory

- What is the difference?

```
int set_key(Key_t key, pin_t seal_pin);
```

vs.

```
int set_key(Key_t* key, pin_t* seal_pin);
```

- Try to limit copies of sensitive data in memory
 - potential unintended disclosure (memory, swap...)
- Pass by value requires more memory erases
 - What about Java's pass by value of reference?

General principles – static factory

- Use static factory instead of class constructor
 - e.g., `javacardx.crypto & class::getInstance()`
 - e.g., `javacardx.crypto & class::buildKey()`

```
import javacardx.crypto.*;

// CREATE DES KEY OBJECT
m_desKey = (DESKey) KeyBuilder.buildKey(KeyBuilder.TYPE_DES,
    KeyBuilder.LENGTH_DES, false);
// CREATE OBJECTS FOR CBC CIPHERING
m_encryptCipher = Cipher.getInstance(Cipher.ALG_DES_CBC_NOPAD, false);
m_decryptCipher = Cipher.getInstance(Cipher.ALG_DES_CBC_NOPAD, false);
// CREATE RANDOM DATA GENERATOR
m_secureRandom = RandomData.getInstance(RandomData.ALG_SECURE_RANDOM);
// CREATE MD5 ENGINE
m_md5 = MessageDigest.getInstance(MessageDigest.ALG_MD5, false);
```

General principles – static factory

- Advantages of static factory over constructors
 - provides named "constructors" (getInstance, buildKey)
 - can return null, if appropriate
 - can return an instance of a derived class, if appropriate
 - reduce verbosity when instantiating variables of generic/template types (no need to write type twice)

```
Map<String, list<String>>* m = new HashMap<String, List<String>>();
```

vs.

```
Map<String, list<String>> m = HashMap.newInstance();
```

- allows immutable classes to use preconstructed instances or to cache instances (speed)
- <http://www.informit.com/articles/article.aspx?p=1216151>

General principles – behave as expected

- Principle of last astonishment
 - user should not be surprised of API behavior
- Be careful with overloading
 - use different names for methods when having same number of arguments
 - same behavior for same (position of) arguments
- Fail fast – report error as soon as possible
 - failure in compile time is better
 - during runtime, first method invocation with bad state should fail
- Provide methods to obtain data elements from results provided originally in strings
 - do not force programmer to parse strings

Example: Avoid long parameter lists

WIN32 API

```

HWND WINAPI CreateWindow(
    _In_opt_ LPCTSTR lpClassName,
    _In_opt_ LPCTSTR lpWindowName,
    _In_     DWORD dwStyle,
    _In_     int x,
    _In_     int y,
    _In_     int nWidth,
    _In_     int nHeight,
    _In_opt_ HWND hWndParent,
    _In_opt_ HMENU hMenu,
    _In_opt_ HINSTANCE hInstance,
    _In_opt_ LPVOID lpParam
);

```



Which one you like more?
Why?

QT API

```

QWidget window;
window.setWindowTitle("Window title");
window.resize(320, 240);
...
window.show();

```

General principles - parameters

- Avoid long parameter lists
 - three or fewer parameters ideal (including default values)
 - mistake in filling arguments might be missed in compile
- When more parameters are required:
 - break method into more methods
 - or encapsulate multiple arguments into single class/struct
- Use consistent parameter ordering (src vs. desc)

```
#include <string.h>
char *strcpy (char* dest, char* src);
void bcopy (void* src, void* dst, int n);
```

Security API

- *“A security API allows untrusted code to access sensitive resources in a secure way.” Graham Steel*
- Interface between different levels of trust
- Security API is designed to enforce a policy
 - certain predefined security properties should always hold
 - e.g., private key cannot be used before user is authenticated
- Security API is not equal to security protocols
 - but closely related
 - security protocol == short program how principals communicate
 - security API == set of short programs called in any order
- http://www.lsv.ens-cachan.fr/~steel/security_APIs_FAQ.html
- <http://www.cl.cam.ac.uk/~rja14/Papers/SEv2-c18.pdf>

Security API attack

- API attack is sequence of commands (function calls) which breach security policy of an interface
- “Pure” API attacks – only sequence of commands
 - e.g., key value is directly revealed
- “Augmented” API attacks – additional brute-force computations required
 - e.g., 128bits key value is exported under 56bits key

Security API – problems in time

- Interfaces get richer and more complex over time
 - pressure from customers to support more options
 - economic pressures towards unsafe defaults
 - failures tend to arise from complexity, **KISS!!!**
 - hard to design secure API, even harder to keep it secure
- Leaks when trusted component talks to less trusted
 - interface often leaks more information than anticipated by designer of trusted component

Security API - typical problems

- Unexpected command sequences
 - methods called in different order than expected
 - use method call fuzzer to test
 - use automata-based programming to verify proper state
- Unknown commands
 - invalid values as method arguments
 - always make extensive input verification
 - use fuzzer to test

Security API - typical problems

- Commands in a wrong device mode
 - sensitive operation (e.g., Sign()) called without previous authentication
 - use methods order fuzzing to test
 - use automata-based programming to ensure proper state
- Existence of undocumented API
 - debugging API not removed (unintentionally)
 - security by obscurity (be aware of reverse engineering)
 - example: Crysalis Luna module (key extraction)
 - <http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-592.pdf>

Security API - typical problems

- Multiple different APIs to single component/storage
 - contracts within one set of API may be broken by second set
 - possible interleaving of function calls from different APIs
 - Example: Ultimaco HSM APIs
 - Microsoft world: CNG, CSP, EKMI
 - JCE, PKCS#11, OpenSSL
 - administration API's: IT monitoring SNMP
 - Example: IBM 4758 HSM APIs
 - IBM CCA, VISA EMV, PKCS#11...
 - IBM proprietary
- Attacks already used in the wild for large scale attacks
 - <http://www.wired.com/threatlevel/2009/04/pins/>

Security API: best practices

- Use API keys, not Username/Password
 - e.g., OAuth instead of Basic Auth
- Don't use sessions (if possible)
 - build API as RESTful services
 - *“Each request from any client contains all of the information necessary to service the request, and any session state is held in the client.” REST Wikipedia*
 - check client input extensively
- Supply methods for secure erase of sensitive data

Security API: best practices

- Always use TLS when secure channel is required
 - or other suitable secure channel, don't build one yourself
- Look at mature APIs for best practice examples
 - Foursquare, Twitter, and Facebook...
- Don't use weak cryptographic algorithms
 - MD5, RC4... Old NSA saying: "Cryptanalysis always gets better. It never gets worse."
- Don't hardcode particular algorithm into API
 - and be prepared for change (e.g., BlockCipher interface instead of AES)

Formal verification of security API

- Harder than security protocol analysis
 - security API typically consist of tens of functions called in any order
 - security protocol only few messages executed in predefined sequence
- Initially applied only to small APIs, now better
- Many interesting practical results
 - real attacks against PKCS#11 devices
 - Ubikey token API problem
 - PKCS#11 RSA's token problem found
- Proofs of security within given model may be given
- http://www.lsv.ens-cachan.fr/~steel/security_APIs_FAQ.html

Formal verification of APIs

- Tookan tool
 - <http://secgroup.ext.dsi.unive.it/projects/security-apis/tookan/>
 - probe PKCS#11 token with multiple function calls
 - automatically create formal model for token
 - run model checker and find attack
 - try to execute attack against real token
- No single “best” tool (Avispa, Proverif...)
- A Generic API for Key Management
 - <http://www.lsv.ens-cachan.fr/~steel/genericapi/>
- International Workshop on Analysis of Security APIs
 - <http://www.lsv.ens-cachan.fr/~steel/asa/>

CODE ANNOTATIONS

Motivation for making annotations

- More semantics of code available for checker
 - Capture otherwise missed bugs
- More explicit documentation of code/API
 - Ideally automatically testable
 - Problems captured in compile time
- Compliancy requirements
 - Driver signature by Microsoft, a must for 64b Windows
- ...

Microsoft SDL C/C++ static checker

- Problems not found by PREfast checker by default
- Achievable via source-code annotation language (SAL)
 - check of return value
 - argument must be not NULL
 - string must be terminated
 - length of data read / written into buffer
- Additional requirements are added to declaration of function, structure... via (non-standard) keywords
- Validity of such requirements are checked by PREfast
 - in pre-state (before fnc call) & in post-state (after fnc call)

SAL – basic terms

- **Element is valid** if contains explicitly assigned value
 - Item of allocated array with unassigned value is invalid
- Valid in *pre-condition* (before function is called)
 - Annotation typically starts with In_xxx
- Valid in *post-condition* (when function ends)
 - Annotation typically starts with Out_xxx
- Number of specified *bytes vs. items*
 - Default is number of items, bytes if `_bytes_` added
 - Number of elements valid

SAL functions basics

Category	Parameter Annotation	Description http://msdn.microsoft.com/en-us/library/hh916383.aspx
Input to called function	<code>_In_</code>	Data is passed to the called function, and is treated as read-only.
Input to called function, and output to caller	<code>_Inout_</code>	Usable data is passed into the function and potentially is modified.
Output to caller	<code>_Out_</code>	The caller only provides space for the called function to write to. The called function writes data into that space.
Output of pointer to caller	<code>_Outptr_</code>	Like Output to caller . The value that's returned by the called function is a pointer.

- Optional version of arguments
 - argument might be NULL
 - `_In_opt`, `_Out_opt`...
 - function must perform check before use

SAL functions basics II.

- Pointer type annotations
- `_Outptr_`
 - should not be NULL
 - should be initialized
- `_Outptr_opt_`
 - can be NULL, must be checked

```
void salDemo(_In_ int* pInArray, _Outptr_ int** ppArray) {  
}  
  
int main(int argc, char* argv[]) {  
    int* pArray = NULL;  
    int* pArray2 = NULL;  
    if (strcmp(argv[1], "alloc") == 0) {pArray = new int[5];}  
    salDemo(pArray, &pArray2);  
  
    return 0;  
}
```

- PREfast analysis

test.cpp(34): warning : C6101: Returning uninitialized memory '*ppArray'. A successful path through the function does **not** set the named _Out_ parameter.

test.cpp(49): warning : C6387: 'pArray' could be '0': **this** does **not** adhere to the specification **for** the function 'salDemo'.

test.cpp(49): warning : C6001: Using uninitialized memory '*pArray'.

SAL annotations – much more

- Annotations of functions
 - <http://msdn.microsoft.com/en-us/library/hh916382.aspx>
- Structs and classes can be also annotated
 - <http://msdn.microsoft.com/en-us/library/jj159528.aspx>
- Locking behavior for concurrency can be annotated
 - <http://msdn.microsoft.com/en-us/library/hh916381.aspx>
- Whole function can be annotated
 - <http://msdn.microsoft.com/en-us/library/jj159529.aspx>
 - `_Must_inspect_result_`
- Best practices
 - <http://msdn.microsoft.com/en-us/library/jj159525.aspx>

SAL – examples (in and out buffer)

```
// read from buffer with size equal to length
int readData( void *buffer, int length );
int readData(_In_reads_(length) void *buffer, int length );

// writes specified amount (length) of data into buffer
int fillData( void *buffer, int *length );
int fillData(_Out_writes_all_(length) void *buffer, const int length );

// writes into buffer maxLength at max, but possibly less and modifies also length argument
int fillData( void *buffer, const int maxLength, int *length );
// Check if no more then maxLength and *length is written, also check range of length
int fillData( _Out_writes_to_( maxLength, *length ) void *buffer,
             const int maxLength, _Out_range_(0, maxLength-1) int *length );

// read AND write from buffer
int readWriteData( void *buffer, int length );
int readWriteData(_Inout_updates_(length) void *buffer, int length );
```

SAL – examples (pointers, strings)

```
// pass argument by value foo pointer
int getInfo( struct thing *thingPtr );
// value is used as input and output => _Inout_
int getInfo( _Inout_ struct thing *thingPtr );

// pass C null-terminated strings
int writeString( const char *string );
// must be null terminated string > _In_z_
int writeString( _In_z_ const char *string );
```

Using SAL with / without PREfast

- Cross-platform code
 - Compiled with MSVC for Windows (SAL is supported)
 - Compiled with GCC for Linux (SAL not supported)
- Issue: SAL annotations makes GCC compilation to fail
- Solution
 - Create custom `#define` for most common SAL annotations
 - Define as empty if not compiled with MSVC
 - Can be also tuned when SAL annotation changes itself
- Peter Gutmann' Experiences with SAL/PREfast
 - <http://www.cs.auckland.ac.nz/~pgut001/pubs/sal.html>

Wrapping defines for SAL

```
#if defined( _MSC_VER ) && defined( _PREFAST_ )
    #define IN_BUFSIZE                _In_reads
    #define IN_BUFSIZE_OPT            _In_reads_opt
    #define OUT_BUFSIZE               _Out_writes
    #define OUT_BUFSIZE_OPT           _Out_writes_opt
    #define OUT_PTR                   _Outptr_
    #define OUT_PTR_OPT               _Outptr_opt_
#else
    #define IN_BUFSIZE( size )
    #define IN_BUFSIZE_OPT( size )
    #define OUT_BUFSIZE( max, size )
    #define OUT_BUFSIZE_OPT( max, size )
    #define OUT_PTR
    #define OUT_PTR_OPT
#endif /* VC++ with source analysis enabled */
```

Modification of <http://www.cs.auckland.ac.nz/~pgut001/pubs/sal.html>

Annotations for GCC/LLVM

- Deputy
 - Not active any more ☹, last update 2006?
 - <http://www.stanford.edu/class/cs295/asgns/asgn5/www/>
- Clang Static Analyzer
 - <http://clang-analyzer.llvm.org/annotations.html>
 - Only few annotations

Splint (is simple to use?)

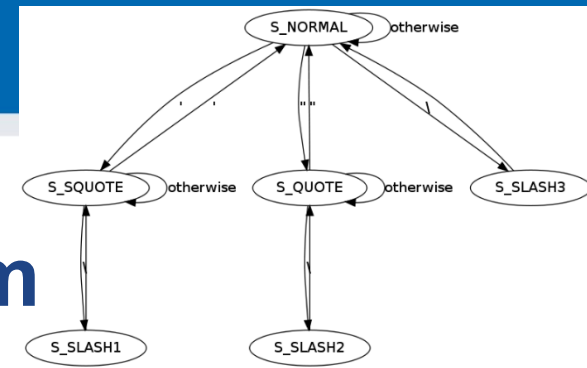
- **SAL version**

```
void strcpy( _Out_z char *s1, _In_z const char *s2 );
```

- **Splint version**

```
void /*@alt char * @*/ strcpy(  
  /*@unique@*/ /*@out@*/ /*@returned@*/ char *s1, char *s2 )  
  /*@modifies *s1@*/ /*@requires maxSet( s1 ) >= maxRead( s2 ) @*/  
  /*@ensures maxRead( s1 ) == maxRead( s2 ) @*/;
```

AUTOMATA-BASED PROGRAMMING



Automata-based style program

- Program (or its part) is thought of as a model of finite state machine (FSM)
- Basic principles
 - Automata state (explicit designation of FSM state)
 - Automata step (transition between FSM states)
 - Explicit state transition table (not all transitions are allowed)
- Practical implementation
 - imperative implementation (switch over states)
 - object-oriented implementation (encapsulates complexity)
- https://en.wikipedia.org/wiki/Automata-based_programming

Example: SimpleSign applet

- Simple smart card applet for digital signature
 - Operation 1: user must verify UserPIN before private key usage for signature is allowed
 - Operation 2: unblock of user pin allowed only after successful AdminPIN verification
- Imperative solution:
 - sensitive operation (Sign()) is wrapped into condition testing successful PIN verification
 - more conditions may be required (PIN and < 5 signatures)
 - same signature operation may be called from different contexts (SignHash(), ComputeHashAndSign())

SimpleSign –imperative solution

```
void SignData(APDU apdu) {  
    // ...  
    // INIT WITH PRIVATE KEY  
    if (m_userPIN.isValidated()) {  
        // INIT WITH PRIVATE KEY  
        m_sign.init(m_privateKey, Signature.MODE_SIGN);  
  
        // SIGN INCOMING BUFFER  
        signLen = m_sign.sign(apdubuf, ISO7816.OFFSET_CDATA,  
                               (byte) dataLen, m_ramArray, (byte) 0);  
  
        // ... SEND OUTGOING BUFFER  
    }  
    else ISOException.throwIt(SW_SECURITY_STATUS_NOT_SATISFIED);  
  
    // ...  
}
```

Test of required condition

Execution of sensitive operation

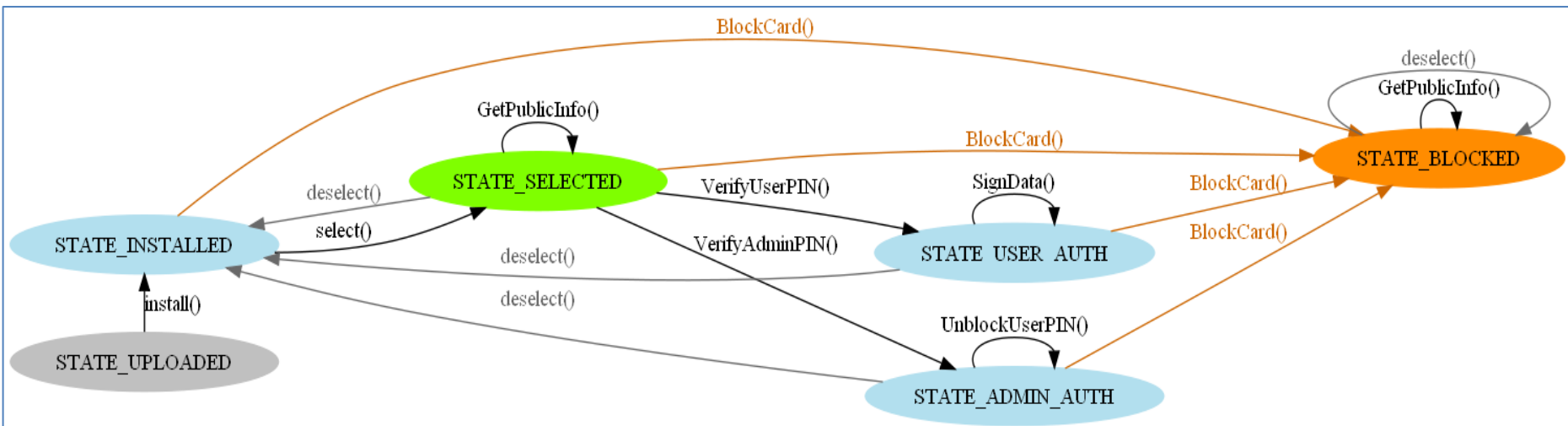
Example: states for smart card applet

```

digraph StateModel {
rankdir=LR;
size="6,6";
node [shape =ellipse color=green, style=filled];
{ rank=same; "STATE_UPLOADED";"STATE_INSTALLED";}
"STATE_INSTALLED";
"STATE_UPLOADED";
"STATE_UPLOADED" -> "STATE_INSTALLED" [label="install()"];
{ rank=same; "STATE_SELECTED";}
"STATE_SELECTED";
{ rank=same;"STATE_USER_AUTH";"STATE_ADMIN_AUTH";}
"STATE_USER_AUTH";
"STATE_ADMIN_AUTH";

"STATE_INSTALLED" -> "STATE_SELECTED" [label="select()" color="black" fontcolor="black"];
"STATE_SELECTED" -> "STATE_USER_AUTH" [label="VerifyUserPIN()" color="black" fontcolor="black"];
"STATE_SELECTED" -> "STATE_ADMIN_AUTH" [label="VerifyAdminPIN()" color="black" fontcolor="black"];
...

```



SimpleSign – automata-based solution

1. Mental model → .dot format (human readable)
 - Graphviz visualization (visual inspection)
 - source code generated for state check and transition check
 - input for formal verification (state reachability)
2. Easy to extend by new states
 - source code is generated again
3. More robust against programming mistakes and omissions

Is transition allowed between given states?

- E.g., is allowed to change state directly from STATE_INITIALIZED to STATE_ADMIN_AUTH?

```
private void SetStateTransition(short newState) throws Exception {
    // CHECK IF TRANSITION IS ALLOWED
    switch (m_currentState) {
        case STATE_UPLOADED: {
            if (newState == STATE_INSTALLED) {m_currentState = STATE_INSTALLED; break;}
            throw new Exception();
        }
        case STATE_INSTALLED: {
            if (newState == STATE_SELECTED) {m_currentState = STATE_SELECTED; break;}
            if (newState == STATE_BLOCKED) {m_currentState = STATE_BLOCKED; break;}
            throw new Exception();
        }
        case STATE_SELECTED: {
            if (newState == STATE_SELECTED) {m_currentState = STATE_SELECTED; break;}
            if (newState == STATE_USER_AUTH) {m_currentState = STATE_USER_AUTH; break;}
            if (newState == STATE_ADMIN_AUTH) {m_currentState = STATE_ADMIN_AUTH; break;}
            if (newState == STATE_BLOCKED) {m_currentState = STATE_BLOCKED; break;}
            if (newState == STATE_INSTALLED) {m_currentState = STATE_INSTALLED; break;}
            throw new Exception();
        }
    }
}
```

Is function call allowed in present state?

- E.g., do not allow to use private key before UserPIN was verified

```
private void checkAllowedFunction(int requestedFunction) {
    switch (requestedFunction) {
        case FUNCTION_VerifyUserPIN:
            if (m_currentState == STATE_SELECTED) break;
            _OperationException(EXCEPTION_FUNCTIONEXECUTION_DENIED);

        case FUNCTION_SignData:
            if (m_currentState == STATE_USER_AUTH) break;
            _OperationException(EXCEPTION_FUNCTIONEXECUTION_DENIED);
        ...
    }
}
```

Sign data only when in
STATE_USER_AUTH

How to react on incorrect state transition

- Depends on particular application
 - create error log entry
 - throw exception
 - terminate process
 - ...
- Error message should not reveal too much
 - side-channel attack based on error content

SimpleSign – additional functionality

- New functionality is now required:
 - signature allowed also after verification of AdminPIN
- Changes required in imperative solution:
 - add additional condition before every Sign()
 - when called from multiple places, developer may forgot to include conditions everywhere
 - not easy to realize, what conditions are required from existing code
- Changes required in automata-based solution:
 - add new state transition (STATE_ADMIN_AUTH <-> SignData())
 - generate new transition tables etc.

SUMMARY

Summary

- Designing good API is hard
 - follow best practices, learn from well-established APIs
- Designing security API is even harder
- Automata-based programming
 - make more robust state and transition validation
 - good to combine with visualization and automatic code generation

Questions 