**Reverse engineering** is the process of discovering the technological principles of a device, object or system through analysis of its structure, function and operation. It often involves taking something (e.g., a mechanical device, electronic component, or software program) apart and analyzing its workings in detail to be used in maintenance, or to try to make a new device or program that does the same thing without using or simply duplicating (without understanding) any part of the original.

Reverse engineering has its origins in the analysis of hardware for commercial or military advantage. The purpose is to deduce design decisions from end products with little or no additional knowledge about the procedures involved in the original production. The same techniques are subsequently being researched for application to legacy software systems, not for industrial or defence ends, but rather to replace incorrect, incomplete, or otherwise unavailable documentation.

**Reverse engineering of software**

Reverse engineering is taking apart an object to see how it works in order to duplicate or enhance the object. The practice, taken from older industries, is now frequently used on computer hardware and software. Software reverse engineering involves reversing a program's machine code (the string of 0s and 1s that are sent to the logic processor) back into the source code that it was written in, using program language statements.

Software reverse engineering is done to retrieve the source code of a program because the source code was lost, to study how the program performs certain operations, to improve the performance of a program, to fix a bug (correct an error in the program when the source code is not available), to identify malicious content in a program such as a virus or to adapt a program written for use with one microprocessor for use with another. Reverse engineering for the purpose of copying or duplicating programs may constitute a copyright violation. In some cases, the licensed use of software specifically prohibits reverse engineering.

Someone doing reverse engineering on software may use several tools to disassemble a program. One tool is a hexadecimal dumper, which prints or displays the binary numbers of a program in hexadecimal format (which is easier to read than a binary format). By knowing the bit patterns that represent the processor instructions as well as the instruction lengths, the reverse engineer can identify certain portions of a program to see how they work. Another common tool is the disassembler. The disassembler reads the binary code and then displays each executable instruction in text form. A disassembler cannot tell the difference between an executable instruction and the data used by the program so a debugger is used, which allows the disassembler to avoid disassembling the data portions of a program. These tools might be used by a cracker to modify code and gain entry to a computer system or cause other harm.

Hardware reverse engineering involves taking apart a device to see how it works. For example, if a processor manufacturer wants to see how a competitor's processor works, they can purchase a competitor's processor, disassemble it, and then make a processor similar to it. However, this process is illegal in many countries. In general, hardware reverse engineering requires a great deal of expertise and is quite expensive.
Another type of reverse engineering involves producing 3-D images of manufactured parts when a blueprint is not available in order to remanufacture the part. To reverse engineer a part, the part is measured by a coordinate measuring machine (CMM). As it is measured, a 3-D wire frame image is generated and displayed on a monitor. After the measuring is complete, the wire frame image is dimensioned. Any part can be reverse engineered using these methods.

The term reverse engineering as applied to software means different things to different people, prompting Chikofsky and Cross to write a paper researching the various uses and defining a taxonomy. From their paper, they state, "Reverse engineering is the process of analyzing a subject system to create representations of the system at a higher level of abstraction."[5] It can also be seen as "going backwards through the development cycle".[5] In this model, the output of the implementation phase (in source code form) is reverse-engineered back to the analysis phase, in an inversion of the traditional waterfall model. Reverse engineering is a process of examination only: the software system under consideration is not modified (which would make it re-engineering). Software anti-tamper technology is used to deter both reverse engineering and re-engineering of proprietary software and software-powered systems. In practice, two main types of reverse engineering emerge. In the first case, source code is already available for the software, but higher-level aspects of the program, perhaps poorly documented or documented but no longer valid, are discovered. In the second case, there is no source code available for the software, and any efforts towards discovering one possible source code for the software are regarded as reverse engineering. This second usage of the term is the one most people are familiar with. Reverse engineering of software can make use of the clean room design technique to avoid copyright infringement.

On a related note, black box testing in software engineering has a lot in common with reverse engineering. The tester usually has the API, but their goals are to find bugs and undocumented features by bashing the product from outside.

Other purposes of reverse engineering include security auditing, removal of copy protection ("cracking"), circumvention of access restrictions often present in consumer electronics, customization of embedded systems (such as engine management systems), in-house repairs or retrofits, enabling of additional features on low-cost "crippled" hardware (such as some graphics card chip-sets), or even mere satisfaction of curiosity.

The Certified Reverse Engineering Analyst (CREA) is a certification provided by the IACRB that certifies candidates are proficient in reverse engineering software.

[edit] Binary software

This process is sometimes termed Reverse Code Engineering, or RCE.[5] As an example, decompilation of binaries for the Java platform can be accomplished using Jad. One famous case of reverse engineering was the first non-IBM implementation of the PC BIOS which launched the historic IBM PC compatible industry that has been the overwhelmingly dominant computer hardware platform for many years. An example of a group that reverse-engineers software for
enjoyment is CORE which stands for "Challenge Of Reverse Engineering". Reverse engineering of software is protected in the U.S. by the fair use exception in copyright law.[7] The Samba software, which allows systems that are not running Microsoft Windows systems to share files with systems that are, is a classic example of software reverse engineering.[8] since the Samba project had to reverse-engineer unpublished information about how Windows file sharing worked, so that non-Windows computers could emulate it. The Wine project does the same thing for the Windows API, and OpenOffice.org is one party doing this for the Microsoft Office file formats. The ReactOS project is even more ambitious in its goals, as it strives to provide binary (ABI and API) compatibility with the current Windows OSes of the NT branch, allowing software and drivers written for Windows to run on a clean-room reverse-engineered GPL free software or open-source counterpart.

[edit] Binary software techniques

Reverse engineering of software can be accomplished by various methods. The three main groups of software reverse engineering are

1. Analysis through observation of information exchange, most prevalent in protocol reverse engineering, which involves using bus analyzers and packet sniffers, for example, for accessing a computer bus or computer network connection and revealing the traffic data thereon. Bus or network behaviour can then be analyzed to produce a stand-alone implementation that mimics that behaviour. This is especially useful for reverse engineering device drivers. Sometimes, reverse engineering on embedded systems is greatly assisted by tools deliberately introduced by the manufacturer, such as JTAG ports or other debugging means. In Microsoft Windows, low-level debuggers such as SoftICE are popular.

2. Disassembly using a disassembler, meaning the raw machine language of the program is read and understood in its own terms, only with the aid of machine-language mnemonics. This works on any computer program but can take quite some time, especially for someone not used to machine code. The Interactive Disassembler is a particularly popular tool.

3. Decompilation using a decompiler, a process that tries, with varying results, to recreate the source code in some high-level language for a program only available in machine code or bytecode.