The objective of this thesis is to design and implement a tool usable for detecting similar code in different projects. The tool should be able to locate code pasted from one project to another and should be able to cope with average attempts to thwart the detection such as symbol renaming, changing the order of unrelated entities, moving entities to different files, adding or removing comments, etc. The tool is implemented in language C++ and is ready to compare source files written in languages C and C++. The tool also enables the comparison of source code written in different languages, which can be compiled by the GNU C Compiler. To obtain good results in these cases, new modules should be added (this is necessitated due to different representations of the GNU C Compiler inner form for different languages).

The first part of this thesis focuses on describing the problem domain, the architecture design and the tools usable for implementation. The second part centers on the implemented solution, a description of data structures and possibilities for application expansion using additional modules. The last part of the thesis sums up the results and outlines future possibilities of implementation.