

ABSTRAK

Kemudahan dalam pertukaran informasi dan perkembangan teknologi informasi sering disalahgunakan oleh sekelompok orang. Penyalahgunaan itu terjadi di lingkungan akademis dan dilakukan oleh siswa/mahasiswa. Salah satu tindakan penyalahgunaan teknologi informasi adalah mencontek dan menyalin tugas/evaluasi yang diberikan tenaga pendidik. Tindakan tersebut merupakan salah satu plagiarisme yang dilakukan oleh siswa. Metode untuk mengantisipasi tindakan tersebut adalah dengan membandingkan tugas siswa secara mandiri. Dalam dunia komputer, evaluasi pemahaman pemrograman sangat berpengaruh terhadap tingkat pemahaman siswa sebagai referensi untuk memahami bahasa pemrograman lain. Dibutuhkan pemahaman dalam bahasa pemrograman untuk menilai kemiripan dari kode program yang ditulis. Penelitian ini menjadi salah satu solusi bagaimana agar dapat mendeteksi kemiripan kode program dengan terkomputerisasi. Sistem pendeteksi kemiripan kode program bahasa C ini bersifat membantu mendeteksi tindakan plagiat dengan memberikan nilai kemiripan antar dua dokumen kode program bahasa C yang dibandingkan. Perubahan variabel, gaya penulisan yang dilakukan siswa guna mencegah deteksi aksi mencontek bisa dicegah dengan adanya *textpreprocessing* dan tokenisasi yang kemudian dilakukan perhitungan dengan algoritma Smith-Waterman. Algoritma Smith-Waterman adalah salah satu algoritma yang bisa diimplementasikan untuk membandingkan dokumen kode program.

Kata Kunci: plagiarisme, bahasa C, *textpreprocessing*, tokenisasi, Smith-Waterman.

ABSTRACT

The ease of information exchange and the rapid growth of information technology is often misused by a group of people. The abuse occurs in academic environment and done by group of students. One of misused of information technology is cheating and copying tasks/evaluations provided by educators. That action is one of the plagiarism done by students. The method to anticipate such action is by comparing the students' tasks independently. In the computer environment the evaluation of programming understanding is very influential on the level of student understanding and very important as a reference to understand other programming languages. It takes an understanding in the programming language to judge the similarity of written program code. This research became one of the solutions how to be able to detect resemblance of program code with computerization. This C language similarity detection system is helpful in detecting plagiarism by giving similarity value between two compared written program code in C language. Changes in variables, writing style that students do to prevent the detection of cheating action can be prevented by the existence of the text preprocessing and tokenization which then performed calculations with Smith-Waterman algorithm. The Smith-Waterman algorithm is one of the method that can be implemented to compare program code documents.

Keywords: plagiarisms, C language, text preprocessing, tokenization, Smith-Waterman.