Similarities in source codes Marek Roštár Supervisor: Ing. Michal Kompan, PhD.

	Objectives:	Method:
	 Problem of rapid increase of popularity of different programming languages 	 Based on a
		Comparisor
	 With this increase comes an increase of plagiarism of source codes 	Source co
	 Propose and implement method for finding similarities in different source codes across multiple programming languages 	Currently 2
		 First just
		 Second r
	 Future Work Adding support for multiple programming languages Adding multiple different levels of abstraction Creating universal parser 	

int sucet(int a,int b){	int sucet(variable,variable){ return variable+variable:
}	}
void main(){	 void main(){
int a; int b=5	variable;
a=6;	variable=value;
printf("%d\n",sucet(a,b));	printf(value,sucet(variable,va
}	}



- bstract syntax trees
- n of using multiple levels of source code abstraction

ode abstraction:

- levels:
- prepares source code for parsing
- removes variables and their values

Evaluation:

- Comparison of results between different source code abstraction levels
- Comparison with multiple existing tools
- Multiple datasets:
 - Fabricated short source codes
 - Dataset for source code plagiarism
 - Dataset from students submissions _

