COMPUTER SCIENCE TRIPOS Part II – 2015 – Paper 7

11 Optimising Compilers (AM)

This question concerns intraprocedural analysis of a flowgraph-style program P whose instructions are in three-address form and labelled $S_1, S_2, \ldots$, with S_1 labelling the first instruction to be executed. Input parameters to the program are simulated by the first few instructions of P being of the form `x = read()`, and the result given by a `return` instruction.

An instruction T is a *semantic reaching definition* at instruction U if, for some execution starting at S_1, instruction T writes to a variable x which does not suffer an intervening assignment when execution reaches instruction U. We write $RD(i)$ for the set of instructions S_j whose definitions reach instruction S_i.

(*a*) By analogy with live variable analysis or available expression analysis, derive dataflow equations for RD and give an algorithm for solving these. Explain any approximation you make, carefully justifying the form of this approximation. [*Hint*: you may find it useful to define *gen* and *kill* for instructions.] [8 marks]

(*b*) Is your analysis for reaching definitions flow-sensitive or flow-insensitive? Give a one-sentence justification of your answer. [2 marks]

(*c*) One use of reaching definitions is for constant propagation: when we know that reading a variable in an operand in a given instruction will always result in the same value k, we may replace the operand with k. Carefully explain how we can use the result of reaching-definitions analysis to perform constant propagation. [*Hint*: you may find it useful to consider the instruction form `z:=x+y`.] [3 marks]

(*d*) Explain how your constant-propagation algorithm would react to the following flowgraph expressed as C code:

```
int t,r,x;
x = read();
if (x>91) t=7; else t=6;
r = t/2;
return r+39;
```

Either explain why your resulting code is optimal, or indicate the source of any information loss which precludes it being optimal. [3 marks]

(*e*) Suppose now the 3-address code were in SSA (single static assignment) form. How would this affect the result of reaching-definitions analysis? [4 marks]