MPhil in Advanced Computer Science

L111: Advanced Data Flow Analysis

- Lecturers: Alan Mycroft and Uday P. Khedker
- Half Module: 6 Lectures (one revision lecture of 60 minutes followed by five core lectures of 90 minutes each).
- Motivation:

  Data flow analysis is a technique of discovering useful information from programs. The applications of this technique range from compiler optimization to software engineering to software verification.

  This module highlights the common abstractions across a large number of data flow analysis problems and covers common techniques of performing data analysis. It shows how the simpler bit vector data flow frameworks and the more advanced frameworks such as pointer and heap analyses are special cases of a general model of data flow frameworks. Time permitting, we will also cover the design and implementation of the *Generic Data Flow Analyser* (gdfa) in GCC (Gnu Compiler Collection).
- Takeaways:

  After attending this course a student will be able to

  - Understand how a carefully chosen representation and a small set of operations facilitate discovering useful properties of programs.
  - Formulate data flow frameworks for new problems.
  - Perform data flow analysis on sample programs.
- Syllabus:

  - *Lecture 0:* Review of background knowledge: Control-Flow Graph, Lattice, Fixed Point Iteration, Liveness, Dataflow Equation, Safety of Analysis.
  - *Lecture 1:* Motivation, Bit vector data flow analysis (Available Expression Analysis and Live Variable Analysis), Round robin iterative method of data flow analysis.
  - *Lecture 2:* Partial Redundancy Elimination, Constant Propagation.
  - *Lecture 3:* Lattice theoretic formulation of data flow frameworks, MFP and MoP assignments, Complexity of iterative data flow analysis, Introduction to GDFA (Generic Data Flow Analyzer in GCC).

- *Lecture 4:* General data flow frameworks: Pointer Analysis, Heap Liveness Analysis.
- *Lecture 5:* Introduction to interprocedural data flow analysis: functional approach, call strings based approach.

- Suggested Reading:
  - Uday P. Khedker, Amitabha Sanyal, and Bageshri Karkare. Data Flow Analysis: Theory and Practice. CRC Press (Taylor and Francis Group), USA, 2009.
  - The Generic Data Flow Analyzer. http://www.cse.iitb.ac.in/grc/gdfa.html
  - Hecht, Matthew S. Flow Analysis of Computer Programs. Elsevier North-Holland Inc. 1977.
  - Flemming Nielson, Hanne Riis Nielson, Chris Hankin: Principles of Program Analysis. Springer, 2005.
  - Alfred V. Aho, Monica S. Lam, Ravi Sethi, and Jeffrey D. Ullman. Compilers: Principles, Techniques, and Tools (2nd Edition), Addison Wesley, 2006.
  - Stephen S. Muchnick and Neil D. Jones (Ed.). Program Flow Analysis : Theory and Applications. Prentice-Hall, 1981.