Reliable Post-link Optimizations Based on Partial Information

Ealan Henis
Gadi Haber
Vadim Eisenberg

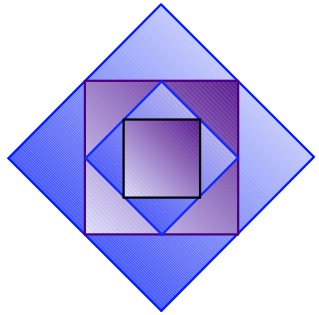
presented by

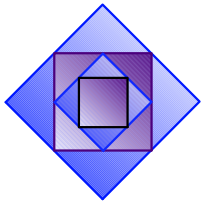
Moshe Klausner

Code Optimization Group

FDDO3, December 2000

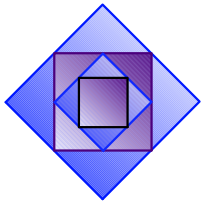
IBM Research Lab in Haifa





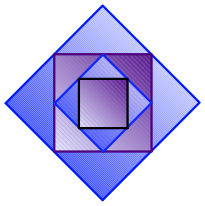
Plan of the Talk

- Introduction
- Goal
- Method of Analysis
- Experiance
- Conclusions



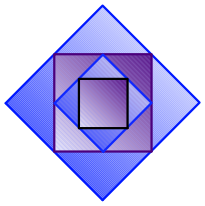
Feedback Based Post-Link Optimizations

- Complement the compiler's optimizations
- Optimize frequently traversed paths at the expense of the infrequent ones
- Global optimizations
- Compiler and linker conventions need not be preserved



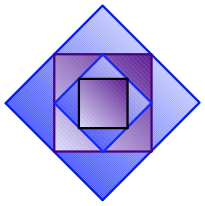
Feedback Based Post-Link Optimizations (cont.)

- **Phases**
  - Analysis --> Instrument --> Profile --> Optimize
- **References**
  - ATOM [Srivastava and Eustace 94]
  - EEL [Larus et al. 95]
  - ETCH [Romer et al. 96]
  - SPIKE [Cohn et al. 97]
  - ALTO [Muth et al. 99]
  - FDPR [Henis et al. 99]
  - .....



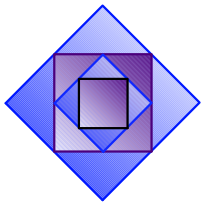
IBM FDPR Tool

- **Feedback Directed Program Restructuring**
  [Heisch 94, Nahshon and Bernstein 96, Henis et al. 99]
  - Dissects into BBs
  - Optimizations:
    - AOPT: repositions BBs,
      requires full information, **more** powerful
    - HeatShrink: repositions BB copies,
      accepts incomplete info, **less** powerful



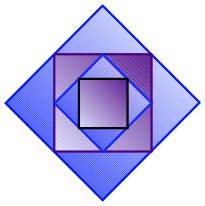
Goal

- Need of an algorithm that enables AOPT to work in incomplete information situation
- Maximize optimization by using hybrid of AOPT and HeatShrink



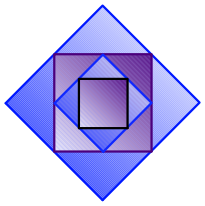
Incomplete information

- **Types of incomplete information:**
  - Basic block could not be classified as code or data
  - Data could not be classified as values or addresses
  - CFG is incomplete (statically unresolved targets)
- **Fully analyzed executable is rarely available**
- **Yet, in most cases only small portion of the information is missing**



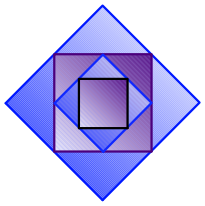
Handling incomplete information

- All BBs marked as analyzed or as Not Fully Analyzed (NFA)
- Confine NFAs within larger Partially Analyzed Areas (PAA)
- Analyzed areas treated using AOPT
- Apply special algorithms to PAAs



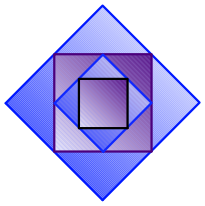
PAA

- **PAA properties and Prerequisite**
  1. Relative references within the PAA (that do not have relocation information) remain semantically correct when PAA repositioned en-bloc
  2. Absolute references from within the PAA are identifiable even for the NFA parts within the PAA
  3. Linker/Loader relocation information available
- **PAA confines the problematic nature of its NFAs**



PAA (cont.)

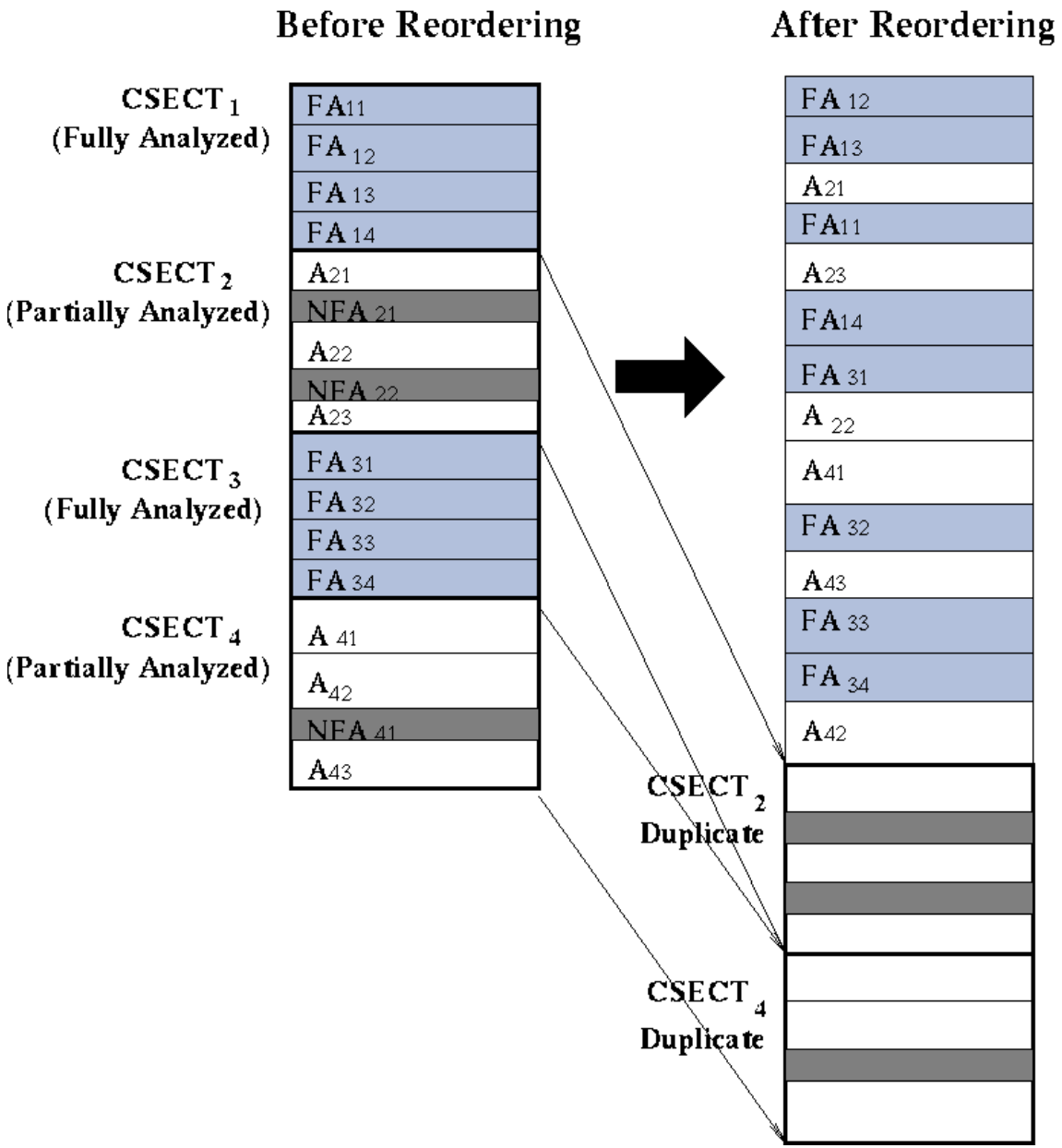
- **In AIX/PowerPC**
  - CSECT
  - Non-stripped relinkable executable
- **In NT/x86**
  - Object module
  - The executable must contain a relocation section
  - Linker map is available

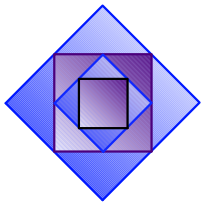


Local Movable HeatShrink Algorithm (LMH)

- **For each NFA define a containing PAA (One PAA may contain several NFAs)**
- **Perform AOPT on fully analyzed areas and on copies of analyzed portions of the PAAs**
  - Assuming the PAA is well behaved
- **Reposition PAAs en-bloc**
- **Update the references to each PAA in the entire executable**
- **Update the addresses in each PAA according to its new position**

LMH algorithm (cont.)

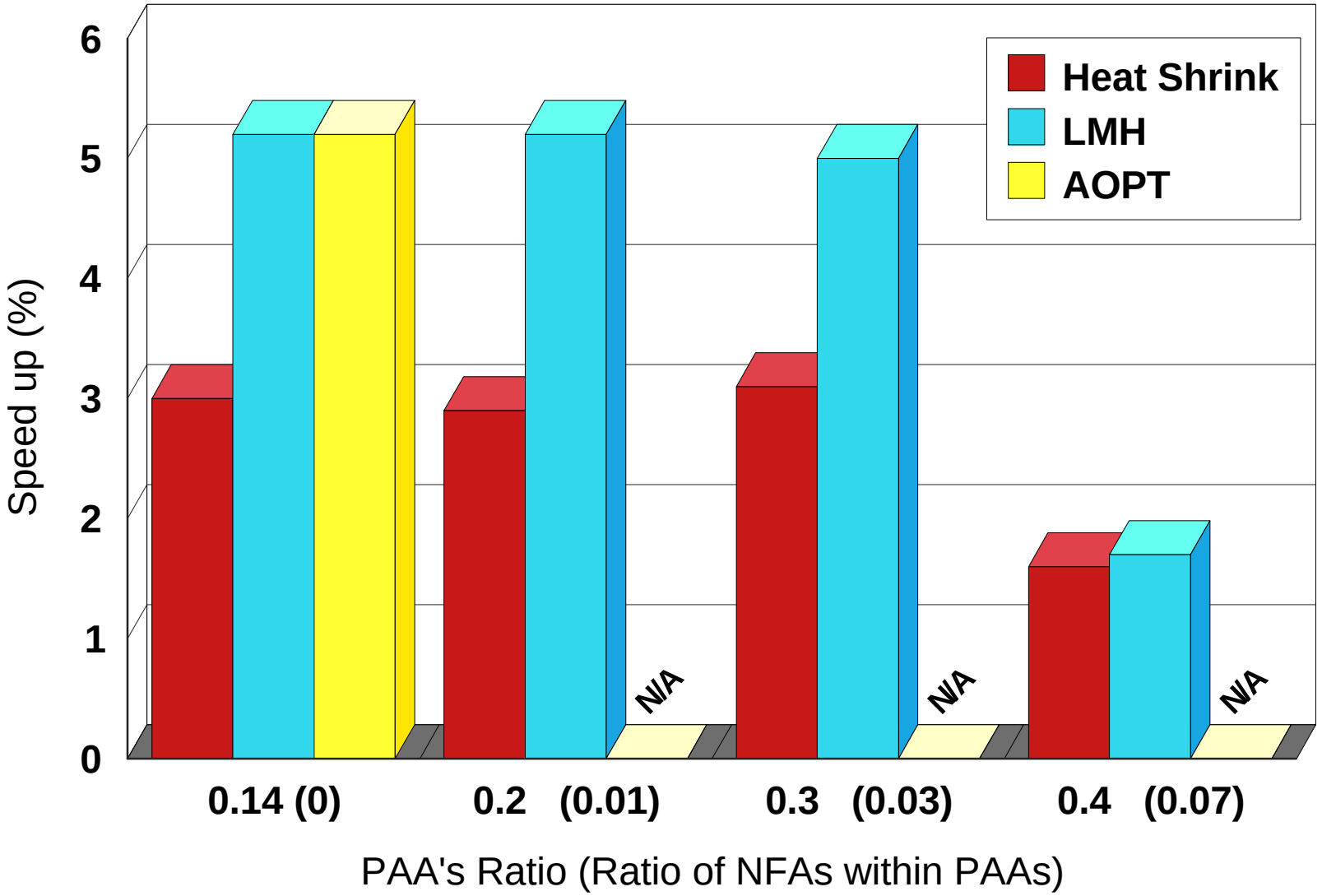




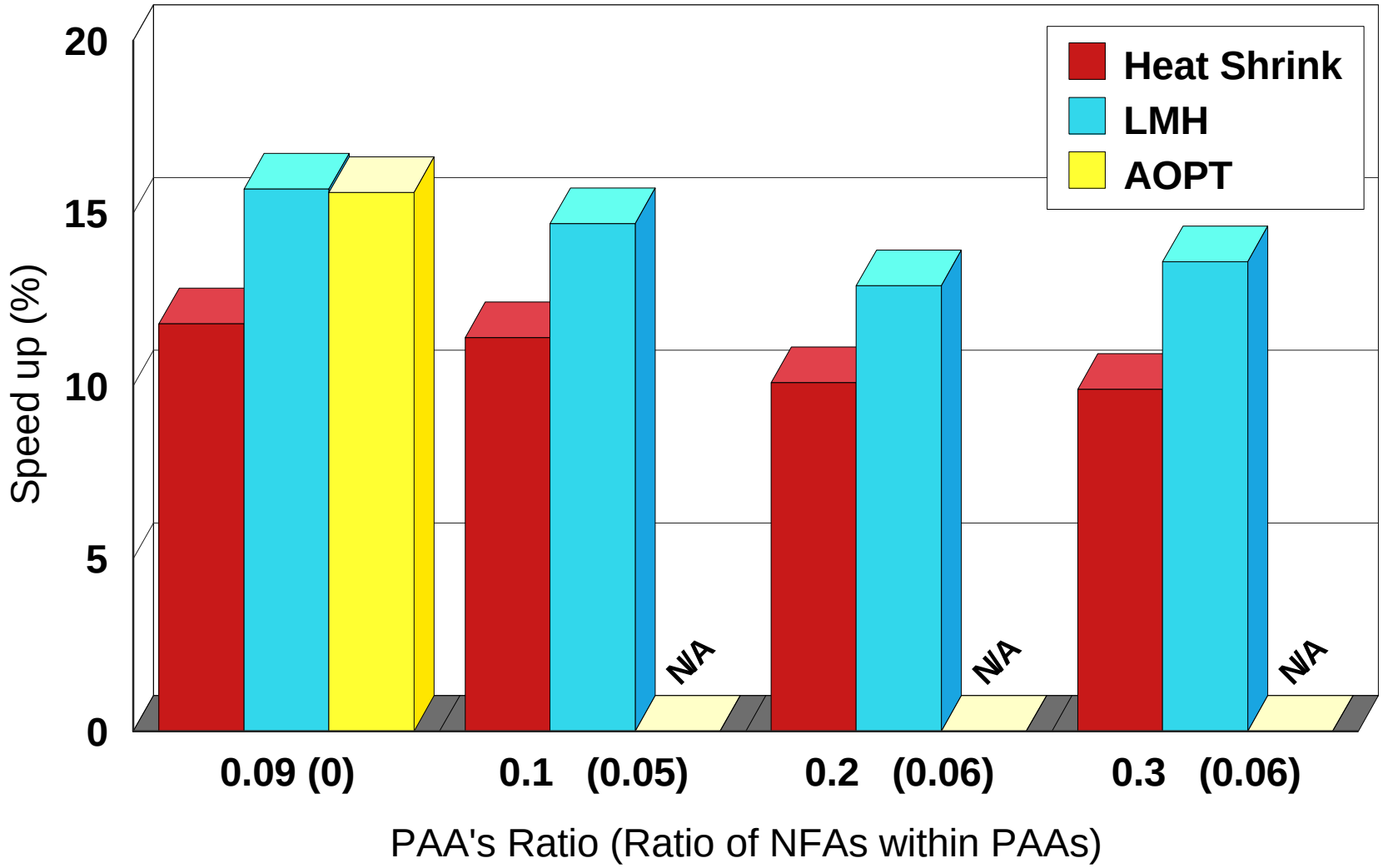
Experiments Framework

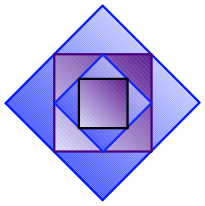
- **Applied to programs with varying amount of missing information**
- **Full information available**
  - Equivalent to AOPT
- **Artificialy information is withheld**
  - Gradual increase of PAAs from zero (full information) to 0.4
  - Up to 7% NFAs within each PAA
- **Typical case in real life applications**
  - In agreement with [Larus and Schnarr 95]
- **Done on AIX/PowerPC**

Performance Gains for perl



Performance Gains for gcc





Conclusions

- **Novel concept to enable combination of AOPT and HeatShrink**
  - Allows to better exploit available information
  - Improve performance gain due to more powerful optimization method (AOPT)
- **Reliable approach for optimization in a partial information situations**
- **Gradual degradation of performance from AOPT to HeatShrink**
  - Significant improvement over a step drop from AOPT to HeatShrink
  - In real life: only small portion of information is missing