Trace analysis and correlation

Isnaldo Francisco de Melo Jr
May 3, 2016

École Polytechnique de Montréal
Laboratoire DORSAL
Outline

● Introduction
● Overview
● Data Structure
● Trace Compare
● Analyzing metrics/properties
  How can we use it for comparisons
● Current Work
● Conclusion
● Future directions
Introduction

Research goals

Correlate and abstract data from a trace in a meaningful way.

Help identifying from the traces the differences among groups of runs
Concepts Overview

Profiling techniques
  e.g. JVMPI

Call Graph
  Gprof profiling with CG

Calling Context Tree
Comparative

Call Graph

Calling Context
Basis of the research

This research is based on:

- Francois work in Trace Compare
- Previous work:
  - Trevis, pattern mining with tracers, trace abstraction
Trace Compare

- Comparing variations of performance using groups of traces

- Basically it highlights delays that appear comparing two groups: normal x slow executions.

- Limitation the developer should try to eliminate these delays manually → expert
Differential Flame Graph

“There is no time here”
Similar tool

Trevis

Trevis is a tree visualization and analysis framework that allows users to visualize, compare, cluster, and intersect context trees, such as calling context trees produced by profilers.
Analysis

Analyzing the compare tool?
So considering this tools, what kind of analysis can we do?
Schema for analysis

- Apply some controlled scenarios and build some techniques or models measuring its characteristics properly.

- Idea: compare two situations and correlate the data using mathematical tools

- Be able to point limitations of the software performance with minimal human assistance
Schema for analysis

Comparing two executions

Start

Task in scenario 1

Task in scenario 2

Compare
Analyzing metrics

Summary of the applied techniques

SVM

Clustering techniques

Analysis of variance module
Analyzing metrics

Support Vector Machine Application:

Introduction
What are the pros and cons
Results so far
Support vector machine

Demonstration of SVM: separating in hyperplane
Analyzing metrics

Agglomerative Clustering

Introduction

What are the pros and cons

Results so far
Agglomerative Clustering

Demonstration with 4 groups
Analyzing metrics

Using Analyze of variance

Introduction

Using ANOVA

What are the pros and cons

Results so far
Anova

Comparing variations of scenarios

Scenario 1

Compare variation within this group

Scenario 2

Compare the variation between groups
Current work

Adapting the models for a more realist approach → real data

Replacing Trace Compare with an integrated Trace compass module
Conclusion

It is possible to correlate some high level information with certain restrictions.

It is mandatory however, to build a problem solver module.

From now on: focus on real data and solve real problems.
- Compare the performance with other similar tools
- Apply the techniques on real issues
- Find graph properties of CCT
Questions

Isnaldo-francisco.de-melo-jr@poly.ca

github.com/FranciscoMeloJr

Any other info?
References

Trevis
  sape.inf.usi.ch/trevis

CCT

PCCT
Obrigado