Union Schemas as a Basis for a C++ Extractor

Thomas Dean
Queen's University
Kingston, Canada

Andrew Malton, Ric Holt
University of Waterloo
Waterloo, Canada
What is CPPX?

- Open Source C++ Extractor
- Minimal Resources
- Minimal Maintenance
- Standard Exchange Format
- Datrix

WCRE October 5, 2001, Stuttgart, Germany
What is CPPX?

• Complete Semantic Facts
• Production Quality **
• Source Complete **
• Performance
Implementation

- Semantic Analysis
- Scanning & Parsing
- Preprocessing (inlining)

GNU C++ compiler provides:

• Based on GNU C++ Compiler
Graph Representations

- GCC representation is a graph
  - GCC graph is for code generation
  - Extra nodes for machine
  - Machine dependent details

- Datrix representation is also a graph
  - More abstract representation
  - Machine independent
• CPPX is a transformation from a GNU C++ graph to a Datrix Graph.

Many small, mostly independent graph transformations—Each targets particular parts of the graph.

CPPX is a transformation from a graph to a Datrix Graph.
Transformations

Graph

Matrix

In

... ...

T1

Graph

GNU

g++

Include

Source

CPPX Structure
- Intermediate graphs have a mixed representation.
- Transformations should be semantic preserving.
- How do we know the result is correct?

Graph Transformations

WCRE October 5, 2001, Stuttgart Germany
Union Schemas

- Combine GCC and Datrix Schema
  - Find node types that represent common concepts (namespace, function, statement)
  - Find relations that represent common concepts (decls, ArcSon)
  - Similar to union grammars for tree transformation (TXL [Cordy])
  - Find relations that represent common concepts (namespace, statement)
  - Find node types that represent common concepts (namespace, function, statement)

Combining GCC and Datrix Schemas
- Text output to CXL, TA, and VCG
- Binary vector representation (g)
- CPPX common representation (g)
- Informally documented GNU schema
- Translators provide specification for concepts
- Mapping between common

Union Schemas
CPPX Status

- http://swag.uwaterloo.ca/~cppx

- PBS now using CPPX as an extractor

• Many small transforms ideal for distributed development (open source)
• CPPX is real

• Looking for volunteers!!

• CPPX Conclusion