

Внутреннее представление Reprise распараллеливающей системы

В данной работе представлен проект основы программной реализации Открытой распараллеливающей системы [1] – внутреннее представление Reprise.

Открытая распараллеливающая система (ОПС) предназначена для автоматического распараллеливания программ с процедурных языков программирования на параллельные компьютеры. С помощью ОПС могут быть разработаны: распараллеливающие компиляторы, распараллеливающие конвертеры (в языки высокого уровня), диалоговые распараллеливающие системы и другие компилирующие продукты [1]. Предполагается, что преобразования из библиотеки ОПС можно будет комбинировать для различных архитектур: конвейерных, векторных, SIMD, MIMD, VLIW, DataFlow, мультитредовых, программируемых... В данный момент в разработке конвертор с языка Си в HDL, конвертеры Си в Си с MPI и OpenMP. Информация об ОПС частично представлена в Интернете: www.ops.rsu.ru

Внутреннее представление (ВП) – это структура данных для хранения информации о программе, а также алгоритмы, облегчающие выполнение преобразований программ [2]. Внутреннее представление Reprise должно быть:

- универсальным, т.е. позволять хранить программы процедурных языков (Си и Фортран). С поддержкой объектно-ориентированных конструкций (характерных для C++);
- многоуровневым. Разные уровни отличаются детализацией исходной программы;
- удобным для описания преобразований (распараллеливающих, оптимизирующих);
- унифицированным; Похожие свойства языков должны быть скрыты за единым интерфейсом, а специфические сохранены для этапа выполнения преобразований;
- расширяемым для подключения компиляторов переднего плана с новых, в том числе, объектно-ориентированных языков программирования к ОПС;
- удобным для построения конвертеров (обеспечивать восстановление из ВП программ в виде, близком изначально написанным пользователем);
- удобным для генерации машинного кода, кода с вызовом OpenMP и MPI.

Внутреннее представление спроектировано с применением Design Patterns [3] и реализовано на языке C++.

При проектировании внутреннего представления необходимо было создать сбалансированное решение. С одной стороны, представление должно вмещать все возможные конструкции входных языков и максимально сохранить присущие им особенности. С другой стороны, необходимо минимизировать количество типов узлов в дереве ВП для удобства выполнения преобразований. Например, известно, что цикл FOR языка Си в общем виде распараллелить практически невозможно. Однако программы на Си содержат множество простых циклов, в стиле цикла DO языка Фортран. Во ВП цикл FOR входного языка отображается на универсальный узел, представляющий удобный для распараллеливания цикл FOR. При этом выполняется автоматически канонизация цикла и в дальнейшем проверки на вид цикла FOR в преобразованиях можно опустить. Во внутреннем представлении также поддерживается канонизация выражений, таких как операторы языка Си: ++, --, += и т.п. Они преобразовываются для использования простых операторов: +, -, =. Подобным образом, канонизированы другие конструкции входных языков.

К сожалению, при канонизации теряется вид исходной программы пользователя. Внутреннее представление содержит специальную метаинформацию, которая добавляется во ВП на этапе работы компилятора переднего плана. Метаинформация позволяет практически полностью восстановить исходную программу пользователя. Это важная возможность для систем полуавтоматического и диалогового распараллеливания, в которых пользователь может проанализировать информационные зависимости в своей программе, а также указать системе особенности программы, которые могут быть полезны при выполнении распараллеливающих преобразований.

Внутреннее представление Reprise является древовидной структурой данных, типы узлов которой делятся на пять групп: узлы для представления типов данных, идентификаторов, операторов, выражений и служебные узлы. Для манипулирования узлами предназначена специальная библиотека, которая тоже является частью ВП. В отличие от библиотеки сервисных функций, существовавшей в предыдущих версиях ОПС [1], в том, что библиотека обеспечивает целостность операций над ВП. Осуществляет контроль за корректностью модифицированного ВП. Библиотека является не просто набором вспомогательных алгоритмов для работы с ВП, она полностью инкапсулирует функции по изменению ВП. Какие-либо существенные изменения ВП можно осуществить только при помощи обращения к этой библиотеке.

Внутреннее представление является многоуровневым. Например, специальный уровень представления выражений, с информацией о потоке управления, вхождениях переменных и коэффициентах в индексных выражениях массивов. Этот уровень используется для построения графов информационных зависимостей в программе, в том числе при построении решетчатого графа [4]. Также отдельный уровень ВП используется для преобразований в пределах функции. Этот уровень скрывает сложность характерную для объектно-ориентированной программы и приводит ВП к виду процедурных языков, что значительно упрощает реализацию преобразований.

В процессе работы над внутренним представлением были изучены внутренние представления наиболее известных распараллеливающих систем: Polaris [5] и SUIF [6], а также семейства компиляторов GNU (GCC).

Внутреннее представление Polaris ориентировано на один язык программирования – Фортран 77. И хотя основные конструкции этого языка удобны для распараллеливания, на них очень трудно отобразить программы с языка Си. В Polaris отсутствует понятие блок операторов, не предусмотрены пометки узлов внутреннего представления. Кроме того, ВП Polaris неразрывно связано с графом информационных связей, который синхронно меняется вместе с изменениями во ВП.

Внутреннее представление SUIF рассчитано на языки программирования Си и Фортран 77. Однако поддержка языка Фортран изначально в систему не закладывалась, программа с него при помощи конвертера f2c переводится в Си, а только потому во внутренний формат. Очевидно, при таком подходе, специфичная для программ на Фортране информация никак не отражается во внутреннем представлении.

GCC – семейство компиляторов с различных языков программирования (C, C++, Fortran, Ada, Java, Objective-C) на множество целевых архитектур (Alpha, ARM, IA-32, IA-64, MIPS и т.д.) [7].

В GCC используется три вида внутренних представлений: GENERIC, GIMPLE и RTL. Представление GENERIC является высокоуровневым представлением, предназначенным для унификации входных языков. В текущей версии GCC большинство оптимизирующих преобразований выполняется на следующем уровне ВП GIMPLE. Этот уровень является упрощенным подмножеством представления GENERIC. В GIMPLE выражения преобразованы в трехадресный код, и структуры управления понижены до операторов IF и GOTO. Представление RTL является низкоуровневым и предназначено для генерации кода.

Во внутреннем представлении Reprise выбран другой подход: все высокоуровневые конструкции входных языков переводятся в узлы ВП. Узлы спроектированы таким образом, что содержат высокоуровневые конструкции языков программирования, удобные для распараллеливания. Если какие-то высокоуровневые конструкции не приводятся к этому виду, они приводятся в наиболее близкие по смыслу конструкции языка.

Литература:

1. Штейнберг Б.Я., Нис З.Я., Петренко В.В., Черданцев Д.Н., Штейнберг Р.Б., Шульженко А.М. Открытая распараллеливающая система 2006. // Труды третьей международной конференции «Параллельные вычисления и задачи управления» РАСО '2006. Москва, 2-4 октября 2006 г.
2. Штейнберг Б.Я., Макошенко Д.В., Черданцев Д.Н., Шульженко А.М. Внутреннее представление в Открытой распараллеливающей системе. // Искусственный интеллект. Научно-теоретический журнал. Институт проблем искусственного интеллекта НАНУ. Украина, Донецк: ДонДИШИ, «Наука и Освита». 2003. № 4. С. 89-97.
3. Гамма Э., Хэлм Р., Джексон Р., Влиссидес Дж. Приемы объектно-ориентированного проектирования. Паттерны проектирования. СПб: Питер, 2001. 368 с.
4. Шульженко А.М. Решетчатый граф и использующие его преобразования программ в Открытой распараллеливающей системе // Научный сервис в сети Интернет: технологии распределенных вычислений. Труды Всероссийской научной конференции. Новороссийск, 19-24 сентября 2005. М.: Изд-во МГУ, 2005. С. 82-85.
5. Faigin K., Hoeflinger J., Padua D., Petersen P., Weatherford S. The Polaris Internal Representation.
6. Aigner G., Diwan A., Heine D., Lam M., Moorey D., Murphy B., Sapuntzakis C. The SUIF Program Representation.
7. <http://gcc.gnu.org/>

Внутреннее представление Kergise распараллеливающей системы, Петренко Виктор Владимирович, Южный
Федеральный Университет, Россия, 344006, г. Ростов-на-Дону, ул. Б. Садовая, 105 +7(863)2633158,
vpetrenko@gmail.com .