

Анализ последовательных программ с помощью средств УБТ

Н.А. Катаев

Московский Государственный Университет имени М.В. Ломоносова, Институт Прикладной Математики имени М.В. Келдыша РАН

Исследование проводится в рамках Системы Автоматизированного Распараллеливания Фортран Программ (САПФОР) [1], разрабатываемой в Институте Прикладной Математики им. Келдыша РАН. САПФОР относится к диалоговым распараллеливающим системам.

В рамках проводимого исследования предполагается разработать анализатор последовательных программ, написанных на языках Fortran и C/C++. На вход анализатора поступает последовательная программа (далее “программа”), возможно содержащая в тексте специальные аннотации, добавленные пользователем и уточняющие свойства программы. На выходе должна быть построена база данных САПФОР, содержащая структуру программы и информацию о зависимостях по данным в каждом цикле исходной программы, необходимые для построения параллельной версии программы. В качестве основного средства анализа используется Универсальная Библиотека Трансляции (УБТ)[2].

Прежде чем поступить на вход анализатора, внутреннее представление программы должно быть сохранено на xml-файл с помощью модифицированного компилятора GCC [2]. Далее средствами УБТ происходит загрузка данного представления в анализатор, выполняется анализ программы, строится база данных САПФОР. Анализ программ выполняется на основе внутреннего представления GIMPLE компилятора GCC. Единство данного представления для различных языков программирования позволяет создать единую систему анализа программ, написанных на языках Fortran и C/C++. Средства УБТ позволяют обнаруживать циклы, реализованные в исходной программе с помощью оператора GOTO, а низкий уровень операторов GIMPLE позволяет статически более точно оценить время выполнения витков цикла.

В силу нацеленности базы данных САПФОР на высокоуровневые языки программирования (для удобства пользователя), использование низкоуровневого представления влечет за собой определенные сложности. Одному оператору исходной программы в общем случае соответствует группа операторов GIMPLE. Для восстановления исходного вида оператора используется граф потока данных программы, строящийся УБТ. Для корректного распределения данных в языке Fortran DVM (C DVM) необходимо знать размерности линейризованных в GIMPLE представлений массивов. Особенности представления COMMON-блоков языка Fortran и необходимость анализа наличия тесной вложенности циклов добавляют дополнительные трудности.

Анализатор был испытан на тестах Якоби и SOR, для которых существующие редуционные зависимости были заданы с помощью аннотаций в тексте программы. Была выполнена ручная проверка корректности баз данных. Полученные базы данных были поданы на вход эксперта САПФОР, построившего параллельные версии указанных тестов.

Проводимое исследование показывает возможность использования анализатора на основе УБТ в качестве компонента САПФОР. Использование GIMPLE представления компилятора GCC в качестве основы анализа позволяет расширить применимость САПФОР для языков Fortran 95 и C/C++. Дополнительно, разрабатываемые в процессе исследования средства анализа GIMPLE представления могут быть использованы для восстановления исходного кода программы по GIMPLE представлению.

Выражаю благодарность В.Е. Владиславлеву за помощь в освоении УБТ.

Литература

1. Система Автоматизированной Параллелизации Фортран Программ.
URL: <http://www.keldysh.ru/dvm/SAPFOR/> (дата обращения: 02.01.2011).
2. Optimizing Technologies.
URL: <http://www.optimitech.com/> (дата обращения: 07.01.2011).